IMPLICATIONS OF SUPPORTIVE AND STRUCTURED TEACHING FOR STUDENT AND TEACHER BEHAVIOR IN K-8 CLASSROOMS SERVING PREDOMINANTLY BLACK STUDENTS

AN ABSTRACT

SUBMITTED ON THE FOURTEENTH DAY OF APRIL 2020

TO THE DEPARTMENT OF PSYCHOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

Elizabeth M. McIntyre, M.S.

Elizabeth McIntyre

Approved: Stacy Overstreet, Ph.D.

Stacy Overstreet, Ph.D.
Chair

Courtney N. Baker, Ph.D.

Courtney N. Baker, Ph.D.

Michael Cunningham, Ph.D.

Michael Cunningham, Ph.D.

Lisa Molix, Ph.D.

Lisa Molix, Ph.D.

Abstract

Positive school climate is associated with a broad range of student benefits across diverse student populations. However, Black[1] students often report less positive perceptions of school climate than their peers, which could lead to decreased school engagement. One important aspect of school climate, teacher-student relationships, may promote positive student outcomes such as engagement. Specifically, supportive and structured teacher behaviors are associated with greater student engagement, which in turn may predict ongoing supportive and structured teaching, but current literature examining these associations underrepresents Black students and their teachers. The current study was completed in elementary and middle school classrooms in four New Orleans public charter schools with majority Black student populations. It was hypothesized that across grade levels and levels of teacher experience in education, supportive and structured teaching at the beginning of the year would be positively associated with student engagement at mid-year, and that mid-year student engagement would positively mediate the association between supportive and structured teaching at the beginning of the year and at the end of the year. Results supported the hypothesized association between beginning-of-the-year supportive and structured teaching and mid-year student engagement. However, results did not support the hypothesized association between beginning-of-the-year supportive and structured teaching and end-of-year supportive and structured teaching, nor was there a mediation effect of mid-year student engagement. The current findings extend the literature promoting supportive and structured teaching as an effective tool for student engagement to include classrooms with predominantly Black students. Results and implications are discussed in the context

of supportive and structured teaching, student engagement, and the broader construct of school climate.

[1]The author of the current study notes the diversity of racial and ethnic backgrounds represented by individuals identifying as Black. As this racial group includes individuals with origins from any of the black racial and ethnic groups of the world, including those from African countries and the Caribbean, the current document adopts the terminology of Kena and colleagues (2015) and uses the term Black as an inclusive term representing these diverse groups.

IMPLICATIONS OF SUPPORTIVE AND STRUCTURED TEACHING FOR STUDENT AND TEACHER BEHAVIOR IN K-8 CLASSROOMS SERVING PREDOMINANTLY BLACK STUDENTS

A DISSERTATION
SUBMITTED ON THE FOURTEENTH DAY OF APRIL 2020
TO THE DEPARTMENT OF PSYCHOLOGY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF SCIENCE AND ENGINEERING
OF TULANE UNIVERSITY
FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
BY

Elizabeth M. McIntyre, M.S.

Elizabeth McIntyre

Approved: Stacy Overstreet, Ph.D.

Stacy Overstreet, Ph.D.
Chair

Courtney N. Baker, Ph.D.

Courtney N. Baker, Ph.D.

Michael Cunningham, Ph.D.

Michael Cunningham, Ph.D.

Lisa Molix, Ph.D.

Lisa Molix, Ph.D.

Table of Contents

List of Tables

List of Figures

Introduction

In April 2007, the National Center for Learning and Citizenship, the Education Commission of the United States, and the Center for Social and Emotional Education came together to generate a comprehensive definition of school climate. This commission endeavored to define school climate beyond imprecise terms such as "the relatively enduring quality of the school environment" (Hoy, 1990) or the school's "openness" (Halpin & Croft, 1963) that prevailed in the school climate literature. The commission defined school climate as,

> "…norms, goals, values, interpersonal relationships, teaching, learning and leadership practices, and organizational structures. A sustainable, positive school climate fosters youth development and learning necessary for a productive, contributing, and satisfying life in a democratic society. This climate includes norms, values, and expectations that support people feeling socially, emotionally, and physically safe. People are engaged and respected. Students, families, and educators work together to develop, live, and contribute to a shared school vision. Educators model and nurture attitudes that emphasize the benefits and satisfactions gained from learning. Each person contributes to the operations of the school and the care of the physical environment" (National School Climate Council, 2007, p. 5).

Accordingly, there seems to be consensus that the core dimensions of school climate represent the major ecologies and functions in the school, which include instruction, interactions between and within groups of individuals in various roles (i.e., students, teachers, administrators, and families), physical facilities, and school-wide norms, beliefs, and values. The National School Climate Center (2012) has further operationalized these features into five dimensions: Safety, Teaching and Learning, Interpersonal Relationships, Institutional Environment, and a Staff Only domain. *Safety* includes rules and norms, physical safety, and emotional safety of all school stakeholders. The *teaching and learning* dimension emphasizes the academic as well as psychosocial

aspects of instruction and learning, including support for learning and growth in understanding of social responsibility. *Interpersonal relationships* indicate the presence of quality, respectful relationships among stakeholders. *Institutional environment* emphasizes features of the physical school environment as well as stakeholder experiences of connectedness and engagement. Finally, the *Staff Only* dimension indicates transparency and effectiveness of leadership, as well as the level of morale, relationships, and collaboration that occurs among school staff. Several researchers (e.g., Cohen, Mccabe, Michelli, & Pickeral, 2009; Thapa, Cohen, Guffy, & Higgins-D'Alessandro, 2013; Wang & Degol, 2016) identified similar categories of school climate in their respective reviews of the school climate literature.

Positive school climate has broad benefits for students, including increased school satisfaction, student socio-emotional well-being, student engagement, academic motivation and achievement, as well as decreased disruptive behaviors, peer victimization, absenteeism, and suspensions (deJung & Duckworth, 1986; Gregory et al., 2010; Hoge, Smit, & Hanson, 1990; Way, Reddy, & Rhodes, 2007; Wilson, Pianta, & Stuhlman, 2007). In a sample of 50,000 high school students across urban, suburban, and rural settings, Konold, Cornell, Shukla, and Huang (2017) examined the associations of Black, White, and Hispanic/Latino students' school climate reports with their school engagement and personal experiences of peer victimization. Across all students, positive perceptions of school climate predicted increased engagement and reduced victimization, with no moderating impact of student racial or ethnic background. Such results indicate the importance of positive school climate across diverse student populations.Yet, student perceptions of school climate often vary by student race. Konold and colleagues (2017)

found that students identifying as Black or African American reported significantly lower school climate ratings than White and Hispanic or Latino students. These disparities in school climate perceptions can even be evident within the same school (Voight, Hanson, O'Malley, & Adekanye, 2015). Furthermore, these discrepancies in school climate perceptions are associated with disparate student outcomes. For example, in the United States, Black students experience exponentially higher suspension rates than students from any other racial or ethnic group (Losen, 2011). Bottiani, Bradshaw, and Mendelson (2017) evaluated school climate predictors of this racial discipline gap among 20,000 adolescent students. They found that across genders, grade levels, and socioeconomic levels, Black students consistently reported lower levels of school equity and personal belongingness than their White peers, and these reports predicted Black students' greater risk for suspension.

Increased awareness of these disparities has incited scrutiny of the features of Black students' school experiences that may contribute to their less positive perceptions of school climate (Penuel, Meyer, & Valladares, 2016). Teacher-student relationships may be one element impacting Black students' experiences. For example, Spilt and Hughes (2015) found that elementary school teachers are more likely to report conflict with Black students in grades 1-5 than with Hispanic and White students. Such findings have engendered questions about teacher approaches that foster knowing, respecting, and engaging with students from diverse backgrounds and are indicative of positive teacher-student relationships.

Teacher behaviors toward students are one critical element of teacher-student relationships that can promote or attenuate positive student perceptions of school climate.

Supportive and structured teacher behaviors represent promising practices for promoting students' positive school experiences (Gregory & Weinstein, 2008; Walker, 2009; Wentzel 2002). Supportive and structured teaching has theoretical grounding in the concept of authoritative teaching, drawn from Baumrind's authoritative parenting style and empirically supported in studies that have primarily drawn from White study populations (Dever & Karabenick, 2011), as well as in the warm demander teaching approach, which applies supportive and structured teaching in the context of diverse student populations (Ford & Sassi, 2014; Irvine & Fraser, 1998; Nelson, 2016; Ware, 2006). Ford and Sassi (2014), Nelson (2016), and Ware (2006) identified supportive and structured teaching behaviors as a promising approach for facilitating Black students' equitable experiences of school climate. Their results warrant systematic evaluation of teacher support and structure in predominantly Black classrooms and the impact of this approach on student outcomes. Specifically, prior research suggests that teacher support and structure are positively associated with student engagement in majority White student samples (Pitzer & Skinner, 2017; Skinner, Furrer, Marchand, & Kindermann, 2008; Wang & Eccles, 2013; Wilson, Pianta, & Stuhlman, 2007). The current study evaluated this relationship in classrooms serving a predominantly Black student population. It assessed observable teacher behaviors indicative of supportive and structured teaching and examined their relationship to student engagement across grades K-8. Furthermore, longitudinal work by Pelletier, Seguin-Levesque, and Legault (2002) and Pitzer and Skinner (2017) showed that the association between supportive and structured teaching and engagement can have lasting positive impacts for students and teachers later in the year, as student engagement predicted teacher maintenance of supportive and structured

behaviors. The current study evaluated this pathway in classrooms with predominantly Black students across grades K-8, testing the association between supportive and structured teaching and student engagement early in the school year, and whether student engagement is associated with the stability of supportive and structured teaching later in the year.

Teacher-Student Relationships as Key Aspects of School Climate

A strong theoretical foundation establishes the Interpersonal Relationships domain of school climate as one important area necessitating further study. Studies adopting this focus prioritize students' relationships with teachers or peers as formative of their school climate perceptions. Ecological systems theory, attachment theory, social motivational theories, and cultural difference theory build a rationale for emphasizing teacher-student relationships in the study of school climate. Two frameworks that apply attachment theory to teaching, authoritative teaching and the warm demander pedagogy, operationalize teacher support and structure as formative for positive teacher-student relationships in the classroom.

First, *relationships in the classroom are proximal to the day-to-day experience of the student.* Per Ecological Systems Theory (Bronfenbrenner, 1977), the environments surrounding the individual weigh greatly on how the individual views, functions in, and develops in them. The student plays a distinct role in each of his or her settings, and each setting is contained within larger ecologies. This nesting facilitates transactions of influences across ecological systems' levels. School climate studies that explore teacher-student relationships emphasize the relationships that are most ecologically proximal to the experience of each student.

Second, *quality teacher-student relationships meet students' needs for autonomy, competence, and relatedness.* The self-system theory of motivation, or self-determination theory (SDT), holds that students have fundamental psychological needs for autonomy, competence, and a sense of belonging and connection with others, or relatedness (Deci, Vallerand, Pelletier, & Ryan, 1991). The most effective teaching practices promote learning in students' individual zones of proximal development by operating within the *relational zone*. The "relational zone" supports students' needs for autonomy, competence, and relatedness by integrating sensitive and nurturing interactions with consistent, predictable structures (Pitzer & Skinner, 2017).

Third, *secure adult attachments influence student perceptions of security and promote healthy exploration and development.* Attachment theory describes the secure bond between a child and his or her caregiver as the central mechanism encouraging the child to seek out and engage in learning experiences that promote development. Secure attachments promote a positive sense of self, inclination to independently explore and engage in the environment, emotional intelligence, and perseverance. Conversely, children experiencing insecure attachments may struggle to develop in these areas (Ainsworth, Blehar, Waters, & Wall, 1978; Bowlby, 1969).

Baumrind (1968) proposed four parenting styles that influence the security of a caregiver-child attachment bond. Each style is each unique in its combination of parental responsiveness and support with parental demandingness and expectations. *Authoritative* parenting reflects high degrees of responsiveness and demandingness; *authoritarian* parenting is characterized by low responsiveness and high demandingness; *permissive* parenting demonstrates high responsiveness and low demandingness; and *neglectful*

parenting is low in both domains. Though attachment theory most explicitly references parenting approaches in early child development, this concept underlies an extended attachment perspective for teachers and school-aged children (Pianta & Steinberg, 1992). This application of attachment theory uses Baumrind's four parenting styles as a framework for measuring the various teacher behaviors present in teacher-student relationships that promote secure attachments. In Baumrind's initial studies examining each parenting style (1966; 1967; 1968), authoritative parenting demonstrated the strongest association with secure attachments among preschool children, which Baumrind attributed to the balance of high support and high structure characteristic of authoritative parenting. This balance fostered secure attachment through setting clear guidelines for behavior while encouraging and reinforcing the child for meeting those expectations (Baumrind, 1966).

Authoritative parenting (Baumrind, 1968) underlies two key concepts commonly measured in teacher-student relationships: teacher *support* and teacher *structure.* Accordingly, authoritative school climate theory emphasizes understanding teacher-student interactions through assessing characteristics of teacher *support*, also termed responsivity, warmth, or connectedness, and teacher *structure*, also termed order, control, or demandingness and representing the teacher's strict, consistent, and fair expectations and practices (Gregory & Cornell, 2009). While teacher support and structure do not represent all the aspects of teacher-student relationships, these constructs are referenced in the National School Climate Center's (2012) Interpersonal Relationships dimension: Support from Adults. This dimension encompasses both support ("patterns of supportive and caring relationships for students") and structure ("high expectations for students'

success"). The warm demander teaching framework likewise emphasizes teacher support and structure and proposes that these constructs are effective across diverse student populations when defined in the context of students' racial and ethnic backgrounds. The warm demander approach, which originated as an approach for culturally responsive pedagogy among Alaska Native students (Kleinfeld, 1975) and eventually extended to urban, low-income, African American student populations (Vasquez, 1989), similarly combines structure, demonstrated through high academic demands and consistent, clear classroom systems (comprising "structure"), with flexible, culturally-specific operationalizations of teacher behaviors representing responsivity to students' needs ("support").

These theories characterize teacher-student relationships as a primary mechanism by which students interact with their school environment and thus experience school climate. Furthermore, the implications for teacher-student relationships span most other main dimensions of school climate, including Teaching and Learning (i.e., support for learning), Safety (i.e., emotional safety), and Institutional Environment (i.e., facilitators of connectedness/engagement), positioning teacher-student relationships as an important area of focus within school climate research. Authoritative teaching and the warm demander pedagogy are two frameworks that highlight teacher support and structure toward students as key aspects of the teacher-student interactions contributing to students' overall perceptions of school climate.

Operationalizing Teacher-Student Interactions: Teacher Support and Structure

Through both the authoritative teaching and warm demander frameworks, teacher support and structure are critical for positive teacher-student relationships and promoting

positive student outcomes. Authoritative teaching holds that the primary indicators of teacher support include anticipating and responding to students' behavioral and academic needs, providing supportive, constructive, and immediate feedback on performance, and communicating openly with students; this framework has been largely applied in samples of predominantly White samples (Wentzel, 2002). Wilson, Pianta, and Stuhlman (2007) described support as teachers' sensitivity in their affective and instructional interactions with students, including offering child-centered activities and choices, providing actionable evaluative feedback, and responding to students' needs. Structure represents behaviors that consistently present and reinforce expectations for student behavior and performance. These behaviors include supervision, monitoring, proactive rule-setting and rule enforcement through classroom behavior management systems, and selective, constructive criticism. Quantitative studies have largely applied this framework among samples of predominantly White students. For example, Wentzel assessed these characteristics in middle school classrooms where 87% of students identified as White. Similarly, Wilson, Pianta, and Stuhlman (2007) assessed the prevalence and impact of these supportive characteristics in 820 first-grade classrooms, 75% of whose students were White.

Similar conceptualizations of support and structure have been highlighted in qualitative work examining teachers' warm demander approaches among Black students. Ware (2006) conducted a qualitative case study of the teaching approaches employed by two Black teachers in Atlanta elementary and middle schools serving primarily Black students from low- and middle-income families. Through interviews and classroom observations, students identified that these two teachers were effective because they

simultaneously maintained high expectations and high responsivity to student needs. Similarly, another qualitative study conducted an ethnography of eighth-grade Black boys' school experiences of teacher-student relationships (Nelson, 2016). Themes in student responses showed strong support for a collaborative working style with clear and consistent expectations and frequent expressions of caring from teachers, and students reported feeling engaged in school because of these behaviors. This qualitative work provides a promising foundation for systematically evaluating supportive and structured teaching in elementary and secondary classrooms of predominantly Black students.

Consistent with the theoretical basis for authoritative approaches, the warm demander approach advances a secure teacher-student relationship through a balance of responsivity and demandingness. However, the warm demander framework also allows for the definitions of support and structure to include teacher behaviors appropriate to the culture of teachers and students in a classroom. Some of these behaviors have historically been excluded from the early parenting literature providing the theoretical grounding for authoritative teaching, which largely drew from White and middle-to-upper class samples (Pettit, Bates, & Dodge, 1997). For example, studies on authoritative parenting have considered acceptance, involvement, and warmth expressed through positive praise as the hallmarks of the construct of support. However, Walker-Barnes and Mason (2001) and Hill and colleagues (2003) have argued that Black and Latino families are more likely than White families to additionally consider high behavioral expectations and "psychological control," such as threat of consequences, as indicators of support. Results from Mason and colleagues (2004) supported this expanded definition of support among Black children, finding that Black youth participants characterized highly demanding and

controlling parental behaviors like threats as controlling and manipulative, but also associated these behaviors with feeling cared for and supported. Chao and Otsuki-Clutter (2011) similarly reported that when studies defined structure (termed "control" in their review) to include monitoring behaviors like supervision and probing for details about children's activities, its benefits are consistently apparent for Black youth as well as youth from other ethnic and racial groups. Such probing behavior aligns with more controlling, authoritarian behaviors in Baumrind's original framework and have not been traditionally considered to represent positive caregiver behaviors but can be an expression of caregiver protectiveness for Black and Latino youth against discriminatory experiences and their potentially dangerous consequences. When studies broaden their operationalization of support and structure to include the cultural context of the study sample, results suggest that high levels of support and structure are associated with positive outcomes for Black youth (Chao & Otsuki-Clutter, 2011; McWayne, Owsianik, Green, & Fantuzzo, 2008).

Therefore, the warm demander pedagogy requires that support and structure be operationalized in the cultural context of the students and teachers comprising the specific individuals in a classroom. In contrast, authoritative teaching wholly excludes certain behaviors from being defined as supportive or structured when they fall outside of a White traditional perspective of parental responsivity and demand. For example, findings from the parenting and teaching literature suggest that for some Black students, these behaviors can include direct "no-nonsense" feedback that can be perceived as harsh in some populations (Ford & Sassi, 2014). Studies have also indicated a cultural practice in some Black and African American populations where Black teachers are more likely to

highlight the commonalities between themselves and their students and adopt a unified approach to learning, and where White teachers tend toward a more traditional characterization of the educator as a unilateral authority figure (Brody & Flor, 1998; Chao & Otsuki-Clutter, 2011; Ford & Sassi, 2014). While the warm demander framework continues to hold that warmth and responsivity are central to an effective approach for students from a diversity of racial and ethnic groups, it also allows for the behaviors described above to be incorporated into teacher behavior when they are balanced by an overall positive regard for students and are culturally appropriate to students' backgrounds as well as that of the teacher. For example, Ford and Sassi (2014) used student reports, teacher interviews, and classroom observations to characterize how a Black teacher and a White teacher, both of whom employed warm demander approaches, distinctly asserted their role in the classroom. The Black teacher developed her authority through strategies such as referencing commonalities between her and her students and using a "direct, no-nonsense discourse style" that occasionally included harsh language followed by expressions of warmth. In contrast, the White teacher indirectly enforced her authority through indirect and warm acknowledgments when students did not meet behavioral expectations and an emphasis on following up on students' emotional states after she delivered behavioral consequences.

Despite the distinctions between authoritative teaching and the warm demander approach, several key characteristics of teacher support and structure are consistent across both frameworks (Table 1). These features include consistently high academic and behavioral expectations, the belief that their students can excel at school, structured discipline systems, and warmth and positive regard toward their students. Additional core

features include encouraging autonomy through choices and providing a rationale for learning objectives and behavioral expectations when proactively presenting them as well as enforcing them following violation, providing adequate supervision for students, providing feedback, and consistent, even-handed systems for behavioral control. Qualitative work provides a promising foundation for systematically evaluating these characteristics and their outcomes in elementary and secondary classrooms of predominantly Black students, but these associations have been minimally assessed across classrooms, grades, and schools. The current study used classroom observations to assess these foundational characteristics of supportive and structured teaching for their impact on positive student behaviors in classrooms with teachers from a diversity of backgrounds serving predominantly Black students.

Supportive and Structured Teaching: Student Outcomes

Authoritative school climate theory maintains that pairing consistent and high levels of support with consistent and high levels of structure promotes the best student outcomes (Buyse, Verschueren, Doumen, Van Damme, & Maes, 2008; Way, 2011). Per Self-Determination Theory, whether an environment meets individuals' psychological needs for autonomy, competence, and relatedness can shape student motivation in the classroom. High levels of teacher support and structure meet students' needs for autonomy, competence, and relatedness in several ways. *Autonomy* represents individuals' tendency to initiate and regulate their own goal-directed behaviors. Supportive and structured teacher practices that promote autonomy include providing opportunities to develop students' proficiency in self-regulation, as by presenting and reinforcing clear behavioral expectations. *Competence* references students' knowledge of

how to engage in goal-directed behavior and access resources for achieving goals. Teachers can foster competence through supportive and structured practices like providing opportunities for students to self-monitor their progress. Finally, *relatedness* indicates connectedness with others in one's social milieu, which teachers can promote through demonstrating warmth, respect, and responsivity toward their students (Deci et al., 1991).

Within this framework, one expected outcome of teacher support and structure is student engagement (Deci et al., 1991). Student engagement is an important mediator of student outcomes, including academic achievement and psychological well-being (Reyes, Brackett, Rivers, White, & Salovey, 2012; Wang & Peck, 2013). Behaviors representing engagement include raising one's hand, taking notes, contributing to class discussions, and behaviorally aligning with school rules and norms (Fredricks, Blumenfeld, & Paris, 2004). As students become autonomous in their educational experience, they demonstrate engagement to facilitate further learning. Similarly, students grow in competence as they acquire skills and confidence in their own self-efficacy, which promotes their engagement in further learning. The secure attachment that characterizes relatedness with one's teacher also promotes engagement in growth experiences in the classroom as students feel confident and supported (Pitzer & Skinner, 2017). Accordingly, studies suggest a link between supportive and structured teacher behaviors and engagement of students. In a sample of classrooms from three middle schools serving 97% White students, Walker (2008) found that students demonstrated and reported the highest levels of engagement in classrooms where teachers made frequent demands *and* nurturing statements at twice the rate of their peer educators. Black student participants in several

ethnographic studies have specifically described teacher support and structure as the optimal teacher approach for engagement (Corbett & Wilson, 2002; Ford & Sassi, 2014; Ware, 2006). However, quantitative studies examining outcomes of supportive and structured teaching for Black students have been minimal and produced mixed results. Studies looking specifically at student engagement are even rarer. For example, while Sandilos, Rimm-Kaufman, and Cohen (2017) found separate main effects for teacher support and structure on Black fifth-and-sixth-grade students' achievement, the significant association of support with achievement disappeared when entered into models with indicators of structure simultaneously. However, using a different outcome variable that is more closely related to engagement, student trust in teachers, Gregory and Weinstein (2008) identified that Black students were most likely to indicate trust when they reported high levels of teacher support *and* structure than when they reported high levels of just one or neither construct. When Pellerin (2005) conducted a similar study examining high schoolers' reports of support and structure as predictors of indicators of school *disengagement* (represented by tardiness and absenteeism), he found that schools with high support/high structure profiles had the weakest association with disengagement. However, this study clustered indicators of teacher support and structure into school-level profiles rather than examining their impact at the classroom level, is limited to a high school sample, and drew from a student sample with 47% of students identifying as White and only 19% of students identifying as Black.

The limited and inconclusive findings related to the impact of teacher support and structure on outcomes for Black students are further complicated by several measurement limitations. Across the broader literature on teacher support and structure, many studies

fail to measure the specific impact of pairing both constructs into a warm demander teaching style, limiting their predictors to only support or only structure (Quin, 2017; Roorda, Koomen, Spilt, & Oort, 2011). There is also variability in the conceptualization and subsequent measurement of teacher support, teacher structure, and student engagement. Furthermore, teacher or student biases in reporting and limited diversity in participant samples can influence survey reports of teacher approaches. Also, studies relying on surveys are often limited to late elementary and older grades, as younger students' personal insight on surveys may be limited by development.

As an example of these limitations, Wang and Eccles (2013) found significant main effects for 7th grade students' reports of teacher support and structure as separate predictors of their engagement at the end of middle school. This study benefited from a diverse and broad student sample from 23 urban public middle schools, where 56% of students in the sample identified as Black, 32% as White, and 12% as Hispanic/Latino, Asian, Native American or biracial. However, the study did not examine the combined impact of support and structure and was limited by shared method variance in its exclusive reliance on student reports. Similarly, in an empirical test of self-determination theory, Skinner, Furrer, Marchand, and Kindermann (2008) examined whether the association of authoritative teaching and student engagement was mediated by students' perceptions of their autonomy, competence, and relatedness. The study sample comprised 805 4th-7th grade students and their 53 teachers in a suburban school district with more than 95% of students identifying as White. Students and teachers reported on a composite scale representing teacher support and structure, and students reported on their autonomy, competence, relatedness, and end-of-year engagement in school. Their findings best

supported the model depicting teacher support and structure as indirectly associated with student engagement via students' feelings of relatedness, competence, and autonomy. Yet although this study considered the simultaneous impact of support and structure, results varied by reporter, demonstrating the need for more objective measurement approaches such as observations to minimize the impact of reporter bias.

Survey methods are common in the study of school climate, as they can be broadly and efficiently administered and capture the subjective experience of school stakeholders (Wang & Degol, 2016). However, observational methods are also helpful for assessing specific teacher behaviors that students experience through elementary and middle school grades. Wilson, Pianta, and Stuhlman (2007) used trained observers to code frequency counts as well as global ratings of patterns of teacher support and structure in their observational study of 820 first-grade teachers' behaviors toward students. Classrooms high in both support and structure were labeled as "high quality," and a student's placement in these classrooms was found to predict students' prosocial behaviors, engagement, and self-reliance above and beyond the other classroom types (high support/low structure and low support/low structure). The authors used both rating scales and observational methods of data collection but cited potential interference of teacher biases in self-report rating scales as a concern. Furthermore, while these findings are promising, they are limited to the context of first-grade classrooms. This study also assessed predominantly White classrooms (~75% of students) and did not report teacher demographics.

In order to evaluate these findings for application to Black students' engagement in elementary and middle school classrooms, further study is needed. In contrast to

frequency, ratings, global ratings of observations can provide a robust measure of patterns that frequency ratings do not necessarily capture. Global ratings can be particularly helpful when observing classrooms from different grade levels, as global ratings operationalize intended domains of behavior while allowing for some flexible application informed by developmental appropriateness (Larson, Pas, Bradshaw, Rosenberg, & Day-Vines, 2018).

The Impact of Students on Teachers

In their review of the school climate literature, Wang and Degol (2016) reported that most school climate studies employ cross-sectional designs that view study variables as stable, justifying cross-sectional study designs. However, the theoretical basis for school climate also characterizes student and teacher outcomes as processes that can feed back to further influence each other over time. As self-determination theory links teacher behaviors to student outcomes, it can also be applied to understand how students might influence teachers' behavior in the classroom. Teachers' needs for autonomy, competence, and relatedness can be challenged by their interactions with students and may contribute to less supportive, more structured teacher behaviors. Difficult student behaviors, such as low engagement, can threaten teachers' sense of competence, or their knowledge that they can bring about a desired behavior; autonomy, or confidence in their own skills; and relatedness, or positive connection with students; subsequently, their responsivity to students may decrease (Larson et al., 2018; Yoon, 2002).

These influences of student behaviors can extend to the reinforcement or modification of teacher approaches later in the school year (Rogosa, Floden, & Willett, 1984). Reeve (2009) suggested that lower student engagement could lead teachers to

demonstrate low support/high structure approaches in the classroom. Pelletier, Seguin-Levesque, and Legault (2002) drew from a sample of 254 French-Canadian teachers from first grade through twelfth grade to examine the association between teacher perceptions of student motivation and their reported use of controlling or supportive teaching behaviors. The study also examined the potential impact of teachers' pressure at work and personal motivation for teaching. Among the three models tested, results supported the fit of a model where perceived student motivation and teachers' pressures at work explained 18% of the variance in teacher motivation, which in turn explained 13% of the variance in teaching behaviors. Overall, their findings supported that teacher perceptions of student motivation, and the behavioral indicators teachers may use to develop these perceptions, play a role in determining subsequent teacher behaviors, and that when teachers perceived student motivation to be high, they tended toward more supportive teaching behaviors.

Studies such as that by Pelletier and colleagues (2002) highlight the complexity of teacher-student patterns of perception and behavior that develop over time. Yet, the pool of longitudinal studies examining these patterns across diverse student samples, teacher samples, and grade levels are sparse. One longitudinal study found that teacher support and structure predicted student engagement, and that students demonstrating more engagement in turn received more support and structure from their teachers than less engaged students. Skinner and Belmont (1993) longitudinally assessed these associations in 14 third-, fourth-, and fifth-grade classrooms serving a 94% White student population from lower middle-class and middle-class socioeconomic groups. The authors hypothesized that teacher support and structure would predict student engagement.

Results supported a model illustrating that higher levels of teacher-perceived student engagement at the beginning of the year predicted higher levels of self-reported teacher support at the end of the year. To a lesser extent, observed student engagement also positively predicted self-reported teacher structure at the end of the year. A strength of this study was its comparison of teacher and student reports of teacher and student behaviors, which highlighted particular divergence relating to perceptions of teacher support and structure. Such divergence illustrates the need for examining the associations between teacher and student behaviors using a consistent, relatively objective methodology such as observations.

Though other studies have not applied a similar methodology to specifically examine student engagement in this pathway, Pitzer and Skinner (2017) found support for a similar model examining feedback between teacher behaviors and several dimensions of student motivation, which the study conceptualized to include student engagement. All constructs were measured using student self-reports. Study hypotheses held that higher teacher support at the beginning of the year would contribute to student motivation and that students with higher motivation would report more teacher support at the end of the year. The study sample included 1020 third through sixth graders from a rural-suburban New York school district, with 95% of students identifying as White and from lower to middle-class families. Students completed several measures in the fall and spring of the school year, including perceived teacher support as well as student self-reports of their own motivation and engagement. The final model demonstrated that teacher support at the beginning of the year predicted indicators of student motivation, including student engagement. Furthermore, students indicating greater motivation

reported received more supportive and structured practices from their teachers at the end of the year.

These studies point toward a likely reciprocal effect of teacher and student behaviors in the classroom, where supportive and structured teacher behaviors promote student engagement, and that engagement reinforces or attenuates the stability of supportive and structured teaching behaviors. The current study extended prior work (Skinner & Belmont, 1993) by assessing the relative influences of engagement in a longitudinal model examining the stability of teacher support and structure in predominantly Black classrooms from the beginning of the school year to the end of the school year.

Teacher and Student Characteristics as Control Variables

The current study considered the potential impact of teacher and student characteristics on the association between teacher support and structure and student engagement. Guided by previous literature addressing the relationship between teacher and student behaviors in the classroom, this study examined and controlled for the potential influence of teacher experience, and student grade level (elementary or secondary). This study used data from observations completed in classrooms where 67% of teachers had five years or fewer of experience in the classroom. Fifty-two percent of classrooms served elementary (grades K-4) students, and 48% served middle school (grades 5-8).

Teacher Experience

Though minimally evaluated across various student outcomes, several studies indicate that teacher experience is positively associated with student achievement

(Clotfelder, Ladd, & Vigdor, 2007; Easton-Brooks & Davis, 2009). Though its impact on classroom engagement has not been evaluated, teacher experience has been negatively associated with several indicators of student *disengagement* in secondary grades, including absenteeism (Ladd & Sorensen, 2017). Furthermore, variations in teacher preparation and experience are associated with student outcomes and may have particularly salient impacts among inexperienced teachers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009).

Student Grade Level

Student grade level (elementary or secondary) represents a developmental correlate that may impact teachers' supportive and structured teaching and student engagement. There is evidence across studies with racially and ethnically diverse student samples that the prevalence of supportive and structured teaching varies by grade level, where middle school students consistently report a marked decrease in support from adults (Bergin & Bergin, 2009; Furrer & Skinner, 2003; Hargreaves, 2000; Roorda et al., 2011; Way, Reddy, & Rhodes, 2007). In comparison to middle school classrooms, several studies have shown elementary school classrooms to include less emphasis on teacher control and discipline, more opportunities for choice, more differentiated instruction, and more positive teacher-student relationships than middle school classrooms (Eccles et al., 1993). This variable carries particular significance when evaluating teaching approaches employed with Black students, as further studies have demonstrated that Black students in secondary grades are less likely to experience effective teaching approaches than White students (Harber, Gengaro, Butisingh, Tsang, & Ouellette, 2012; Wentzel, 2002).

The Current Study

Black students regularly report more negative perceptions of school climate than their peers (Bottiani, Bradshaw, & Mendelson, 2017; Konold et al., 2017; Voight, Hanson, O'Malley, & Adekanye, 2015). The prevalence and nature of school climate features that promote positive perceptions and associated outcomes among Black elementary and middle school students warrant further study. Teacher-student relationships, and specifically teachers' supportive and structured behaviors that play into those relationships, comprise a key facet of students' school experiences. Furthermore, student behavioral responses to supportive and structured teaching behaviors indicate their experiences of the classroom without interference from the potential bias that survey methods can present. Engagement is one such response that has been shown to mediate student outcomes including academic achievement and psychological well-being (Reyes, Brackett, Rivers, White, & Salovey, 2012; Wang & Peck, 2013). Teacher response to student engagement further shapes classroom patterns of behavior by teachers as well as students. The current study contributes to the existing literature on effective teaching behaviors and student engagement in several important ways.

First, the current study contributes to the existing literature with a study design that allows for the simultaneous evaluation of teacher support and teacher structure. Theoretical frameworks from the parenting literature, authoritative school climate theory, and the warm demander pedagogical approach predicate that teacher support must be paired with structure in order to foster the best outcomes for children (Baumrind, 1968; Gregory & Cornell, 2009; Kleinfeld, 1975). Despite its theoretical importance, studies

often measure only teacher support or only teacher structure, providing a limited picture of effective teaching approaches. (Quin, 2017; Roorda, Koomen, Spilt, & Oort, 2011).

Second, the current study's use of observational data collection with global ratings contributes several methodological advantages. Observational methodology limits the interference of bias from teacher or student reporters. It should be noted that this methodology is not immune to observer bias related to observer characteristics such as training experience and racial and ethnic background; it is critical for studies using this methodology to speak to that likely presence of bias within the observation team. Observational methodology also allows for data collection across a breadth of student developmental levels. While many studies have collected data directly from students, younger student participants may be developmentally limited in their ability to sustain attention to surveys. In contrast, behavior observations are sensitive to short-term changes in behavior as well as patterns over time (Snyder et al., 2006). As such, observational data collection methods that use global ratings allow for intentional adjustment to classrooms' developmental levels and context while maintaining a consistent approach.

Third, the current study builds upon the previous literature examining teacher behaviors as predictors of student behaviors, and vice versa. Qualitative work provides an impetus for systematically evaluating teacher support and structure and their association with student outcomes in elementary and secondary classrooms of predominantly Black students (Ford & Sassi, 2014; Nelson, 2016; Ware, 2006). Studies that have examined associations between teacher behaviors and Black student behaviors are rare and often cross-sectional (e.g., Debnam et al., 2015; Larson et al., 2018); those few that have

employed longitudinal methods have drawn from samples serving predominantly White student samples (Pelletier, Seguin-Levesque, & Legault, 2002; Pitzer & Skinner, 2017). The current study collected observational data at four timepoints per year across two years in order to assess the pathway from early teacher and support structure to student engagement to teacher support and structure at the end of the year. This study also examined teacher support, teacher structure, and student engagement within a sample of educators from a diversity of backgrounds working in four urban public charter schools that served K-8 students from predominantly low-income backgrounds and identifying as Black or African American. Data was collected across time, grades, and in classrooms serving a student population that is currently underrepresented in the literature examining teacher support, teacher structure, and student engagement. Furthermore, characteristics of the teacher (teacher experience) as well as students' developmental level (grade level) were considered for their potential influence on associations within this pathway. Figure 1 presents the hypothesized structural equation model.

Study Aims and Hypotheses

1. **Assess the association between supportive and structured teacher behaviors and student engagement in K-8 classrooms serving primarily (>85%) Black students across four charter schools in New Orleans.** Student grade level (elementary or middle) and years of teacher experience were entered as controls. It was hypothesized that higher levels of supportive and structured teaching at the beginning of the school year would predict higher student engagement in the middle of the year.

2. **Evaluate student engagement as a mediator of the association between beginning-of-the-year supportive and structured teaching and end-of-year supportive and structured teaching.** It was hypothesized that student engagement would mediate the stability of teachers' supportive and structured teaching over the course of the school year when controlling for student grade level, years of teacher experience in education, and observation year. The mediation effect was hypothesized to be positive, such that student engagement would account for the positive association between beginning-of-the-year teacher support and structure and end-of-the-year teacher support and structure.

Method

The current study is based on secondary data analysis of a larger study, Safe Schools NOLA. Safe Schools NOLA is a four-year study funded by the National Institute of Justice that utilized a multiple baseline design to evaluate the effectiveness of a multi-component implementation strategy for trauma-informed care in six New Orleans K-8 public charter schools serving students from predominantly low-income and racial and ethnic minority backgrounds (The Cowen Institute, 2015; New Orleans Parents' Guide, 2016). The six schools were managed by two charter management organizations (CMOs), with each CMO managing three of the study schools. Study data was collected from all six schools during all four years of the larger study, beginning with a baseline year in the 2016-2017 academic year. In each subsequent year of the larger study, two schools received the intervention. Training for observational data collection began partway through the baseline year and was ongoing for the remainder of the study. The current study used observational data collected in the first two intervention years of the study

(2017-2018 and 2018-2019) from the four schools (2017-2018; observation year one; two schools from each CMO; Cohort 1), then two schools (2018-2019; observation year two; one school from each CMO; Cohort 2), that had not yet received the intervention.

Participants

One hundred thirty-seven classroom teachers from four schools comprised the teacher study sample, with teachers from each school representing 20-30% of the total sample. Demographic information was available for 106 of 137 (77%) participants. Of the 106 teachers who reported demographic information, 75% identified as female and 79% were between the ages of 18 and 34. Fifty-one percent identified as White, 43% identified as Black or African American, 2% identified as "other" races (including Latino/a, Cajun, Arab, and Mexican American), and the remaining 4% identified with multiple groups. Across racial groups, 5% of teachers identified ethnically as Hispanic or Latino. Teachers were divided relatively evenly across elementary (K-4th grades; 52%) and middle school grades (5th-8th grades; 48%). Sixty-one percent of demographic reporters had trained in an alternative certification education program, and 67% had 5 or fewer years of experience working in education.

Classroom observations were completed for 72% of teachers during observation year one (Cohort 1), with the remaining 27% of observations completed during observation year two (Cohort 2). Chi square tests of independence indicated that across almost all demographic areas, observation data did not differ between teachers observed during year one and those observed during year two. Data only differed by which schools were represented each year ($X^2 = 35.28$, $df = 3$, $p < .01$); the two schools providing data across both observation years of the study comprised more of the study sample than the

other two schools. Further details on teacher participant demographics are available in Table 2.

Although the current study did not utilize individual student data, classroom-level demographics were noted by observers during each observation. On average, teachers taught classrooms of approximately 24 students ($M = 23.64$, $SD = 5.18$), with generally about half of the class comprised of male students ($M_{number\ of\ male\ students} = 12.29$, $SD = 12.79$), and approximately zero to one student per class presenting as White or Caucasian ($M_{number\ of\ White\ students} = .15$, $SD = .50$).

Procedure

At the outset of the 2016-2017 school year, teacher participants provided formal consent to participate in the larger study. Teacher recruitment continued as new teachers were hired at each participating school over the study lifespan. Data collection was structured using a random planned missingness design in order to reduce data collection cost and participant burden while preserving the validity of the study. This design, also termed wave-missing longitudinal design, allows for complete data collection from each participant within a randomly assigned subset of measurement waves. By intentionally assigning random incidences of missing data, this study design reduces the likelihood that data might be missing from the subsequent dataset by some other, likely not random, mechanism (Harel, Stratton, & Aseltine, 2011; Little, Jorgensen, Lang, & Moore, 2013). Within the current study, consenting participants were randomly assigned to two of four waves of data collection per study year. At each data collection wave, teachers were observed and completed self-reports as well as student reports. As teachers were recruited on a rolling basis over all years of the study, teachers completed demographic surveys at

their first assigned wave in the study. Teachers participating in both observational data collection and self-report surveys received $10 in compensation for each wave of completed surveys. Table 3 provides the breakdown of number of participants by timepoint and wave combination.

Teachers were observed during general instruction in classrooms of approximately 20-30 students. Observational data was collected using the Assessing School Settings: Interactions of Students and Teachers (ASSIST) Observation System (Rusby, Taylor, & Milchak, 2001), which is a school-based observation system that includes components for assessing student behaviors in the classroom as well as teacher classroom management practices. For each observation, one to two trained observers from the study team completed a 23-minute classroom observation consisting of three minutes orienting to the classroom context, followed by twenty minutes of formal observation.

Observational training procedures followed those used in previous observation studies (Pas et al., 2015) and included one full day of coding and discussing practice videos to develop reliable coding. In order to maintain inter-observer agreement following the training, the observer team met on a monthly basis, then a quarterly basis, to complete and assess inter-rater reliability on coding practice videos that had been previously coded by a master coding team with at least one year of observer experience in the study. During the year, observer pairs were assigned to double-code approximately one-third of observations to inform ongoing calculations of interrater reliability. Interrater reliability was calculated as a measure of consistency between each pair within their assigned double-coded observation, and the individuals comprising each pair varied

randomly across observations and timepoints. In the current sample, 137 teachers were observed two times across eight timepoints (two observation years of four timepoint each). Across the eight timepoints, 222 discrete observations were completed, and 63 of those observations (28%) were double-coded. Inter-observer agreement was calculated by dividing the total number of agreements by the total number of agreements and disagreements (Pas et al., 2015). Agreement at or above 80% was considered sufficient. In the current study, agreement was lowest but acceptable at timepoint 1 (83%); at timepoints 2 and 4, agreement was 88% and 90% respectively. Data from both observations completed for each teacher within an observation year were used in study analyses. Across both years of observational data collection used in the current study, 24 teachers provided data in both years. A coinflip was used to randomly select which year of data for each of these teachers would be used in the study.

Across Cohorts 1 and 2, ten observers comprised the observation team (Table 4). This team included six masters-level school psychology doctoral students and two graduate students completing their masters' degrees in school psychology, all of whom who had completed coursework and practica in behavioral observation and assessment. Additionally, the team included one postdoctoral research coordinator with a doctorate in school psychology as well as one bachelor's-level research coordinator with undergraduate coursework in behavioral assessment. Eight members of the ten-person observation team identified as White or Caucasian (80%), and two members identified as Black or African American (20%). 70% of the team did not have classroom teaching experience; the three individuals with teaching experience reported having 1-5 years of

experience. However, 70% of the team endorsed having worked in a school in some capacity for at least one year.

Given the homogenous backgrounds of the predominantly female, predominantly White observation team, maintenance procedures for reliability of observational coding included open discussions about observer reactions to teacher and student behaviors and monitoring potential biases in these reactions by honing how teacher and student behaviors, including teacher support, teacher structure, and student engagement were operationally defined in the code manual. For example, through these conversations, the team decided to discount "tone of voice" from operationalized definitions of teacher behaviors, as perceptions of the relative harshness or warmth of tone are subjective and rooted in individuals' experiences with tone (Ford & Sassi, 2014). The regularity of coding meetings allowed observational procedures to be iteratively honed and monitored for relevance, bias, and inconsistencies, and most of the observation team completed coursework in developmental child psychology across diverse populations. However, no formal procedures were used to directly mitigate observer bias. As such, results should be interpreted in the context of observer demographics and the likelihood that observation procedures did not fully remove observer bias from the study.

Measures

This section describes the measures used to examine each construct of interest (Table 5). For all constructs, the current study utilized data from observers' global ratings of several dimensions of teacher and student behaviors that were completed immediately following each observation. The background, description, and reliability statistics from previous literature are presented here for each measure. For the factors representing

teacher support and structure, confirmatory factor analyses were used to respecify their measurement to improve fit in the structural equation model used to assess study hypotheses (see Analytic Approach and Appendix for further description of respecification). Descriptive statistics and reliability statistics for each original measure in the current sample, as well as each respecified measure, are included below. Table 6 presents correlations between study demographics and the original measures as designed, prior to respecification. The Appendix section presents details on the decisions and results of steps taken for respecification of the measurement model.

Teacher Support

Teacher support was represented using global ratings from the ten-item Sensitivity subscale of the Caregiver Interaction Scale (CIS; Arnett, 1989) collected during Timepoint 1 and Timepoint 4. The CIS is a rating scale designed for observers to measure teacher behavior toward students. Four subscales comprise the CIS: Harshness (9 items) reflects overly critical and hostile teacher behaviors and includes items such as, "speaks with irritation or hostility to the children"; Permissiveness (three items) indicates passivity and excessive tolerance of misbehavior and includes items such as, "doesn't reprimand children when necessary; Detachment (four items) represents low levels of warmth and includes items such as, "Spends considerable time in activity not involving children"; and Sensitivity (ten items) indicates teacher warmth and responsivity and includes items such as, "seems enthusiastic about the children's efforts." Immediately following the observation period, observers rated teachers on each item of the Sensitivity subscale, as well as the larger CIS scale, using a rating scale from 1-4, where 1 = not at all true and 4 = very much true. The measure was slightly modified for the current study

to ensure that the items were relevant for older elementary and middle school students. Additionally, a fifth option to indicate "not applicable" was provided but never endorsed for the CIS Sensitivity subscale in the current study sample.

Prior studies have used the CIS Sensitivity subscale to validate other measures of teacher responsivity (e.g., de Kruif, McWilliam, Ridley, & Wakely, 2000). This subscale has been found to be a reliable and valid indicator of teacher sensitivity toward low-income, racial and ethnic minority preschoolers, although studies reporting its validity and reliability have been limited to early childhood settings ($\alpha = 0.85$; Colwell, Gordon, Fujimoto, Kaestner, & Korenmann, 2013). In the current sample, the internal consistency of the ten-item subscale was considered good at timepoint 1 ($\alpha = 0.86$) and at timepoint 4 ($\alpha = 0.80$). The mean score for the ten-item subscale was 2.68 (*SD* = .68) at timepoint 1 and 2.30 (*SD* = .50) at timepoint 4.

The respecification process supported a revised factor structure for the CIS Sensitivity subscale that retained six items from the original subscale and collapsed response categories into the following revised categories: 1 = not at all true; 2 = somewhat true; and 3 = quite a bit or very much true. The mean score for the revised six-item subscale was 2.09 (*SD* = .47) at timepoint 1 and 1.95 (*SD* = .46) at timepoint 4, with good internal consistency at both timepoints ($\alpha_{timepoint1} = 0.83$; $\alpha_{timepoint4} = 0.81$). Refer to the Appendix for item-level statistics obtained through confirmatory factor analysis for the original ten-item subscale as well as the revised six-item subscale.

Teacher Structure

Teacher structure was represented using the seven items of the Teacher Proactive Behavior Management subscale of the Assessing School Settings: Interactions of

Students and Teachers Observation System (ASSIST; Rusby, Taylor, & Milchak, 2001). The ASSIST includes frequencies tallies and global ratings of various aspects of classroom management on several subscales including: Teacher Proactive Behavior Management (seven items) and Student Compliance (six items). Observers provided global ratings of Teacher Proactive Behavior Management immediately following each observation using a scale ranging from 1 (Never) to 5 (Almost Continuously). A sixth option to indicate "not applicable" was provided but never endorsed for the Teacher Proactive Behavior Management subscale in the current study sample. Previous studies have used the ASSIST system across elementary, middle, and high schools. However, the student populations served in these studies ranged from 52% to 82% White-identifying students respectively, reflecting the limited application of the ASSIST system in schools serving racially and ethnically diverse students (Pas et al., 2015; Rusby, Crowley, Sprague, & Biglan, 2011). A recent generalizability study of the ASSIST demonstrated strong internal consistency among classrooms from six high schools with a racial and ethnic minority rate of 48% among students (Abry, Cash, & Bradshaw, 2018). Larson and colleagues (2018) recently used the Teacher Proactive Behavior Management subscale to represent evidence-based and positive classroom management practices in their study across 18 elementary and middle schools with 41% of students identifying as White and 34% of students identifying as Black or African American. Though not specifically termed *structure* in their study, the authors articulated that this subscale was selected for its correspondence to teacher behaviors that include setting clear expectations, anticipating student difficulties, and creating structures that pre-correct and prevent undesired behaviors. Their study, which used a four-item subset of the Teacher

Proactive Behavior management subscale, reported its internal consistency as questionable (α = .66); in a less diverse sample, Rusby and colleagues (2011) reported the internal consistency of the full seven-item subscale as good (α = .87). In the current study sample, the internal consistency of the original seven-item subscale was considered very poor at Timepoint 1 (α = .20) and good at Timepoint 4 (α = .75). Poor reliability at timepoint 1 was due to substantial missing data and lack of variance in responses. The mean score for the seven-item subscale was 2.78 (*SD* = 1.03) at timepoint 1 and 2.25 (*SD* = .70) at timepoint 4.

The respecification process supported a revised factor structure for the ASSIST Proactive Behavior Management subscale that retained five items from the original subscale and collapsed response categories into the following revised categories: 1 = never; 2 = at times; and 3 = a lot or almost always. The mean score for the revised five-item subscale was 2.35 (*SD* = .43) at timepoint 1 and 2.06 (*SD* = .42) at timepoint 4, with adequate internal consistency at both timepoints ($\alpha_{timepoint1}$= 0.64; $\alpha_{timepoint4}$ = 0.70). Refer to the Appendix for item-level statistics obtained through confirmatory factor analysis for the original seven-item subscale as well as the revised five-item subscale.

Student Engagement

Student engagement was represented by Timepoint 2 global ratings on the six-item ASSIST Student Compliance subscale, which has been previously used to represent student engagement through assessing student behaviors of compliance and involvement in classroom activities (Larson et al, 2018). Also termed Student Cooperation, this subscale includes items representing student responsivity to prompts and classroom structures (“students comply”), student enthusiasm (“students are interested, enthusiastic,

and involved,") and sustained attention to classroom content ("students are focused and engaged"). Observers provided global ratings of Student Compliance immediately following each observation using a scale ranging from 1 (Never) to 5 (Almost Continuously). A sixth option to indicate "not applicable" was provided but never endorsed for the Student Compliance subscale in the current study sample. Pas and colleagues (2015) reported strong internal consistency for this subscale (α = .92). The internal consistency of the scale at Timepoint 2 in the current study was acceptable (α = .70) with a mean score of 3.61 (*SD* = .85). The factor structure for the original subscale demonstrated acceptable fit in confirmatory factor analysis. Therefore, it was retained for use in the final structural model with no creation of parcels or revisions to response categories. Refer to the Appendix for item-level statistics obtained through confirmatory factor analysis supporting the retained subscale.

Grade Level

Following each observation, observers recorded various classroom characteristics including the grade level of the students in the classroom. In the current study, observations were recoded to grade categories representing elementary school (grades K-4) and middle school (grades 5-8).

Teacher Experience

Teachers provided their demographic information, including their years of experience in education, as part of their self-report survey responses. Other demographic data included their gender, racial and ethnic group identification, age, education background, certification pathway (traditional or alternative, such as through Teach for

America), years as a teacher, years at their school, and years of experience in the education field.

Observation Year

Observation year in the study (Cohort 1 or Cohort 2) corresponded to timestamps for observations.

Analytic Approach

Structural Equation Modeling (SEM) was used to evaluate the proposed latent structural regression model (Figure 1) representing study hypotheses that beginning-of-the-year teacher support and structure would positively predict student engagement in the middle of the year (Hypothesis 1), and that student engagement would fully mediate a significant positive association between teacher support and structure at the beginning of the year and at the end of the year (Hypothesis 2). Teacher structure and teacher support were indicators for a higher-order factor comprising both constructs, hereafter referred to as *supportive and structured teaching.* The current study evaluated Hypothesis 1 by assessing for a significant and positive association between timepoint 1 supportive and structured teaching and timepoint 2 student engagement. This study assessed Hypothesis 2 by testing for a significant and positive direct effect between timepoint 1 supportive and structured teaching and timepoint 4 supportive and structured teaching and a significant mediating effect of student engagement on that association.

Dichotomous variables representing observation year (Cohort 1 or 2), grade level (elementary, grades K-4; and middle, grades 5-8) and teacher experience in education ($\leq$5 years of experience and 6+ years of experience) were entered as categorical control variables.

Several features of the proposed model warranted the use of SEM to analyze study hypotheses. First, SEM allows for the evaluation of direct and indirect associations between latent variables of interest. It allows the researcher to account for measurement error that is random or due to poor reliability of scores by explicitly including it in the analyses. Furthermore, SEM can analyze complex longitudinal models with multiple types of data and levels of missingness. Given the planned missingness design of the study, SEM's compatibility with Full Information Maximum Likelihood (FIML) estimation allows for all possible data to be included in analyses. FIML uses subsets of cases with available data to calculate parameter estimates and their standard errors (Collins, Schafer, & Kam, 2001; Kline, 2011). The proposed model also included ordered-categorical item-level indicators, categorical control variables, higher-order latent factors, and multiple timepoints. SEM allowed for the measurement models and structural model to incorporate these various representations of data, preserving close integrity to the data as originally collected.

Initial estimation of the proposed mediation model to assess Hypotheses 1 and 2 (Figure 1) indicated poor model fit. Subsequently, theoretically- and empirically grounded revisions to the measurement of latent factors demonstrating poor model fit were employed. These revisions, referred to as respecification of the measurement model, were informed by results from confirmatory factor analyses completed using IBM SPSS Amos 26. This respecification process enhances the measurement model's representation of each factor and improves overall model fit (Kline, 2011). For the current study, the respecification process involved assessing the factor loadings, correlation matrices, and theoretical relevance of items comprising each factor. This information informed

decisions for trimming, or adding item indicators for a latent factor, as well as averaging items into groups that form indicators ("parceling"; Little, 2013). The Appendix details the general steps, decision points, and results of specifying and respecifying the measurement model for each construct. Following respecification, structural equation modeling assessed three iterations of measurement models to achieve optimum model fit and conceptual strength for assessing the hypothesized structural model. All models imposed measurement invariance of supportive and structured teaching across timepoints. Each iteration was tested using Mplus Version 8.1 with maximum likelihood estimation and FIML. The section below presents the fit indices for each of the three measurement models, describes the differences between models, and provides the rationale for determining which model was most viable for producing valid and interpretable analyses of study hypotheses.

Model 1: Item-Level Indicators Created Latent Factors Representing Supportive and Structured Teaching and Student Engagement

The first measurement model used item-level indicators from the CIS Sensitivity subscale at timepoint 1 and 4 and the ASSIST Proactive Behavior Management subscale at timepoints 1 and 4 to comprise the supportive and structured teaching higher-order latent factor. The latent factor for timepoint 2 student engagement was represented by item-level indicators from the ASSIST Student Compliance subscale at timepoint 2. This initial measurement model allowed for all study measures to retain the item structure established in the extant literature. Furthermore, a fully latent model benefits from accounting for measurement error when representing each construct, providing more accurate parameter estimates. However, one downside to representing all factors with

item-level indicators is that structural equation modeling requires that available data, or "observations," equal or outnumber the parameters in the model. This condition is necessary for the model to be identified and successfully produce unique parameter estimates for the proposed structural model. As each item introduces a parameter into the model, use of item-level indicators risks including too many parameters and precluding analysis of the model when the model is underpowered (Kline, 2011). The current model treated all item-level indicators as ordered-categorical data. For observer-report Likert scale items, ordered-categorical treatment of data reduces issues related to using continuous indicators, including skewness and limited response categories that lead to biased estimates of parameters (Liu et al., 2017). Five iterations of this model returned continued parameter errors that precluded interpretation of results.

Model 2: Mean Subscale Scores Created Manifest Variables Representing Teacher Support, Teacher Structure, and Student Engagement

The second model used mean subscale scores to represent manifest variables for each separate construct of teacher support, teacher structure, and student engagement. Subscale scores for teacher support and teacher structure still comprised supportive and structured teaching as a higher-order factor. Given the parameter errors of Model 1, Model 2 took a conservative approach to estimating the measurement model of study constructs. Representing each construct using subscale scores greatly reduces the complexity of the model by substantially eliminating parameters. However, such reduced complexity also reduces the accuracy of parameter estimates, as manifest variables only reflect observed data and cannot account for measurement error (Ayán, Ramirez, & Díaz, 2009). The initial attempt at this model returned a parameter error that recommended a

fixed correlation of zero between years of experience in the education field and observation year in the study. Following this recommendation returned a model that demonstrated adequate fit across most indices ($\chi 2(10, N = 137) = 4.73, p > .05$; RMSEA > .05, CFI=1.00, TLI=1.21, SRMR=0.12).

Model 3: Parcels Created Latent Factors Representing Supportive and Structured Teaching; Item-Level Indicators Created the Latent Factor Representing Student Engagement

The third model used the identified parcels for teacher support and structure as manifest indicators of the higher-order latent factor representing supportive and structured teaching at timepoints 1 and 4, item-level indicators of the latent factor representing student engagement at timepoint 2, and a fixed correlation of zero between years in the field of education and year observed. This measurement model provided the benefit of allowing for item-level indicators to comprise a latent factor for at least one of the variables of interest, student engagement, using the items as delineated by the originally established subscale. Attempts to use the originally established subscales representing teacher support and teacher structure indicated poor fit and excessive parameters, which precluded identification of the model. Subsequently, respecification was determined to be appropriate for these factors. Respecification used item parcels representing teacher support and teacher structure to comprise the latent higher-order factor of supportive and structured teaching. Item parcels allowed for improved reliability and distributions as well as reduced model complexity, parameter estimates, and sampling variability (Kline, 2011; Little et al., 2013). However, the use of parcels can reduce the breadth and depth of a construct's representation, as the relative contributions

of individual items are reduced (Little, Cunningham, Shahar, & Widaman, 2002). Following consideration of these strengths and limitations relative to those of the other measurement models tested, this model was retained over Models 1 and 2 as the most viable analysis of study hypotheses. Results are presented below for Model 3 (Figure 2), with standardized coefficients reported.

Results

Sample Descriptive Analyses

This section presents descriptive statistics for demographics and variables of interest in the study sample. Table 7 presents the correlations among all observed variables including demographic control variables, subscale scores, and parcels. Statistics for support and structure are described separately rather than for the second-order latent factor of supportive and structured teaching to better characterize the data.

Overall, ratings of teacher support were moderate. On the 4-point Likert scale used for this measure, the mean score at both timepoints fell between the "somewhat true" and "quite a bit true" response categories, suggesting that teachers generally demonstrated supportive behaviors at least some of the time ($TP1_{support}$ $M = 2.68$, $SD = .67$; $TP4_{support}$ $M = 2.30$, $SD = .50$). Teacher support ratings at Timepoint 4 were not significantly different from those at Timepoint 1. Timepoint 1 support was adequately correlated with timepoint 1 structure ($r = .68$) and timepoint 2 student engagement ($r = .43$) and poorly correlated with timepoint 4 support ($r = .10$) and timepoint 4 structure ($r = .10$). Mean ratings of teacher structure were similar ($TP1_{structure}$ $M = 2.8$, $SD = 1.03$; $TP4_{structure}$ $M = 2.25$, $SD = .70$) but because they were collected on a 5-point Likert scale, a

mean score falling between 2 (“seldom”) and 3 (“some of the time”) suggests teachers showed slightly lower tendencies toward structure. Teacher structure ratings at Timepoint 4 were not significantly different from those at Timepoint 1. Timepoint 1 structure was poorly correlated with timepoint 2 student engagement (r = .18), timepoint 4 support (r = .12). and timepoint 4 structure (r = .12). Student engagement ratings at Timepoint 2 were high (M = 3.61, SD = .85) and suggested that students were generally engaged at least some of the time during the observation period.

Study participants did not significantly differ on variables of interest based on gender, observation year, education level, years spent as a teacher, years spent at their school, or identification as Hispanic/Latino. Though demographic information was not provided for 31 of the 137 teachers comprising the study sample, teachers did not significantly vary on study variables based on whether they had reported their demographics. Crosstab analyses examined associations between demographic categories. Results indicated several differences in demographic composition across schools. School 1 and School 4 were overrepresented in the sample due to contributing data from two observation years instead of one ($X^2(3, N=106) = 35.28, p < .01$). Additionally, schools also varied by teacher age ($X^2(3, N=106) = 18.81, p = .03$) and race ($X^2(3, N=106) = 17.25, p = .05$). Schools 1 and 4 had a greater proportion of teachers under the age of 25 than schools 2 and 3. The majority of teachers in Schools 1, 2, and 4 identified as White or Caucasian, and the majority of teachers in School 3 identified as Black or African American. Results also indicated gender differences in grade level ($X^2(3, N=106) = 6.15, p = .01$) and certification ($X^2(3, N=106) = 9.57, p = .02$), where males were more likely to teach at the middle school level and pursue an alternative

education degree than females. Alternative certifications were also more common among teachers in their first five years of working in the education field ($X^2(3, N=106) = 30.32$, $p = .01$) and White teachers; the majority of Black teachers and more experienced teachers reported pursuing a traditional education degree ($X^2(3, N=106) = 24.00$, $p < .01$).

Associations between teacher demographics and study variables are further discussed here using subscales calculated from the revised items and response categories for teacher support and teacher structure. Timepoint 2 engagement did not require any revision through the respecification process. Within the sample of 106 teachers providing demographic information, teacher demographic characteristics including age and race were associated with distinct patterns of observer ratings for teachers.

Although teachers did not differ in timepoint 1 observer ratings of support and structure based on race, timepoint 4 observer ratings indicated significantly more support and structure among White teachers than Black teachers and teachers identifying with multiple racial or ethnic groups ($F_{support}(3,1) = 3.25$, $p = .03$; $F_{structure}(3,1) = 3.44$, $p = .03$).

Teacher professional characteristics including school, their experience in education, and whether they taught at an elementary or middle school were also associated with observer ratings for teachers during timepoint 4. Although teachers did not vary in their ratings of support for timepoint 1 or timepoint 4 based on their experience in education, teachers with five or fewer years working in education had significantly higher ratings for structure at the end of the year than teachers with more experience ($t(46) = 2.19$, $p = .03$). Furthermore, while teachers at the elementary school level and middle school level demonstrated similar levels of support and structure at the beginning of the year, elementary teachers demonstrated significantly more support and

structure at the end of the year than middle school teachers ($t_{support}(57) = 2.53$, $p = .02$; $t_{structure}(57) = 2.45$, $p = .02$). Finally, support and structure ratings at timepoint 4 varied by school. For timepoint 4 support as well as timepoint 4 structure, School 1 earned the highest ratings and School 3 earned the lowest ($F_{support}(3,1) = 3.87$, $p = .02$; $F_{structure}(3,1) = 4.37$, $p = .01$).

Student engagement significantly varied by teacher age, where ratings were highest in classrooms with teachers between ages 18 and 24, and between ages 45 and 54 ($F(3,1) = 4.21$, $p = .01$). However, these groups altogether comprised approximately 20 percent of the sample providing demographic information and were therefore minimally represented.

Structural Equation Models

Hypothesis 1 posited a positive association between beginning-of-the-year supportive and structured teaching and mid-year student engagement. As predicted, the association between early teacher behaviors and mid-year student engagement was significant ($\beta = .42$, $p < .01$). However, the other associations within the proposed mediation model were not supported. Hypothesis 2 proposed that student engagement would mediate the stability of teachers' supportive and structured teaching over the school year, where beginning-of-the-year supportive and structured teaching would be associated with end-of-the-year supportive and structured teaching, and that this relationship would be fully mediated by mid-year student engagement. Contrary to hypothesis 2, a direct effect of beginning-of-the-year teacher behaviors was not observed, as beginning-of-the-year supportive and structured teaching was not associated with end-of-year supportive and structured teaching ($\beta = .21$, $p > .05$). Furthermore, an association

between mid-year student engagement and end-of-year supportive and structured teaching was not supported (β = .16, p > .05). Accordingly, an indirect effect of student engagement was not supported (β = .07, p > .05). As noted above in descriptive analyses regarding control variables, grade level did not predict a significant portion of the variance for beginning-of-the-year supportive and structured teaching (β = -0.14, p >.05) or student engagement (β = .01, p > .05), but it did impact end-of-the-year supportive and structured teaching (β = -.26, p = .04). Likewise, teacher experience did not predict beginning-of-the-year supportive and structured teaching (β = .06, p >.05) or student engagement (β = .02, p > .05), but it was significantly associated with end-of-the-year supportive and structured teaching (β = -.31, p = .02). Lastly, observation year did not predict a significant portion of the variance for beginning-of-the-year supportive and structured teaching (β = -.03, p >.05), end-of-the-year supportive and structured teaching (β = .06, p > .05), or student engagement (β = -.02, p > .05).

Supplemental Exploratory Analyses

The proposed model hypothesized longitudinal relationships between teacher support and teacher structure, and between teacher behaviors and student engagement. Results only supported the longitudinal relationship between beginning-of-the-year supportive and structured teaching and mid-year student engagement, and no other longitudinal relationships were observed. To further characterize the proposed associations between constructs of interest, the study author examined associations within the same timepoint. These exploratory analyses provided information on whether cross-sectional associations were consistent with those observed in other studies. Furthermore, while study hypotheses posited that the proposed associations between constructs would

build over time, cross-sectional associations may indicate that reciprocity between teacher and student behaviors also occurs acutely within specific moments in time. These supplemental exploratory analyses examined the cross-sectional associations between the second-order supportive and structured teaching factor and the subscale score representing teacher engagement for timepoint 1 (beginning-of-year), timepoint 2 (mid-year), and timepoint 4 (end-of-year; Table 8). Covariates representing grade level, teacher experience in education, and observation year were included in the model. Within timepoints 1 and 2, supportive and structured teaching was significantly associated with student engagement (timepoint 1: $\beta = .50$, $p < .01$; timepoint 2: $\beta = .67$, $p < .01$). At timepoint 4, supportive and structured teaching did not predict student engagement ($\beta = .47$, $p > .05$).

Discussion

The current study examined associations between teacher and student behaviors as indicators of teacher-student relationships, a key element of school climate. This work aimed to broaden the evidence base for the core components of supportive and structured teaching to include teachers in classrooms with predominantly Black student populations. This study also assessed the reciprocal nature of teacher and student behaviors, where student engagement might explain continued teacher support and structure over the course of the school year. To achieve these aims, the current study used a longitudinal design and teacher observations in classrooms with primarily Black student populations across four public charter schools in New Orleans. Results supported the first hypothesis, which held that beginning-of-the-year supportive and structured teaching would positively predict mid-year student engagement when controlling for grade level and

teacher experience in education. The second hypothesis further examined student engagement as a positive mediator between beginning-of-the-year and end-of-the-year supportive and structured teaching, above and beyond what could be accounted for by teacher experience in education and grade level. Results did not support a direct effect between supportive and structured teaching at the beginning of the year and at the end of the year, nor an indirect effect of mid-year student engagement.

Teacher Support and Structure as a Predictor of Student Engagement

As hypothesized, beginning-of-the-year support and structure predicted mid-year student engagement. This finding aligns with previous studies utilizing primarily White study samples suggesting a long-term impact of teacher behaviors on student engagement (e.g., Buyse et al., 2008; Skinner et al., 2008; Way, 2011). Wang and Eccles (2013) found that this effect persists even as students move to the next grade, as their study showed that middle school students' seventh grade perceptions of four domains of classroom climate (structure, opportunities for choice, relevant instruction, and support from peers and teachers) predicted their engagement at the end of eighth grade. The current study suggests that the core components of the supportive and structured teaching construct can likewise promote positive outcomes among Black students.

The identified association occurred across time, as teacher behaviors were measured at timepoint 1 and student behaviors were measured at timepoint 2. This longitudinal association suggests that the behaviors teachers demonstrate at the start of the year do impact student behaviors, and that students' overall engagement in class may be a response to patterns of teacher behavior. Within the theoretical framework underlying the study, this finding extends the notion that adults demonstrating support

and structure toward students can promote one of the behavioral indicators of secure relationships: positive engagement. The previous literature examining school climate has identified disparities in student perceptions of school climate that vary by student racial and ethnic background, although the associations between school climate and positive student outcomes are consistent across diverse student populations (Konold, Cornell, Shukla, and Huang, 2017). The current study highlights that the core components of supportive and structured teaching are effective for promoting engagement among Black students, just as they have been shown to do for White students. Taken together with the existing literature, findings reinforce the idea that students largely respond well to positive regard and enjoyment from teachers, warmth, consistent structures, and clear expectations (Ford & Sassi, 2014; Gregory & Weinstein, 2008; Nelson, 2016; Walker, 2009; Ware, 2006; Wentzel 2002).

This finding affirms the existing literature indicating that a teaching approach that combines warmth and demandingness leads to positive outcomes for students. The theoretical foundation for the current study draws from ecological systems theory, self-determination theory, and attachment theory, positing that student-teacher relationships influence student behaviors, and that supportive and structured teaching is one aspect of teacher behaviors that promotes positive student outcomes through meeting students' needs for autonomy, competence, and relatedness. This finding builds upon empirical studies supporting these theories in several important ways. First, it establishes that across grade levels, supportive and structured teaching promotes student engagement. This is important to consider in light of previous studies indicating that middle school students experience more structured, but less supportive teaching approaches (Bergin &

Bergin, 2009; Furrer & Skinner, 2003; Hargreaves, 2000; Roorda et al., 2011; Way, Reddy, & Rhodes, 2007). The current study challenges the idea that older students require more structure and less support, instead indicating that a consistent blend of support and structure is appropriate across developmental levels. Second, the current study supports the positive association between supportive and structured teaching and student engagement in classrooms of predominantly Black students. Despite the breadth of literature indicating that supportive and structured teaching is effective in White or diverse student samples, classrooms serving Black students have not been represented in quantitative studies examining this approach. Furthermore, the current study established this finding with a unique teacher sample, which represented a diversity of racial and ethnic groups who were considered early-career educators. Third, beginning-of-the-year supportive and structured teaching showed an impact on students at mid-year, indicating that teachers' effects on their classroom and students begins right when the school year starts. Authoritative teaching and the warm demander frameworks highlight the importance of a consistent blend of support and structure; findings from the current study indicate that it is necessary for teachers to demonstrate that consistency right from the beginning in order to promote student engagement.

Engagement Did Not Mediate Supportive and Structured Teaching Across Time

Contrary to hypothesis 2, results demonstrated that timepoint 1 supportive and structured teaching was not associated with timepoint 4 supportive and structured teaching, and that there was no mediating effect of mid-year student engagement. This finding contradicts previous studies suggesting an association between student engagement and end-of-year teacher support and structure (Pelletier et al., 2002), as well

as studies examining similar models of feedback between teacher and student behaviors over time (e.g., Pitzer & Skinner, 2017; Skinner & Belmont, 1993).

Despite limited longitudinal effects in the structural equation models assessed, supplemental exploratory analyses showed that supportive and structured teaching was positively associated with student engagement within timepoint 1 and timepoint 2. These results reflect those from cross-sectional models examining relationships between teacher behaviors and student engagement in less diverse samples. One such study, which also used the ASSIST Proactive Behavior Management Subscale and ASSIST Student Compliance subscale, found the same cross-sectional association (Larson et al., 2018). Their study sample differed from the current study, as it comprised 274 teachers across 18 elementary and middle schools who collectively served a student population that was approximately 40% White, 34% Black, and 13% Hispanic or Latino. Taken together, teacher and student behaviors may be more sensitive to the dynamic nature of the classroom and the acute interactions occurring in the day-to-day rather than to patterns of behavior that develop across the school year.

Despite these findings, timepoint 4 supportive and structured teaching was not significantly associated with timepoint 4 student engagement. While teacher behaviors and student behaviors may be sensitive to acute interactions and respond accordingly, these direct responses may be increasingly drowned out as the school year progresses. School climate constructs such as teacher support and structure are typically represented in the literature as stable and constant, measured at a single timepoint, but it is likely that school climate fluctuates as a function of specific events occurring in the classroom as well as aggregated impacts within and outside of the classroom (Evans et al., 2009).

Results suggest that across the school year, many factors at the levels of individual students and teachers, the broader classroom dynamic, and the school all could impact teacher and student behaviors above and beyond the patterns of support and structure established at the beginning of the year.

The current study did not identify an association between beginning-of-the-year supportive and structured teaching and end-of-the-year supportive and structured teaching, which provides an interesting consideration for the application of ecological systems theory to understanding teacher behavior over the course of the school year. The absence of this association suggests that teachers' behaviors are fluid over the course of the year rather than stable. The relative inexperience of the teacher sample in the current study may partly explain this result, as new teachers likely intend to grow and develop their craft over the course of the school year. However, the unique influence of control variables on end-of-the-year teacher behaviors, but not on teacher behaviors earlier in the year, also suggests that as the school year goes on, student and teacher characteristics begin to have a greater impact on teachers than they did at the year's outset. Teachers may begin the year with a firm commitment to a clear vision of their practice, but this vision may shift in response to ecological influences such as their relationships with their students, decisions made by their administrators, and the organizational climate of their school. The current study looked at student engagement as one feature within the classroom that may influence teacher behavior. However, results from the current study did not identify this hypothesized association, and it may be that other factors present in the larger school environment carry more influence on teacher behavior than factors within the classroom. Further study is needed to identify what features of teachers'

environments outside of the classroom contribute to the relative stability and instability of their behaviors as the year goes on.

Results and Characteristics of the Teacher Sample

Characteristics of the unique teacher sample may have impacted the stability of constructs and associations over the course of the school year. Our sample consisted primarily of young female teachers who had five or fewer years of experience working in education following their training through alternative certification programs. Considering teacher support and teacher structure separately, mean ratings across timepoints for teacher support were moderate, and mean ratings for teacher structure were generally slightly lower. Given the accelerated yet abbreviated training that teachers in alternative certification programs often receive before leading their own classroom, it may be the case that young teachers from such programs receive limited strategies for building positive structures in their classrooms and are less equipped to do so. Interestingly, by the end of the year, teachers with five or fewer years working in education (who comprised much of the teacher sample) had significantly higher ratings for structure at the end of the year than teachers with more experience. In this way, there may be an effect among newer teachers where on-the-job training, as well as time spent with their students, promotes a tendency toward more structure by the end of the year. Though promising when paired with teacher support, there is likely to be limited effectiveness for teachers who introduce new structures once the school year has already begun. Alternatively, results may also indicate that newer teachers may begin the year with supportive and structured approaches but gravitate toward a more punitive, authoritarian approach as the year goes on. Fry (2007) reported that newer teachers often experience inadequate

support from their schools when they begin teaching. This lack of support may fuel a tendency toward traditional classroom management methods in which the overarching goal is student compliance rather than connection (Veenman, 1984).

Results also indicated that many demographic and professional characteristics were associated with timepoint 4 ratings of teacher support and structure. Timepoint 4 teacher support and structure ratings were higher among elementary teachers than middle school teachers. These results may reflect the greater emphasis and supports for elementary teachers to demonstrate support and structure through formal behavior systems. Although universal classroom supports for behavior such as Positive Behavior Interventions and Supports (PBIS) and social-emotional learning (SEL) have been developed across grade levels, middle school structures such as different teachers for each subject introduce complexity into maintaining such systems across classrooms. This finding is commensurate with previous studies such as the work of Eccles and colleagues (1993), who reported that teacher control and discipline are heightened in middle school classrooms in comparison to elementary school classrooms.

It is important to discuss that end-of-year teacher ratings favored elementary classrooms, in which White female teachers comprised a larger portion of the teacher staff than at the middle school level. Overall, timepoint 4 teacher support and structure ratings were also higher among White teachers than Black teachers and teachers identifying with multiple racial or ethnic groups. Results must be considered in the context of the observation team, which was also comprised of primarily White females, and the potential impact of observer bias, as well as measurement bias, on ratings that became more pronounced at the end of the year. Improved inter-rater reliability over the

course of the school year indicated increased alignment within the observation team. Given the heightened ratings for both support and structure for White teachers at the end of the year, it is possible that the team aligned to a narrow authoritative teaching framework when operationalizing teacher behaviors (see Table 1).

Reviewing the extant literature that reports on development and application of the scales used in this study provides valuable insights for areas of potential bias in the current study. While the CIS and ASSIST measures have been used in diverse samples across some studies, they were not developed with the demographics of the current study in mind. In particular, the CIS was developed for early childhood settings and drew from classrooms with primarily White students. As such, its subscales distinguish support, as defined by White norms, and penalizes harshness, also defined by White norms. For example, direct, ostensibly critical feedback can be valued as a supportive behavior among some Black students and teachers. And, while the ASSIST has been used in more diverse settings (e.g., Larson et al, 2018), studies examining its validity have not specifically examined classrooms serving Black students. Future studies that expand from the core components of supportive and structured teaching to observing characteristics of the warm demander pedagogy might see stronger associations with engagement, as well as with more tangible outcomes such as achievement (Sandilos et al., 2017). It should be noted, however, that the original ASSIST Student Compliance subscale measured quite well as an indicator of student engagement. Measurement bias in this study may be more apparent in the measures capturing teacher behaviors than student behaviors.

Overall, the current study set out to measure the foundational characteristics of supportive and structured teaching that are consistent across the authoritative teaching

and warm demander frameworks, which were identified to include warmth, positive regard for students, high expectations, and proactive, consistent behavioral management. However, applying the warm demander framework requires expanding the conceptualization of supportive and structured teaching to a definition that is rooted in the culture of the teacher and students in the classroom. Findings from the current study provide a basis for comparison to student and teacher outcomes when teams in future studies define and observe for characteristics of supportive and structured teaching that incorporate culturally validated features of teacher and student behavior. In support of an expanded conceptualization for supportive and structured teaching, authors such as Skinner and Belmont (1993) suggest that there are also dimensions within support and within structure that add complexity to the measurement of these constructs. For example, they distinguished emotional support, which aligns with support as conceptualized in this study, from autonomy support, or giving opportunities to grow independently. Although the current study included items reflecting both ideas ("seems enthusiastic about the children's activities and efforts"; "encourages children to exhibit prosocial behavior"), they were compiled into one core measure of support. Measures that more thoroughly capture the dimensionality of support and structure may yield different results than those found in the current study.

Limitations and Future Directions

The current study was impacted by several limitations that require discussion. The first limitation relates to the apparent instability of teacher behaviors over time, as analysis of the full model included constrained factor loadings for timepoint 1 and timepoint 4 teacher support, and for timepoint 1 and timepoint 4 teacher structure. This

decision was grounded in the theory that the same constructs were being measured by the tools for measuring teacher support and teacher structure, and that this measurement would be robust to time. This idea, termed longitudinal measurement invariance, is conceptually sound and was reflected in confirmatory factor analysis results indicating adequate fit of the measurement models for support and structure across timepoints. However, longitudinal measurement invariance may not hold when major transitions occur between timepoints in which data is collected using the same measure (Liu et al., 2017). Given that teacher support and teacher structure were examined at the beginning of the year, within the first month of school, then within the last month of school, many correlates of time had the potential to impact teachers' behaviors. Considering that teachers in the current study were relatively inexperienced, the course of one school year provides great opportunities for growth and adaptation of their practice as they gain experience in the classroom. Within this context, teachers may have broadened their repertoire of supportive and structured teaching behaviors as the year went on, and study measures may not have been robust to these potential changes.

The current study experienced limitations related to limited power and measurement concerns for the teacher support and teacher structure constructs. Regarding teacher support, the CIS Teacher Sensitivity measure was developed in early childhood settings. Although items were slightly adapted to allow for observation of school-aged settings, the measure has not been validated for this purpose. Although initial validation studies of the CIS identified four dimensions for caregiver interactions, recent studies suggest that the measure may represent one larger construct of teacher warmth, and that the apparent distinctions between dimensions are only due to some scales comprising

negatively worded items. Work by Biales (2018) and Pessanha and colleagues (2017) suggests reverse coding these items and including all twenty-six items in an analysis of teacher support. However, power concerns for the current study precluded testing a twenty-six-item model for improved fit. The current study also struggled to determine adequate fit for the factor structure of the teacher structure construct, as several items on the scale posed issues. Per Kline (2011), poor measurement models preclude adequate structural models. Various steps were taken to improve the measurement model, including confirmatory factor analyses, addressing zero-frequency cells, and parceling. However, decisions for improving the measurement model were generally limited by power.

The sample size of the current study comprised an additional, significant limitation. The current study assessed a complex model that iteratively reduced parameter estimates where possible. However, SEM guidelines recommend ten participants for each planned parameter (Schreiber et al., 2006). The current study had a sample of only 137 teachers to test a model with 70 free parameters. Furthermore, reducing parameters in a low-powered sample often meant losing a large portion of the data and variance. For example, the teacher experience in education covariate was collapsed from five response categories into two in order to address zero frequency cells, but the variable lost its normal distribution in becoming dichotomous. Low power precluded examining several further potential parameters that may more adequately characterize the interplay between teacher support and structure and student engagement, and which future well-powered studies might examine further.

The grade level covariate in the current study were significantly associated with end-of-year supportive and structured teaching, but not beginning-of-the-year supportive and structured teaching. The inconsistency of this result was surprising considering the breadth of developmental levels and classroom norms from kindergarten to eighth grade. Future studies should continue to consider and work to operationalize key features of these constructs across grades. In a well-powered study, looking at results by individual grade level may lend insight into these features and contribute a flexible framework for operationalization of supportive and structured teaching. There may also be alternate measures of observed teacher support, teacher structure, and student engagement that more adequately characterize these behaviors in school-level settings. The current study used scales that have been selected by previous studies to represent these constructs, but alternative subscales of the ASSIST and use of the full CIS may comprise a better-fitting measurement model for the data collected in this study.

Finally, quantitative studies should continue to follow the lead of qualitative work in examining the warm demander pedagogy in classrooms serving Black students. Future studies in this area should move to measure and address observer bias in teacher observations as well as validate observation coding schemas for the warm demander construct. One method of doing so is using multi-informant validation, where observations are examined for their correspondence to teacher and student reports of their interactions in the classroom. Larson and colleagues (2018) found significant associations between observed teacher behaviors and student engagement, but not between teacher reports of their own behaviors and perceived student engagement. From these results and others, it is evident that there is a wealth of insight to be gained from identifying and

exploring potential convergence and divergence of stakeholder perceptions of behavior from observed behavior.

Conclusion

In sum, the association between supportive and structured teaching and Black student engagement is grounded in theory and ethnographic study but necessitates further systematic assessment. Studies that systematically assess outcomes of supportive and structured teaching for Black students are limited and have produced mixed results. Studies specifically examining engagement are even less common, limited by the demographics of their study samples, and variable in their conceptualizations and measurement of key constructs. The current study builds upon the existing literature by using in-vivo observations to assess the association between teacher support, teacher structure, and student engagement in classrooms of Black students. Results from the current study reflect important considerations for teachers who are new to the field of education and working in public charter schools with a predominantly Black student population. Findings tell us that teacher support and structure, as defined without cultural context, have limited implications for students across time, but do impact their day-to-day interactions with students. Regarding the "relational zone" put forth in self-determination theory as the state in which students are most optimally engaged when their needs for autonomy, competence, and connectedness are met, results suggest that student behaviors are sensitive to how acutely these needs are met, rather than to a pattern of whether these needs have been met across time.

Regarding our understanding of the larger construct of school climate, teacher-student behaviors as measured in the current study comprise just one part of the elements

comprising students' experiences of school. The current study suggests that within this one dimension of school climate, students are sensitive to the day-to-day occurrences in the classroom. Over time, many influences likely combine with the impact of teacher behaviors to comprise their school climate perceptions. It is the task of future studies to capture and examine those influences in concert with each other.

Furthermore, the current study demonstrates the necessity of expanding how teacher support and structure are considered in the literature. Though results from the current study supported a narrow framework of the core components of support and structure for student outcomes, the warm demander literature highlights that we must evaluate the presence of support and structure flexibly. It is critical to continue drawing from the valuable findings of qualitative work that suggest and define what teacher support and structure look like across cultural contexts, so that we can be sure we are conducting studies with cultural relevance. Black students are underrepresented in the quantitative literature but comprise 16% of public school students in the United States (National Center for Education Statistics, 2017). We need to expand how we consider and operationalize these constructs that have been promoted as best practices, so that we may thoroughly and effectively serve all students.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Oxford, England: Lawrence Erlbaum.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179-211.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*(5), 888-918.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. *The Handbook of Attitudes*, 173-222.

Anderson, J., & Gerbing, D. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411-423.

Arnett, J. (1989). Caregivers in day-care centers: Does training matter?. *Journal of Applied Developmental Psychology*, *10*(4), 541-552.

Ayán, M., Ramirez, M., & Díaz, M. (2009). Path analysis models versus latent variable models: Examples from educational and health approaches. *International Journal of Hispanic Psychology, 2*(2), 91-110.

Baker, C. N., Brown, S. M., Wilcox, P. D., Overstreet, S., & Arora, P. (2016). Development and psychometric evaluation of the Attitudes Related to Trauma-Informed Care (ARTIC) scale. *School Mental Health*, *8*(1), 61-76.

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child development*, 887-907.

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs, 75*(1), 43-88.

Baumrind, D. (1968). Authoritarian vs. authoritative parental control. *Adolescence*, *3*(11), 255-272.

Bergin, C., & Bergin, D. (2009). Attachment in the classroom. *Educational Psychology Review*, *21*(2), 141-170.

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and white adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, *109*(4), 532-545.

Bowlby, J. (1969). Attachment and loss v. 3 (Vol. 1). *Random House.*

Brody, G. H., & Flor, D. L. (1998). Maternal resources, parenting practices, and child competence in rural, single-parent African American families. *Child Development*, *69*, 803–816.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513-531.

Buyse, E., Vershueren, 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,* 367-391.

Chao, R. K., & Otsuki-Clutter, M. (2011). Racial and ethnic differences: Sociocultural and contextual explanations. *Journal of Research on Adolescence*, *21*(1), 47-60.

Cohen, J., Mccabe, E.M., Michelli, N.M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record, 111*(1), 180-213.

Cole, S. F., Eisner, A., Gregory, M., & Ristuccia, J. (2013). *Helping traumatized children learn, volume 2: Creating and advocating for trauma-sensitive schools*. Massachusetts Advocates for Children.

Collins, L.M., Schafer, J.L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330-351.

Colwell, N., Gordon, R. A., Fujimoto, K., Kaestner, R., & Korenman, S. (2013). New evidence on the validity of the Arnett Caregiver Interaction Scale: Results from the early childhood longitudinal study-birth cohort. *Early Childhood Research Quarterly*, *28*(2), 218-233.

Cooper, P. M. (2003). Effective white teachers of black children: Teaching within a community. *Journal of Teacher Education*, *54*(5), 413-427.

Corbett, D., & Wilson, B. (2002). What urban students say about good teaching. *Educational Leadership*, *60*(1), 18-23.

The Cowen Institute (2015). NOLA by the numbers: 2014 School performance scores. New Orleans, LA: Tulane University Cowen Institute for Public Education Initiatives. Retrieved from www.coweninstitute.org.

Deci, E.L., Vallerand, R.J., Pelletier, L.G., & Ryan, R.M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*(3&4), 325-346.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review*, *95*(2), 158-165.

deJung, J. & Duckworth, K. (1986). High school teachers and their students' attendance: Final report. Eugene: University of Oregon Center for Education Policy and Management, College of Education. (ERIC Document Reproduction Service No. ED 266 557).

Delpit, L. (2006). *Other people's children: Cultural conflict in the classroom*. The New Press.

Dever, B. V., & Karabenick, S. A. (2011). Is authoritative teaching beneficial for all students? A multi-level model of the effects of teaching style on interest and achievement. *School Psychology Quarterly*, *26*(2), 131-144.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, *48*(2), 90-101.

Ford, A. C., & Sassi, K. (2014). Authority in cross-racial teaching and learning (re) considering the transferability of warm demander approaches. *Urban Education*, *49*(1), 39-74.

Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 71*(1), 59-109.

Fry, S.W. (2007). First-year teachers and induction support: Ups, downs, and in-betweens. *The Qualitative Report, 12*(2), 216-237.

Furman, W., & Buhrmester, D. (2009). Methods and measures: The network of relationships inventory: Behavioral systems version. *International Journal of Behavioral Development*, *33*, 470-478.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148-162.

Gay, G. (2002). Culturally responsive teaching in special education for ethnically diverse students: Setting the stage. *Qualitative Studies in Education*, *15*, 613-629.

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Gregory, A., Cornell, D., Fan, X., Sheras, P., Shih, T. H., & Huang, F. (2010). Authoritative school discipline: High school practices associated with lower bullying and victimization. *Journal of Educational Psychology*, *102*(2), 483-496.

Gregory, A., & Cornell, D. (2009). "Tolerating" adolescent needs: Moving beyond zero tolerance policies in high school. *Theory into Practice*, *48*(2), 106-113.

Gregory, A., & Weinstein, R. S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology*, *46*(4), 455-475.

Halpin, A., & Croft, D. (1963). The organizational climate and individual value systems upon job satisfaction. *Personnel Psychology*, *22*, 171-183.

Harel, O., Stratton, J., & Aseltine, R. (2011). Designed missingness to better estimate efficacy of behavioral studies (Technical Report 11-15). Storrs, CT: Department of Statistics, University of Connecticut.

Hargreaves, A. (2000). Mixed emotions: Teachers' perceptions of their interactions with students. *Teaching and Teacher Education*, *16*(8), 811-826.

Hindman, A. H., Pendergast, L. L., & Gooze, R. A. (2016). Using bifactor models to measure teacher–child interaction quality in early childhood: Evidence from the Caregiver Interaction Scale. *Early Childhood Research Quarterly*, *36*, 366-378.

Hoge, D. R., Smit, E. K., & Hanson, S. L. (1990). School experiences predicting changes in self-esteem of sixth-and seventh-grade students. *Journal of Educational Psychology*, *82*(1), 117-127.

Howard, M. (2016). A Review of Exploratory Factor Analysis (EFA) Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? International Journal of Human-Computer Interaction, 32(1), 51.

Hoy, W. K. (1990). Organizational climate and culture: A conceptual analysis of the school workplace. *Journal of Educational and Psychological Consultation*, *1*(2), 149-168.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1 - 55.

Irvine, J., & Fraser, J. (1998, May 13). "Warm demanders": Do national certification standards leave room for the culturally responsive pedagogy of African American teachers? Education Week, p. 56.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., ... & Velez, E. D. V. (2015). The Condition of Education 2015. NCES 2015-144. *National Center for Education Statistics*.

Kleinfeld, J. (1975). Effective teachers of Eskimo and Indian students. *The School Review*, *83*(2), 301-344.

Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling (2nd Ed.).* New York, NY, US: Guilford Press.

Konold, T., Cornell, D., Shukla, K., & Huang, F. (2017). Racial/ethnic differences in perceptions of school climate and its association with student engagement and peer aggression. *Journal of Youth and Adolescence*, *46*(6), 1289-1303.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology*, *100*(1), 96-104.

Larson, K. E., Pas, E. T., Bradshaw, C. P., Rosenberg, M. S., & Day-Vines, N. L. (2018). Examining how proactive management and culturally responsive teaching relate to student behavior: Implications for measurement and practice. *School Psychology Review*, *47*(2), 153-166.

Little, T. D. (2013). Overview and foundations of structural equation modeling. In Kenny & Little (Eds.), *Longitudinal structural equation modeling* (pp. 20-25). Guilford Press.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural equation modeling*, *9*(2), 151-173.

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, *39*(2), 151-162.

Little, T.D., & Rhemtulla, M. (2013). Planned missing data designs for developmental

researchers. *Child Development Perspectives, 7*(4), 199-204.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological methods*, *18*(3), 285-300.

Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, *22*(3), 486-506.

Losen, D. J. (2011). Discipline Policies, Successful Schools, and Racial Justice. *National Education Policy Center*.

Mason, C. A., Walker-Barnes, C. J., Tu, S., Simons, J., & Martinez-Arrue, R. (2004). Ethnic differences in the affective meaning of parental control behaviors. *Journal of Primary Prevention*, *25*(1), 59-79.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling*, *11*(3), 320-341.

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, *44*(6), 764-802.

McWayne, C. M., Owsianik, M., Green, L. E., & Fantuzzo, J. W. (2008). Parenting behaviors and preschool children's social and emotional skills: A question of the

consequential validity of traditional parenting constructs for low-income African Americans. *Early Childhood Research Quarterly*, *23*(2), 173-192.

Mitchell, M. M., & Bradshaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: The role of classroom management and exclusionary discipline strategies. *Journal of School Psychology*, *51*(5), 599-610.

Murray, C., Murray, K., & Waas, G. (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,* 49–61.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

Nelson, J. D. (2016). Relational teaching with Black boys: Strategies for learning at a single-sex middle school for boys of color. *Teachers College Record*, *118*(6), 1-30.

New Orleans Parents' Guide (2016). New Orleans parents' guide to public schools: Tenth edition. New Orleans, LA: New Orleans Parents' Guide. Retrieved from http://neworleansparentsguide.org/wp-content/uploads/2017/12/NOPG-Guide-to-Public-Schools2016.pdf.

Obidah, J. E., & Teel, K. M. (2001). *Because of the kids: Facing racial and cultural differences in schools* (Vol. 18). Teachers College Press.

Pas, E. T., Cash, A. H., O'Brennan, L., Debnam, K. J., & Bradshaw, C. P. (2015). Profiles of classroom behavior in high schools: Associations with teacher behavior

management strategies and classroom composition. *Journal of School Psychology*, *53*(2), 137-148.

Pellerin, L. A. (2005). Applying Baumrind's parenting typology to high schools: toward a middle-range theory of authoritative socialization. *Social Science Research*, *34*(2), 283-303.

Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, *94*(1), 186-196.

Penuel, W., Meyer, E., & Valladares, M. R. (2016). Making the Most of the Every Student Succeeds Act (ESSA): Helping States Focus on School Equity, Quality and Climate. *National Education Policy Center*.

Pettit, G. S., Bates, J. E., & Dodge, K. A. (1997). Supportive parenting, Ecological Context, and Children's Adjustment: A seven-Year Longitudinal Study. *Child development*, *68*(5), 908-923.

Pianta, R.C., & Steinberg, M.S. (1992). Teacher-child relationships and adjusting to school. In R.C. Pianta (Ed.), *New directions in child development (Vol. 57). Beyond the parent: The role of other adults in children's lives* (pp. 61–79). San Francisco: Jossey-Bass.

Pitzer, J., & Skinner, E. (2017). Predictors of changes in students' motivational resilience over the school year: The roles of teacher support, self-appraisals, and emotional reactivity. *International Journal of Behavioral Development*, *41*(1), 15-29.

Proctor, P. (1984). Teacher expectations: a model for school improvement. *The Elementary School Journal, 84,* 468-481.

National School Climate Council. (2007). *The School Climate Challenge: Narrowing the gap between school climate research and school climate policy, practice guidelines and teacher education policy*. Retrieved from http://www.schoolclimate.org/climate/advocacy.php

National School Climate Center (2017). *The 12 Dimensions of School Climate Measured.* Retrieved from http://www.schoolclimate.org/programs/documents/dimensions_chart_pagebars.pdf

Quin, D. (2017). Longitudinal and contextual associations between teacher–student relationships and student engagement: A systematic review. *Review of Educational Research*, *87*(2), 345-387.

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159-175.

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In *Handbook of Research on Student Engagement* (pp. 3-19). Springer, Boston, MA.

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

Rhemtulla, M. (2016, February 1). Population Performance of SEM Parceling Strategies Under Measurement and Structural Model Misspecification. *Psychological Methods*. Advance online publication.

Rogosa, D., Floden, R., & Willett, J. B. (1984). Assessing the stability of teacher behavior. *Journal of Educational Psychology*, *76*(6), 1000.

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.

Rubie-Davies, C. M. (2007). Classroom interactions: exploring the practices of high and low-expectation teachers. *British Journal of Educational Psychology, 77,* 289-306.

Rusby, J. C., Taylor, T. K., & Milchak, C. (2001). Assessing school settings: Interactions of students and teachers. *Eugene, OR: Oregon Research Institute*.

Rusby, J. C., Crowley, R., Sprague, J., & Biglan, A. (2011). Observations of the middle school environment: The context for student behavior beyond the classroom. *Psychology in the Schools*, *48*(4), 400-415.

Sandilos, L. E., Rimm-Kaufman, S. E., & Cohen, J. J. (2017). Warmth and demand: The relation between students' perceptions of the classroom environment and achievement growth. *Child development*, *88*(4), 1321-1337.

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23-74.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571-581.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*(4), 765-781.

Snyder, J., Reid, J., Stoolmiller, M., Howe, G., Brown, H., Dagne, G., & Cross, W. (2006). The role of behavior observation in measurement systems for randomized prevention trials. *Prevention Science*, *7*(1), 43-56.

Spilt, J. L., & Hughes, J. N. (2015). African American children at risk of increasingly conflicted teacher–student relationships in elementary school. *School Psychology Review*, *44*(3), 306-314.

Stuhlman, M. W., & Pianta, R. C. (2009). Profiles of educational quality in first grade. *The Elementary School Journal*, *109*(4), 323-342.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357-385.

Tobin, T. J., & Vincent, C. G. (2011). Strategies for preventing disproportionate exclusions of African American students. *Preventing School Failure: Alternative Education for Children and Youth*, *55*(4), 192-201.

Vasquez, J. (1989). Contexts of learning for minority students. The Education Forum, 52, 243–253.

Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research, 54*(2), 143-178.

Voight, A., Hanson, T., O'Malley, M., & Adekanye, L. (2015). The racial school climate gap: Within-school disparities in students' experiences of safety, support, and connectedness. *American Journal of Community Psychology*, *56*(3-4), 252-267.

Walker, J. M. (2008). Looking at teacher practices through the lens of parenting style: Three case studies at entry to middle school. *Journal of Experimental Education*, *76*, 218-240.

Walker-Barnes, C. J., & Mason, C. A. (2001). Ethnic differences in the effect of parenting on gang involvement and gang delinquency: A longitudinal, hierarchical linear modeling perspective. *Child Development*, 72, 1814 – 1831.

Wang, M. T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*(2), 315-352.

Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, *28*, 12-23.

Wang, M. C., & Gordon, E. W. (Eds.). (1994). *Educational resilience in inner-city America: Challenges and prospects*. Hillside, NJ: Lawrence Erlbaum.

Wang, M. T., & Peck, S. C. (2013). Adolescent educational success and mental health vary across school engagement profiles. *Developmental Psychology*, *49*(7), 1266-1276.

Ware, F. (2006). Warm demander pedagogy: Culturally responsive teaching that supports a culture of achievement for African American students. *Urban Education*, *41*(4), 427-456.

Way, S. M. (2011). School discipline and disruptive classroom behavior: The moderating effects of student perceptions. *The Sociological Quarterly*, *52*(3), 346-375.

Way, N., Reddy, R., & Rhodes, J. (2007). Students' perceptions of school climate during the middle school years: Associations with trajectories of psychological and behavioral adjustment. *American journal of community psychology*, *40*(3-4), 194-213.

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*(1), 287-301.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *25*(4), 41-78.

Wigfield, A., & Guthrie, J. T. (2000). Engagement and motivation in reading. *Handbook of Reading Research*, *3*, 403-422.

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

Yoon, J.S. (2002). Teacher characteristics as predictors of teacher-student relationships: Stress,

negative affect and self-efficacy. *Social Behavior and Personality, 30*, 485-494.

Zullig, K. J., Koopman, T. M., Patton, J. M., & Ubbes, V. A. (2010). School climate: Historical review, instrument development, and school assessment. *Journal of Psychoeducational Assessment*, *28*(2), 139-152.

Table 1.

Components of Authoritative Teaching and Warm Demander Approaches

Teacher Behavior	Authoritative Teaching	Warm Demander
Teacher Support CAN include…		
Frequent positive praise	X	X
Communicating openly with students, such as through providing rationale for learning objectives and behavioral expectations	X	X
Constructive immediate feedback that is usually neutral and sparingly negative	X	X
Unconditional positive regard for students as individuals; authentic interest in their well-being	X	X
Providing students with choices	X	X
Frequent negative feedback followed by warmth		X
Highlighting commonalities between students and teacher		X
Psychological control/coercion (e.g., threat of consequences)		X
Teacher Structure CAN include…		
Behavioral control via proactive rule-setting and rule-enforcement	X	X
Consistently high expectations for behavioral and academic performance	X	X
Adequate supervision	X	X
Direct, negative verbal monitoring of behavior		X
Indirect, warm acknowledgements of behavioral expectations that are or are not being met		X
Strategic use of irritation/hostility when expectations violated; balanced by responsivity		X

Note. Baumrind; Bondy & Ross, 2008; Brody & Flor, 1998; Chao & Otsuki-Clutter, 2011; Ford & Sassi, 2014; Gregory & Cornell, 2009; Nelson, 2016; Pellerin, 2005; Pettit, Bates, & Dodge, 1997; Sandilos, Rimm-Kaufman, & Cohen, 2017; Vasquez, 1989; Ware, 2006; Wentzel, 2002; Wilson, Pianta, & Stuhlman, 2007.

Table 2.

Demographic Characteristics of Teacher Sample

Demographic Category *N* (%)[a]		CMO 1		CMO 2	
	Overall *N* = 137)	School 1 (*n* = 45)	School 2 (*n* = 26)	School 3 (*n* = 29)	School 4 (*n* = 37)
Observation Year					
Cohort 1	**99 (72.3)**	24 (53.3)	26 (100.0)	29 (100.0)	20 (54.1)
Cohort 2	**38 (27.7)**	21 (46.7)	--	--	17 (45.9)
Gender					
Female	**80 (75.5)**	32 (82.1)	15 (68.2)	15 (68.2)	18 (78.3)
Male	**26 (19.0)**	7 (17.9)	7 (31.8)	7 (31.8)	5 (21.7)
Grade Category					
Elementary (K-4$^{th\ grade}$)	**71 (51.8)**	25 (55.6)	13 (50.0)	15 (51.7)	18 (48.6)
Middle (5th-8th grade)	**66 (48.2)**	20 (44.4)	13 (50.0)	14 (48.3)	19 (51.4)
Age Category					
18-24	**16 (15.1)**	8 (20.5)	1 (4.5)	--	7 (30.4)
25-34	**68 (64.2)**	26 (66.7)	17 (77.3)	15 (68.2)	10 (43.5)
35-44	**16 (15.1)**	3 (7.7)	4 (18.2)	6 (27.3)	3 (13.0)
45-54	**6 (5.7)**	2 (5.1)	--	1 (4.5)	3 (13.0)
Race/Ethnicity*					
Black/African American	**45 (42.5)**	13 (33.3)	8 (36.4)	16 (72.7)	8 (34.8)
White/Caucasian	**54 (50.9)**	24 (61.5)	11 (50.0)	4 (18.2)	15 (65.2)
Hispanic or Latino	**5 (4.7)**	--	2 (9.1)	1 (4.5)	2 (8.7)
Multiple races selected	**5 (4.7)**	2 (5.1)	2 (9.1)	1 (4.5)	--
Other	**2 (1.9)**	--	1 (4.5)	1 (4.5)	--

Note. Table continues onto next page.

Table 2 Continued.

Demographic Characteristics of Teacher Sample

Demographic Category *N* (%)[a]		CMO 1		CMO 2	
	Overall *N* = 137)	School 1 (*n* = 45)	School 2 (*n* = 26)	School 3 (*n* = 29)	School 4 (*n* = 37)
Education					
Completed college	**52 (49.1)**	22 (56.4)	7 (31.8)	14 (63.6)	9 (39.1)
Some graduate school	**23 (21.7)**	9 (23.1)	7 (31.8)	4 (18.2)	3 (13.0)
Completed graduate school	**31 (29.2)**	8 (20.5)	8 (36.4)	4 (18.2)	11 (47.8)
Years in Role					
<1	**35 (33.0)**	16 (41.0)	6 (27.3)	4 (18.2)	9 (39.1)
1-5	**51 (48.1)**	14 (35.9)	12 (54.5)	13 (59.1)	12 (52.2)
6-10	**16 (15.1)**	8 (20.5)	4 (18.2)	4 (18.2)	1 (4.3)
11-15	**1 (0.9)**	--	--	--	1 (4.3)
16-20	**3 2.8)**	1 (2.6)	--	1 (4.5)	--
Years in School					
<1	**50 (47.2)**	19 (48.7)	8 (36.4)	9 (40.9)	7 (30.4)
1-5	**51 (48.1)**	17 (43.6)	12 (54.5)	13 (59.1)	9 (39.1)
6-10	**5 (4.7)**	3 (7.7)	2 (9.1)	--	--
Years in Field					
<1	**18 (17.0)**	7 (17.9)	2 (9.1)	2 (9.1)	7 (30.4)
1-5	**54 (50.9)**	20 (51.3)	15 (68.2)	10 (45.5)	9 (39.1)
6-10	**25 (23.6)**	10 (25.6)	4 (18.2)	7 (31.8)	4 (17.4)
11-15	**5 (4.7)**	--	1 (4.5)	2 (9.1)	2 (8.7)
16-20	**2 (1.9)**	1 (2.6)	--	1 (4.5)	--
20+	**2 (1.9)**	1 (2.6)	--	--	1 (4.3)

[a]percentage of the total number of teachers who reported demographics (n = 106; 77% of the total study sample).

Table 3.

n by Randomly Assigned Timepoint and Timepoint Wave

Wave	**1**	**2**	**3**	**4**
1	52			
2	20	59		
3	11	13	52	
4	15	8	19	59

Table 4.

Demographic Characteristics of Observer Team

Demographic Category	n (%)
Gender	
Female	9 (90)
Male	1 (10)
Age Category	
18-24	3 (30)
25-34	6 (60)
35-44	1 (10)
Race/Ethnicity	
Black/African American	2 (20)
White/Caucasian	8 (80)
Number of Semesters on Observation Team	
0	1 (10)
1-2	5 (50)
3-4	1 (10)
5-6	3 (30)
Number of Semesters as Master Coder	
0	5 (50)
1-2	1 (10)
3-4	4 (40)
5-6	---
Years of Teaching Experience	
0	7 (70)
1-5 years	3 (30)
Years of Experience Working in Schools	
0	3 (30)
<1-5 years	4 (40)
6-10 years	3 (30)

Table 5.

Description of Study Variables

Construct	Measure	Time Point (s) used in current study	*N*
Years of Experience in Education[a], Grade Level[a], Race, Gender, Age, Education Background, Certification, Years in Role, Years at School	Teacher Demographic Information	N/A	106
Teacher Support	CIS Sensitivity Subscale	Timepoints 1 and 4	Timepoints 1 (*N* =52) and 4 (*N* = *59*)
Teacher Structure	ASSIST Proactive Behavior Management Subscale	Timepoints 1 and 4	Timepoints 1 (*N* =52) and 4 (*N* = 59)
Student Engagement	ASSIST Student Compliance Subscale	Timepoint 2	59

[a]Entered into model as dichotomous control variables (years of experience in education coded as 1= $\leq$ 5 years of experience, 2 = > 5 years of experience; grade level coded as 1 = K-4th grade; 2 = 5th-8th grade).

Table 6.

Correlations prior to Respecification

Variable	1	2	3	4	5	6	7	8
1. Teacher experience in education field (original)	—							
2. Grade Level	-.04	—						
3. Observation year[a]	.10	-.14	—					
Timepoint 1								
4. T1 Teacher Support subscale score (original)	.004	-.25	-.11	—				
5. T1 Teacher Structure subscale score (original)	.03	-.06	-.55	.74	—			
Timepoint 2								
6. Student Engagement subscale score	.03	-.13	-.11	.40	.03	—		
Timepoint 4								
7. T4 Teacher Support subscale score (original)	-.23	-.35	.14	.15	-.05	.41	—	
8. T4 Teacher Structure subscale score (original)	-.27	-.28	-.19	.36	.39	.32	.64	—

[a]Entered into model as dichotomous control variable (year 1 = 1; year 2 = 2).

Table 7.

Correlations Using Revised Factor Structures and Response Categories

Variable	1	2	3	4	5	6	7	8	9	10
1. Teacher experience in education field (recoded)[a]	—									
2. Grade Level (recoded)[a]	-.02	—								
3. Observation year[a]	0.03	-0.14	—							
4. Timepoint 1 Teacher Support subscale score (revised)	-.02	-0.17	-0.05	—						
5. Timepoint 1 Support Parcel 1	-.11	-0.12	0.06	0.90	—					
6. Timepoint 1 Support Parcel 2	0.03	-0.20	-0.14	0.82	0.49	—				
7. Timepoint 1 Support Parcel 3	0.02	-0.12	-0.02	.87	.80	.54	—			
8. Timepoint 1 Teacher Structure subscale score (revised)	0.12	-0.23	-0.23	.68	.44	.71	.55	—		
9. Timepoint 1 Structure Parcel 1	0.08	-0.21	-0.04	.54	.39	.45	.57	.62	—	
10. Timepoint 1 Structure Parcel 2	0.12	-0.19	-0.26	.61	.38	.67	.45	.96	.37	—
11. Timepoint 2 Student Engagement subscale score	-.01	-0.13	-0.11	0.43	0.5	0.22	0.39	0.18	0.31	0.09
12. Timepoint 4 Teacher Support subscale score (revised)	-0.93	-.33	0.09	0.10	0.20	-0.04	0.21	-0.03	-0.01	-0.04
13. Timepoint 4 Support Parcel 1	-0.22	-.30	0.15	-0.34	-0.17	-0.35	-0.13	-0.27	0.03	-0.35
14. Timepoint 4 Support Parcel 2	-0.23	-0.24	-0.01	0.28	0.37	0.08	0.25	0.03	-0.21	0.14
15. Timepoint 4 Support Parcel 3	-0.35	-.27	0.14	0.25	0.25	0.12	0.42	0.12	0.25	0.03
16. Timepoint 4 Teacher Structure subscale score (revised)	-0.27	-.30	0.09	0.29	0.38	0.09	0.32	0.12	0.09	0.10
17. Timepoint 4 Structure Parcel 1	-0.21	-.29	0.09	0.42	0.36	0.27	0.51	0.38	0.48	0.24
18. Timepoint 4 Structure Parcel 2	-0.26	-0.25	0.07	0.12	0.30	-0.05	0.11	-0.10	-0.20	-0.02

Note. Table continues onto next page.

Table 7 Continued

Correlations Using Revised Factor Structures and Response Categories

Variable	11	12	13	14	15	16	17	18
1. Teacher experience in education field (recoded)[a]								
2. Grade Level (recoded)[a]								
3. Observation year[a]								
4. Timepoint 1 Teacher Support subscale score (revised)								
5. Timepoint 1 Support Parcel 1								
6. Timepoint 1 Support Parcel 2								
7. Timepoint 1 Support Parcel 3								
8. Timepoint 1 Teacher Structure subscale score (revised)								
9. Timepoint 1 Structure Parcel 1								
10. Timepoint 1 Structure Parcel 2								
11. Timepoint 2 Student Engagement subscale score	—							
12. Timepoint 4 Teacher Support subscale score (revised)	0.29	—						
13. Timepoint 4 Support Parcel 1	0.32	.85	—					
14. Timepoint 4 Support Parcel 2	0.24	.82	.48	—				
15. Timepoint 4 Support Parcel 3	0.24	.89	.70	.62	—			
16. Timepoint 4 Teacher Structure subscale score (revised)	0.28	.65	.46	.61	.61	—		
17. Timepoint 4 Structure Parcel 1	[b]	.56	.43	.42	.61	.80	—	
18. Timepoint 4 Structure Parcel 2	0.28	.57	.38	.60	.49	.92	.50	—

[a]Entered into model as dichotomous control variables (years of experience in education coded as 1= ≤ 5 years of experience, 2 = > 5 years of experience; grade level coded as 1 = K-4th grade, 2 = 5th-8th grade; year 1 = 1; year 2 = 2).

[b]Correlation not calculated; pairwise deletion resulted in prohibitively small *n* which precluded calculation.

Table 8.

Results of Cross-Sectional Regression Analyses

Variable	B	*p*
T1 Supportive and Structured Teaching and Engagement	.50	.001
T2 Supportive and Structured Teaching and Engagement	.67	< .001
T4 Supportive and Structured Teaching and Engagement	.47	.169

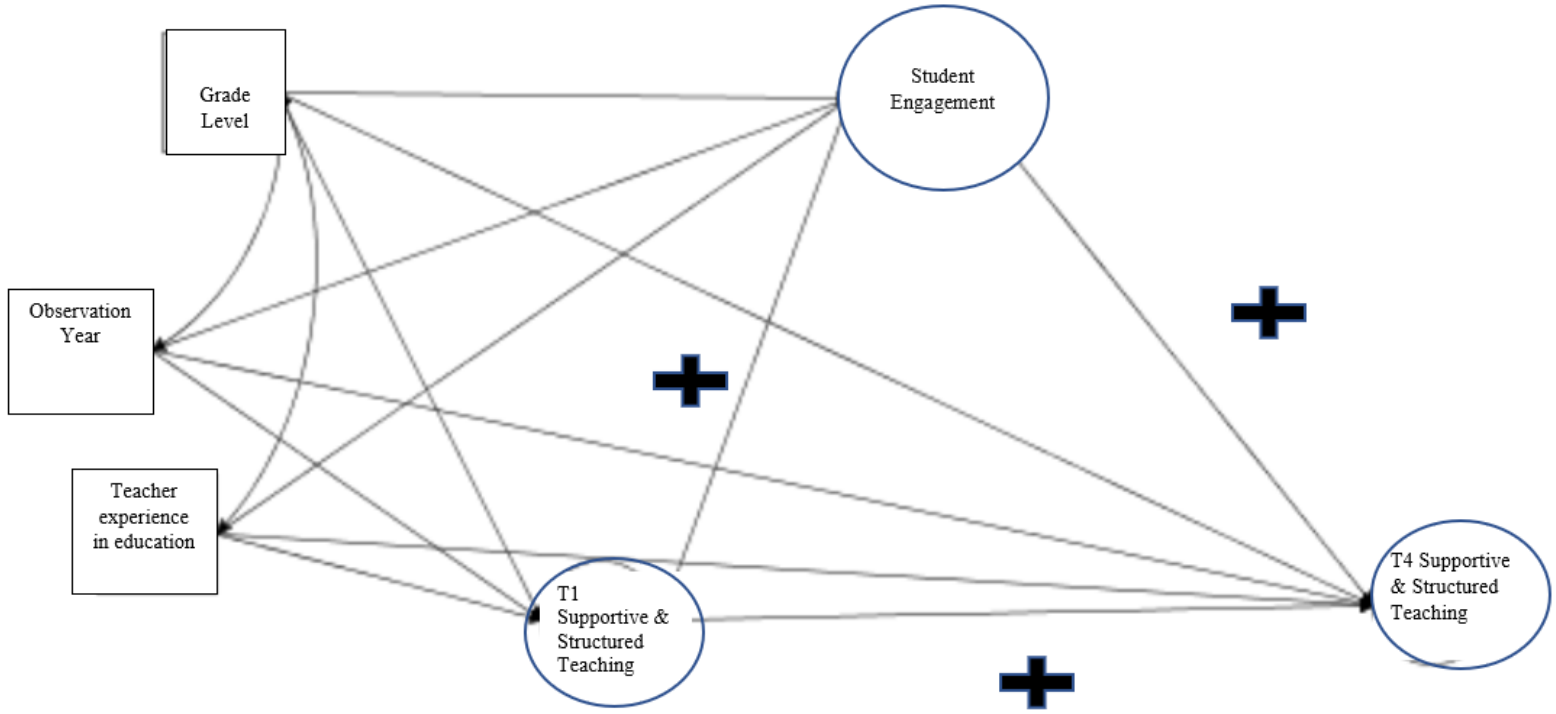

Figure 1. *Hypothesized Model*

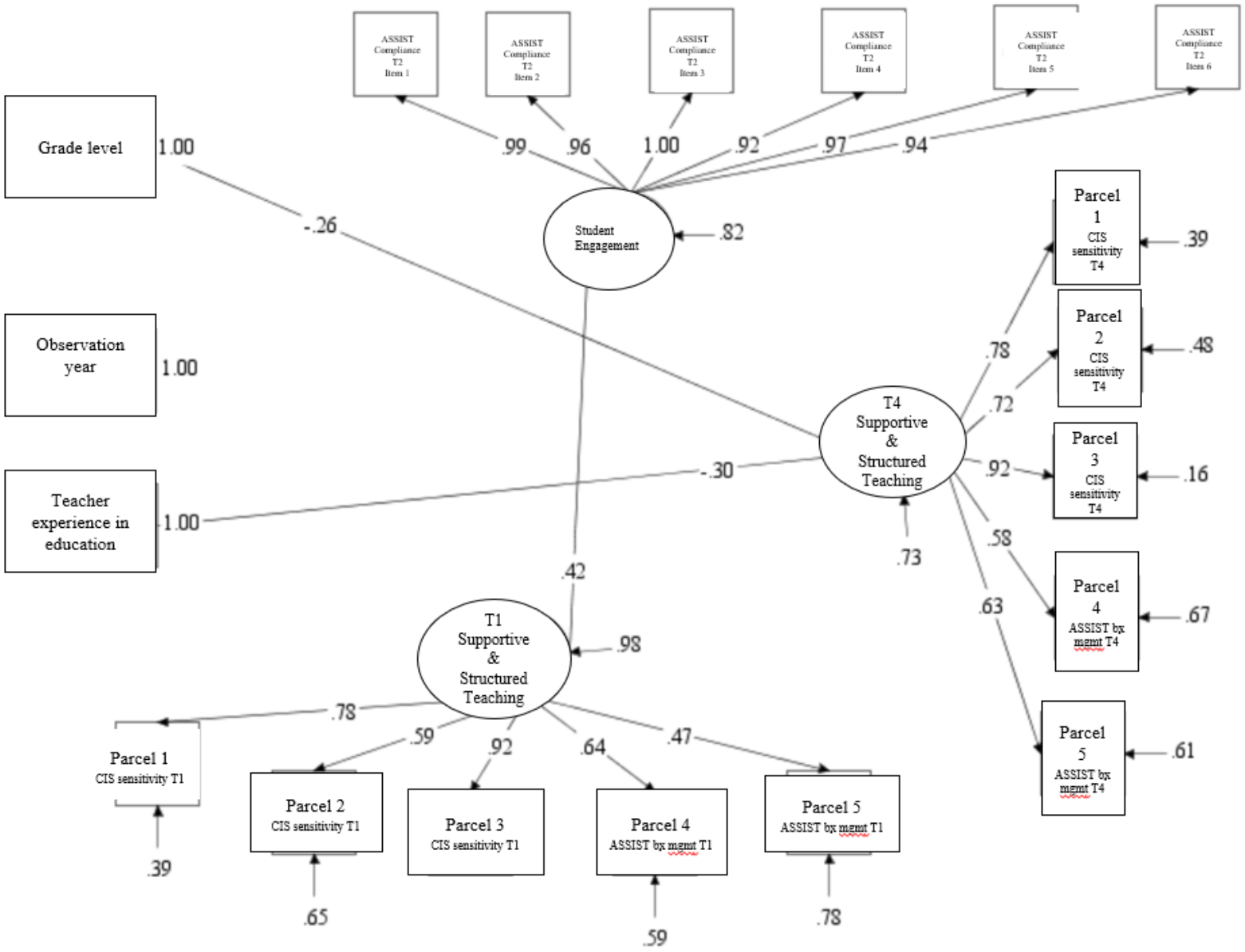

Figure 2. *Structural Equation Model Results*

Appendix

Specification of the Measurement Model

General steps taken for specification and respecification of the measurement model are described at the beginning of this section. Specific results and decision points for each construct in the model follow, presented by variable. The a priori model developed using the theoretical and empirical literature is presented in Figure 1.

One item from the Proactive Behavior Management Scale (Item 6: "When aggression occurred, teacher acknowledged student's feelings") presented limited variance that prohibited model analysis. Upon inspection, this limited variance was largely due to the rarity of student aggression occurring during observations. As a result, over 60% of data was missing from this item for both the Timepoint 1 Proactive Behavior Management Scale as well as Timepoint 4 Proactive Behavior Management scale. As such, this item was removed from the scale at both timepoints.

Subsequent attempts to analyze the model returned errors related to various parameters in the model. Several steps were taken to address initial errors. First, a Monte Carlo integration was employed. Integrations are necessary in mediation models with missing data in order to reduce model complexity (Muthén & Muthén, 1998-2017). The current study variables used varying scales, with teacher support, teacher structure, and student engagement using ordered-categorical scales in the initial model, and control variables using categorical scales. Furthermore, model complexity was increased due to the level of missingness in the dataset. Integration points were set at 3000, and covariance coverage was set at 0.05. Factor loadings were constrained across timepoint 1 and

timepoint 4 for teacher support and teacher structure. This step specified measurement invariance for the teacher support and teacher structure factors across time, where measurement at both timepoints was assumed to measure the same construct. Factor loading equality constraints are beneficial in datasets with sparse data. Additionally, teacher experience was collapsed into a dichotomous variable to address sparse data and reduce model parameters. The subsequent response categories for teacher experience grouped together teachers with five or fewer years of experience in education, and teachers with six or more years of experience.

Following these steps, confirmatory factor analyses were completed for all latent variables of interest to inform model respecification that could address continued parameter errors. Confirmatory factor analyses inform model respecification through assessing the factor structure of latent variables. Confirmatory factor analyses were completed because Mplus parameter errors referenced possible nonidentification of the structural equation model. Identification errors are related to an imbalance in observations and parameters, where models can have multiple solutions and a unique set of estimates cannot be derived. In order for structural regression models to be identified, in addition to the model needing zero or more degrees of freedom where observations equal or exceed parameters, it must also be possible to estimate the unique variances and covariances between the factors in a structural regression model. Identification errors in a structural regression model can be related to poor identification of the measurement model. Although confirmatory factor analysis is not required to identify a structural regression model, it can be used to inform respecification of the measurement model and improve the balance of parameters and observations (Kline, 2011). Given that the

measures used in the current study have been minimally evaluated in previous studies with similar teacher demographics, classroom demographics, and grade levels, confirmatory factor analyses provided the additional benefit of validating the measures in the study population.

Several indicators of model fit informed confirmatory factor analyses and subsequent revisions to the factor structures of latent variables. Statisticians recommend a holistic consideration of multiple fit indices in confirmatory factor analysis (Kline, 2011). For each analysis, the Chi-Square goodness-of-fit (X^2) statistic, was examined. The X^2 statistic provides a corresponding *p* value to assess model fit, where statistically significant *p* values are considered to indicate poor model fit. However, given that this statistic is sensitive to model complexity and sample size, Hu and Bentler (1999) and Marsh, Hau, and Wen (2004) recommend additional consideration of the RMSEA, CFI, and TLI global fit indices. An RMSEA of $\leq .05$, CFI $\geq .95$, and TLI $\geq .95$ generally indicate good model fit (Schermelleh-Engel, Moosbrugger, & Muller, 2003). These fit statistics are reported below for each of their respective variables. Factor loadings and communalities were considered for potential removal of item indicators within each scale, in conjunction with empirical and theoretical bases for their inclusion or exclusion. Generally, items with factor loadings less than 0.4 and communalities less than 0.5 were excluded (Howard, 2016). Each latent variable was scaled by setting the variance for one reference item to 1 (Kline, 2011). All confirmatory factor analyses used maximum likelihood estimation.

Following confirmatory factor analyses, parceling was employed to some variables of interest. Specific decisions about which variables to be parceled and the

subsequent creation of parcels are described below. Parcels are averaged item scores within a subscale that use the same rating scale. They are treated as continuous indicators of latent variables with generally improved reliability and distributions over individual items (Kline, 2011). The purpose of parceling is to reduce the complexity of structural models. This reduced complexity has several advantages for analysis of a structural equation model. First, in models analyzed using data from small samples, biases in Chi square fit statistics are reduced by model simplification. Second, the increased psychometric properties of parcels increase the likelihood that a model will converge and that parameter estimates will be stable. One caveat to this benefit is that parceling reduces the number of indicators per factor which can threaten model convergence and stability rather than improve them. Researchers should also employ caution in interpreting analyses of measurement models that use parceling, as estimated factor loadings can be biased and mask misspecification of the measurement model. For the current study's purpose of assessing the proposed structural model, parceling is an acceptable analytic tool (Little, Rhemtulla, Gibson, & Schoemann, 2013; Rhemtulla, 2016). However, the current study minimized its use of parceling by retaining item-level indicators for any latent factors whose original factor structure demonstrated acceptable model fit in confirmatory factor analyses, as described in more detail below. To select items to average into parcels, inter-item and residual correlations were examined for items within a scale, within a timepoint (Little et al., 2013; Rhemtulla, 2016). Where possible, parceling decisions followed recommendations from Matsunaga (2008) to create three parcels per factor, with two or more items per parcel. However, some final parcels did not meet these criteria if a factor was comprised of too few items. Results from confirmatory

factor analyses and decision points for parceling are reported below, organized by study variable. Item-level data screening for outliers, missingness, and violations of normality are also reported for items included in the revised factor structures. For all items with missing responses, missingness was calculated to represent less than 5% of data for that item. As such, missing data was addressed using substitution with the overall sample mean for that item (Kline, 2011). To evaluate for normality, item-level skewness and kurtosis were assessed using skewness and kurtosis statistics. Problematic values for skewness were considered those that exceeded ± 2, and problematic values for kurtosis were considered those that exceeded ± 7 (Curran, West, & Finch, 1996). Z scores were calculated to identify outlier responses, with z scores exceeding 3.29 considered outliers (Tabachnik & Fidel, 2001). Using this criterion, no outliers were identified in study variables. Following respecification of the measurement model for each latent construct, remaining zero-frequency cells were identified and addressed through collapsing response categories so that each response category had at least one observation in every pair of items in the variance covariance matrix. Zero frequency cells, where a response category is not endorsed for either item in a variance covariance matrix, are common in datasets with planned missingness and can impact the accuracy of parameter estimates (Liu et al., 2017; Savalei, 2011). Revised response categories were consistent for all items within a subscale and across time.

Timepoint 1 Teacher Support and Timepoint 4 Teacher Support

Item-level statistics and inter-item correlations for the original ten-item CIS Sensitivity subscale are presented in Table A1 (timepoint 1) and Table A2 (timepoint 4). A confirmatory factor analysis assessed Timepoint 1 Teacher support as a latent factor of

the original ten-item CIS Sensitivity Subscale (Table A3). The measured data demonstrated poor fit to this model across all indices examined ($\chi 2(35, N = 55) = 114.41$; $p < .05$; RMSEA=.21; CFI=.76; TLI=.64). One item with a factor loading under 0.4 (Item 8 - "encourages the children to try new experiences") was removed from the model. An additional item with a low factor loading (Item 19 – "encourages children to exhibit prosocial behavior such as sharing or helping") was considered for removal but retained for its theoretical relevance. Additionally, three items with communalities below 0.5 were removed: (Item 7 – "When the children misbehave, explains the reason for the rule they are breaking"; item 25 – "When talking to children, kneels to establish better eye contact"; item 16 – talks to children on a level they can understand), all of which may decrease in theoretical relevance and variance across teachers outside of the early childhood settings in which the CIS has been most rigorously validated. The subsequent six-item factor structure was assessed for model fit and showed adequate to acceptable fit for the CFI (.95) and TLI (.91) (Table A4). Although the six-item model was not acceptable for the X^2 statistic ($\chi 2(9, N = 52) = 17.44$; $p = .04$) or RMSEA (.14), these indices have been reported to be more sensitive to sample size. Subsequently, this model was retained. Given that measurement invariance was assumed for timepoint 1 teacher structure and timepoint 4 teacher structure, the six-item model that demonstrated the best fit to timepoint 1 CIS Sensitivity subscale data was assessed for its fit to timepoint 4 CIS Sensitivity subscale data (Table A5). Consistent with results for the six-item model using timepoint 1 data, the six-item model run with Timepoint 4 data demonstrated adequate to acceptable fit for the CFI (0.93) and TLI (.882), and poor fit with the X^2 statistic ($\chi 2(9, N = 59) = 18.51$; $p = .03$) and RMSEA (.14).

Following the confirmatory factor analysis and revision of the factor structure, crosstabs analyses identified remaining zero-frequency cells in the revised six-item CIS Sensitivity subscale at both timepoint 1 and timepoint 4. Subsequently, response categories for the six items comprising this scale in Timepoint 1, as well as the six items comprising this scale in the Timepoint 4, were collapsed to reflect the following scale: 1 = not at all true; 2 = somewhat true; and 3 = quite a bit or very much true. With the revised and recoded item-level indicators comprising the latent factors of Timepoint 1 support and Timepoint 4 support, data screening identified one instance of missing data for Item 19 that was addressed using mean substitution. Item 3 posed a concern for negative skewness at timepoint 1 as well as timepoint 4. This concern was addressed through use of the MLR (Maximum Likelihood Estimation with Robust Standard Errors) estimator in Mplus analyses, which accounts for non-normal data (Muthén & Muthén, 1998-2007).

Individual items were averaged into parcels, with the intention being for parcels to contain highly correlated values and demonstrate adequate internal consistency at or above $\alpha = 0.60$. Additional assessment of the face validity of items informed parceling. Teacher support was represented by three parcels of two items each. CIS Sensitivity items 1and 14 comprised one parcel at each timepoint ($\alpha_{Timepoint1} = 0.86$; $\alpha_{Timepoint4} = 0.77$), items 6, 19 comprised the second parcel at each timepoint ($\alpha_{Timepoint1} = 0.50$; $\alpha_{Timepoint4} = 0.42$), and items 3 and 11 comprised the third parcel at each timepoint ($\alpha_{Timepoint1} = 0.60$; $\alpha_{Timepoint4} = 0.60$). Skewness and kurtosis values for each of the six parcels indicated normal distributions. Parcel statistics are presented in Table A6 (timepoint 1) and Table A7 (timepoint 4).

Timepoint 1 Teacher Structure and Timepoint 4 Teacher Structure

Item-level statistics and inter-item correlations for the original seven-item ASSIST Proactive Behavior Management subscale are presented in Table A8 (timepoint 1) and Table A9 (timepoint 4). Following deletion of item 6 due to significant missing data and limited variance, a confirmatory factor analysis assessed Timepoint 1 Teacher structure as a latent factor comprised of the six remaining items of the ASSIST Teacher Proactive Behavior Management subscale (Table A10). Results indicated a poor model fit ($\chi2(9, N = 52) = 39.64$; $p < .01$; RMSEA=.26; CFI=.68; TLI=.47). One item with a standardized factor loading under 0.4 (Item 7 – "Teacher positively acknowledges or praises students who are treating peers kindly") was removed from the model. While two additional items demonstrated communalities below 0.5 (Item 3 – "The teacher clearly explains learning objectives" and Item 5 – "the teacher made clear to students how they were expected to treat their peers"), these items demonstrated significant correlations with most other items on the scale and referenced relevant behaviors of preemptively setting classroom structures that communicate expectations for student behavior. Subsequently, these items were retained. The resulting five-item model showed improved but still poor fit ($\chi2(5, N = 52) = 18.92$; $p = .02$; RMSEA=.23; CFI=.82; TLI=.63). Modification indices suggested allowing correlated residuals for items 3, 4, and 5, suggesting similarity in the items. This step substantially improved model fit and demonstrated acceptable values across fit indices ($\chi2(2, N = 52) = .22$; $p = .89$; RMSEA < .01; CFI=1.00; TLI=1.12) (Table A11). Given that measurement invariance was assumed for timepoint 1 teacher structure and timepoint 4 teacher structure, the five-item model that demonstrated the best fit to timepoint 1 ASSIST Teacher Proactive Behavior

Management subscale data was reassessed for its fit to the same scale in Timepoint 4 (Table A12). Consistent with results for the five-item model using timepoint 1 data, the model run with Timepoint 4 data demonstrated acceptable values across fit indices ($\chi 2(2, N = 59) = .12$; $p = .94$; RMSEA < .01; CFI=1.00; TLI=1.15).

Following the confirmatory factor analysis and revision of the factor structure, crosstabs analyses identified remaining zero-frequency cells in the revised five-item ASSIST Teacher Proactive Behavior Management subscale at both timepoint 1 and timepoint 4. Subsequently, response categories for the subscale items across both timepoints were collapsed to reflect the following scale: 1 = never; 2 = at times; and 3 = a lot or almost always. With the revised and recoded item-level indicators comprising the latent factors of Timepoint 1 structure and Timepoint 4 structure, data screening identified four instances of missing data that were addressed using mean substitution. Skewness and kurtosis values were all within normal limits, falling between ± 2 (Curran, West, & Finch, 1996).

The methodology used create parcels of items to represent Teacher Support was applied here. Teacher structure was represented by one parcel of three items and one parcel of two items.

Teacher Proactive Behavior Management items 1 and 2 comprised one parcel at each timepoint ($\alpha_{Timepoint1} = 0.55$; $\alpha_{Timepoint4} = 0.78$), and items 3, 4, and 5 comprised the other parcel at each timepoint ($\alpha_{Timepoint1} = 0.62$; $\alpha_{Timepoint4} = 0.51$). Skewness and kurtosis values for each of the four parcels indicated normal distributions. Parcel statistics are presented in Table A13 (timepoint 1) and Table A14 (timepoint 4).

Timepoint 2 Student Engagement

A confirmatory factor analysis assessed the factor structure of Timepoint 2 Student Engagement using the six items of the ASSIST Student Compliance subscale (Table A15). The initial model was retained, as no items demonstrated standardized factor loadings below 0.4 or communalities below 0.5; in fact, all factor loadings and communalities were at or above 0.7. CFI and TLI global fit indices were acceptable ($\chi 2(9, N = 59) = 32.28$; $p <. 01$; RMSEA=.21; CFI=0.95; TLI=0.92). This scale demonstrated no missing data, zero-frequency cells requiring collapsing of response categories, or violations of normality. Given the viability of the initial factor structure for Timepoint 2 Student Compliance in the current sample, it was not further divided into parcels. In preserving the initial factor structure for the structural model, the latent factor of engagement could be calculated with the most integrity to the data as it was originally collected. In contrast, teacher support and teacher structure were meant to comprise a higher order latent factor at timepoint 1 and at timepoint 4 and were constrained by an assumption of measurement invariance across time. These additional parameters created model complexity that parceling helped to reduce.

Table A1.

Timepoint 1 Item-Level Statistics and Inter-Item Correlations for Original CIS Sensitivity Subscale

Item	Content	Item-Level Statistics		Inter-Item Correlations									
		M	SD	1	3	6	7	8	11	14	16	19	25
1	Speaks warmly to the children.	2.85	.98	1.00									
3	Listens attentively when children speak to him or her.	3.40	.87	.58	1.00								
6	Seems to enjoy the children.	2.77	1.06	.80	.42	1.00							
7	When the children misbehave, explains the reason for the rule they are breaking.	2.13	1.01	.49	.35	.45	1.00						
8	Encourages the children to try new experiences.	1.66	.94	.10	-.12	.31	-.02	1.00					
11	Seems enthusiastic about the children's activities and efforts.	2.60	1.00	.76	.40	.75	.38	.32	1.00				
14	Pays positive attention to the children as individuals.	2.90	.98	.72	.58	.59	.40	.01	.67	1.00			
16	Talks to children on a level they can understand.	3.44	.87	.52	.81	.30	.3	-.22	.36	.52	1.00		
19	Encourages children to exhibit prosocial behavior, e.g. sharing, helping.	1.84	1.05	.27	.07	.38	.51	.16	.42	.30	.15	1.00	
25	When talking to children, kneels, bends, or sits at their level to establish eye contact (as appropriate).	2.94	1.14	.38	.62	.28	.32	-.13	.39	.59	.55	.26	1.00

Note. Original subscale used the following response categories for all items on the subscale: 1 = not at all true; 2 = somewhat true; 3 = quite a bit true; 4 = very much true.

Table A2.

Timepoint 4 Item-Level Statistics and Inter-Item Correlations for Original CIS Sensitivity Subscale

Item	Content	Item-Level Statistics		Inter-Item Correlations									
		M	SD	1	3	6	7	8	11	14	16	19	25
1	Speaks warmly to the children.	2.59	.87	1.00									
3	Listens attentively when children speak to him or her.	3.07	.72	.60	1.00								
6	Seems to enjoy the children.	2.39	.79	.66	.41	1.00							
7	When the children misbehave, explains the reason for the rule they are breaking.	2.23	.91	.37	.32	.23	1.00						
8	Encourages the children to try new experiences.	1.09	.28	.07	-.03	.08	.13	1.00					
11	Seems enthusiastic about the children's activities and efforts.	2.14	.94	.60	.42	.70	.36	.22	1.00				
14	Pays positive attention to the children as individuals.	2.64	.89	.59	.56	.45	.21	.12	.58	1.00			
16	Talks to children on a level they can understand.	3.16	.64	.36	.43	.19	.17	-.07	.37	.49	1.00		
19	Encourages children to exhibit prosocial behavior, e.g. sharing, helping.	1.53	.75	.17	.13	.32	.25	-.05	.36	.18	-.06	1.00	
25	When talking to children, kneels, bends, or sits at their level to establish eye contact (as appropriate).	2.17	.93	.26	.32	.26	.24	.14	.41	.37	.36	-.06	1.00

Note. Original subscale used the following response categories for all items on the subscale: 1 = not at all true; 2 = somewhat true; 3 = quite a bit true; 4 = very much true.

Table A3.

Timepoint 1 CFA of Original CIS Sensitivity Subscale

Item	Content	Standardized Regression Weight	S.E.	*p* value
1	Speaks warmly to the children.	.92	---	---
3	Listens attentively when children speak to him or her.	.64	.62	<.01
6	Seems to enjoy the children.	.82	.12	<.01
7	*When the children misbehave, explains the reason for the rule they are breaking.* [a]	*.49*	*.52*	*<.01*
8	*Encourages the children to try new experiences.* [a]	*.1*	*.14*	*.36*
11	Seems enthusiastic about the children's activities and efforts.	.83	.92	<.01
14	Pays positive attention to the children as individuals.	.80	.87	<.01
16	*Talks to children on a level they can understand.*[a]	*.58*	*.53*	*<.01*
19	Encourages children to exhibit prosocial behavior, e.g. sharing, helping.	.27	.43	.01
25	*When talking to children, kneels, bends, or sits at their level to establish eye contact (as appropriate).* [a]	*.51*	*.66*	*<.01*

[a]Items removed from the model.

Table A4.

Timepoint 1 CFA of Revised CIS Sensitivity Subscale with Item-Level Statistics and Inter-Item Correlations

Confirmatory Factor Analysis					Item-Level Statistics		Inter-Item Correlations					
Item	Content	Standardized Regression Weight	S.E.	*p* value	*M*	*SD*	1	3	6	11	14	19
1	Speaks warmly to the children.	.93	---	---	2.54	.67	1.00					
3	Listens attentively when children speak to him or her.	.58	.12	<.01	2.81	.52	.58	1.00				
6	Seems to enjoy the children.	.85	.11	<.01	2.42	.70	.72	.44	1.00			
11	Seems enthusiastic about the children's activities and efforts.	.84	.11	<.01	2.37	.71	.73	.45	.75	1.00		
14	Pays positive attention to the children as individuals.	.77	.11	<.01	2.56	.64	.75	.56	.56	.66	1.00	
19	Encourages children to exhibit prosocial behavior, e.g. sharing, helping.	.35	.16	.01	1.73	.82	.16	.01	.34	.30	.14	1.00

Note. IFinal model used revised response categories for all items on the subscale: 1 = not at all true; 2 = somewhat true; and 3 = quite a bit or very much true.

Table A5.

Timepoint 4 CFA of Revised CIS Sensitivity Subscale with Item-Level Statistics and Inter-Item Correlations

Confirmatory Factor Analysis					Item-Level Statistics		Inter-Item Correlations					
Item	Content	Standardized Regression Weight	S.E.	*p* value	*M*	*SD*	1	3	6	11	14	19
1	Speaks warmly to the children.	.82	---	---	2.42	.65	1.00					
3	Listens attentively when children speak to him or her.	.63	.13	<.01	2.80	.45	.42	1.00				
6	Seems to enjoy the children.	.79	.14	<.01	2.31	.65	.63	.46	1.00			
11	Seems enthusiastic about the children's activities and efforts.	.80	.16	<.01	2.10	.83	.52	.46	.63	1.00		
14	Pays positive attention to the children as individuals.	.70	.16	<.01	2.48	.68	.6	.55	.49	.62	1.00	
19	Encourages children to exhibit prosocial behavior, e.g. sharing, helping.	.32	.15	.02	1.53	.75	.10	.22	.27	.36	.21	1.00

Note. Final model used revised response categories for all items on the subscale: 1 = not at all true; 2 = somewhat true; and 3 = quite a bit or very much true.

Table A6.

Timepoint 1 CIS Sensitivity Parcels

Parcel	CIS item number and Content	Standardized regression weight	S.E.	*p* value	α	Skewness	Kurtosis
1	Item 1, Item 14	.85	---	--	.86	-1.32	.92
2	Item 6, Item 19	.58	.16	<.01	.50	.09	-1.02
3	Item 3, Item 11	.94	.17	<.01	.60	-1.49	1.84

Table A7.

CIS Sensitivity Parcels (Timepoint 4)

Parcel	CIS item number and Content	Standardized regression weight	S.E.	*p* value	α	Skewness	Kurtosis
1	Item 1, Item 14	.73	---	---	.77	-.85	-.19
2	Item 6, Item 19	.66	.17	<.01	.42	.5	-.70
3	Item 3, Item 11	.95	.24	<.01	.60	-.67	-.55

Table A8.

Timepoint 1 Item-Level Statistics and Inter-Item Correlations for Original ASSIST Proactive Behavior Management Subscale

Item	Content	Item-Level Statistics		Inter-Item Correlations						
		M	SD	1	2	3	4	5	6*	7
1	Teacher gives clear instructions and directives to students.	4.21	.72	1.00						
2	Teacher is consistent, even-handed, and firm.	4.17	.76	.75	1.00					
3	The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements.	3.23	1.38	.26	.32	1.00				
4	Teacher praised students for specific behaviors or using social skills.	2.77	1.38	.40	.43	.51	1.00			
5	The teacher made clear to students how they were expected to treat their peers.	2.29	1.47	.23	.28	.45	.43	1.00		
6[a]	When aggression occurred, teacher acknowledged student's feelings (e.g., stating that the behavior is hurtful or offensive).	1.00	.000	---	---	---	---	---	1.00	
7	Teacher positively acknowledges or praises students who are treating peers kindly (e.g., including others, listening respectfully).	1.63	1.17	.14	.23	.37	.47	.57	---	1.00

Note. Original subscale used the following response categories for all items on the subscale: 1 = never; 2 = seldom; 3 = some of the time; 4 = a lot of the time; 5 = almost continuously.

[a]Correlations could not be computed because variable is constant.

Table A9.

Timepoint 4 Item-Level Statistics and Inter-Item Correlations for Original ASSIST Proactive Behavior Management Subscale

Item	Content	Item-Level Statistics		Inter-Item Correlations						
		M	SD	1	2	3	4	5	6	7
1	Teacher gives clear instructions and directives to students.	3.49	.88	1.00						
2	Teacher is consistent, even-handed, and firm.	3.75	.58	.70	1.00					
3	The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements.	2.09	1.16	.37	.28	1.00				
4	Teacher praised students for specific behaviors or using social skills.	2.00	1.07	.41	.28	.02	1.00			
5	The teacher made clear to students how they were expected to treat their peers.	1.78	.97	.31	.24	.30	.42	1.00		
6	When aggression occurred, teacher acknowledged student's feelings (e.g., stating that the behavior is hurtful or offensive).	1.56	.91	-.24	-.17	-.34	.19	.25	1.00	
7	Teacher positively acknowledges or praises students who are treating peers kindly (e.g., including others, listening respectfully).	1.46	.90	.24	.16	.22	.63	.76	.23	1.00

Note. Original subscale used the following response categories for all items on the subscale: 1 = never; 2 = seldom; 3 = some of the time; 4 = a lot of the time; 5 = almost continuously.

Table A10.

Timepoint 1 CFA of Original ASSIST Proactive Behavior Management Subscale sans Item 6

Item	Content	Standardized Regression Weight	S.E.	*p* value
1	Teacher gives clear instructions and directives to students.	.80	---	---
2	Teacher is consistent, even-handed, and firm.	.85	.21	<.01
3	The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements.	.46	.36	.002
4	Teacher praised students for specific behaviors or using social skills.	.59	.35	<.01
5	The Teacher made clear to students how they were expected to treat their peers.	.42	.38	.01
7[a]	*Teacher positively acknowledges or praises students who are treating peers kindly (e.g., including others, listening respectfully).*	*.36*	*.30*	*.02*

[a] Item removed from the model.

Table A11.

Timepoint 1 CFA of Revised ASSIST Proactive Behavior Management Subscale with Item-Level Statistics and Inter-Item Correlations

					Item-Level Statistics		Inter-Item Correlations				
Item	Content	Standardized Regression Weight	S.E.	*p* value	*M*	*SD*	1	2	3	4	5
1	Teacher gives clear instructions and directives to students.	.83	---	---	2.83	.38	1.00				
2	Teacher is consistent, even-handed, and firm.	.91	.27	<.01	2.79	.41	.39	1.00			
3	The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements.	.34	.34	.02	2.35	.76	.14	.24	1.00		
4	Teacher praised students for specific behaviors or using social skills.	.48	.33	<.01	2.04	.82	.34	.32	.39	1.00	
5	The Teacher made clear to students how they were expected to treat their peers.	.31	.36	.04	1.76	.81	.17	.19	.0	.36	1.00

Note. Final model used revised response categories for all items on the subscale: 1 = never; 2 = at times; and 3 = a lot or almost always.

Table A12.

Timepoint 4 CFA of Revised ASSIST Proactive Behavior Management Subscale with Item-Level Statistics and Inter-Item Correlations

					Item-Level Statistics		Inter-Item Correlations				
Item	Content	Standardized Regression Weight	S.E.	*p* value	*M*	*SD*	1	2	3	4	5
1	Teacher gives clear instructions and directives to students.	.97	---	---	2.61	.59	1.00				
2	Teacher is consistent, even-handed, and firm.	.72	.12	<.01	2.75	.44	.68	1.00			
3	The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements.	.35	.19	.02	1.70	.72	.37	.26	1.00		
4	Teacher praised students for specific behaviors or using social skills.	.42	.19	.01	1.71	.70	.40	.26	.11	1.00	
5	The Teacher made clear to students how they were expected to treat their peers.	.32	.16	.02	1.5	.62	.30	.32	.27	.41	1.00

Note. Final model used revised response categories for all items on the subscale: 1 = never; 2 = at times; and 3 = a lot or almost always.

Table A13.

Timepoint 1 ASSIST Proactive Behavior Management Parcels

Parcel	ASSIST Proactive Behavior Management item number and Content	Standardized regression weight[a]	S.E.	*p* value	α	Skewness	Kurtosis
1	Item 1, Item 2	---	---	---	.55	-1.50	1.00
2	Item 3, Item 4, Item 5	---	.05	<.01	.62	-.02	-.98

Note. Error variance was set to 1 in the CFA assessing the fit of parcels as indicators of the latent construct Timepoint 1 Structure in order to identify the model with two indicators.
[a]Standardized regression weights were not calculated due to holding variance equal across parcels.

Table A14.

ASSIST Proactive Behavior Management Parcels (Timepoint 4)

Parcel	ASSIST Proactive Behavior Management item number and Content	Standardized regression weight[a]	S.E.	*p* value	α	Skewness	Kurtosis
1	Item 1, Item 2	---	---	---	.78	-1.04	-.44
2	Item 3, Item 4, Item 5	---	.09	<.01	.51	.60	.06

*Note. E*rror variance was set to 1 in the CFA assessing the fit of parcels as indicators of the latent construct Timepoint 1 Structure in order to identify the model with two indicators.
[a]Standardized regression weights were not calculated due to holding variance equal across parcels.

Table A15.

Timepoint 2 CFA of ASSIST Student Compliance Subscale with Item-Level Statistics and Inter-Item Correlations

Confirmatory Factor Analysis					Item-Level Statistics		Inter-Item Correlations					
Item	Content	Standardized Regression Weight	S.E.	*p* value	*M*	*SD*	1	2	3	4	5	6
1	Students comply.	.97	---	---	3.76	.86	1.00					
2	Students consistently follow rules appropriate to settings.	.92	.06	<.01	3.70	.88	.91	1.00				
3	Students cooperate.	.96	.05	<.01	3.81	.82	.94	.88	1.00			
4	Students are interested, enthusiastic, and involved.	.85	.09	<.01	3.39	1.03	.81	.74	.82	1.00		
5	Students are focused and engaged.	.89	.08	<.01	3.42	1.00	.84	.80	.85	.87	1.00	
6	Students treat their peers with respect (i.e., listen when peers are talking).	.86	.08	<.01	3.59	.89	.82	.81	.81	.77	.85	1.00

Items maintained original response categories for all items on the subscale: 1=Never, 2=Seldom, 3=Some of the Time, 4=A lot of the Time, 5=Almost continuously.

Biography

Elizabeth McIntyre is a sixth-year student completing her Ph.D. in the Tulane University doctoral program in School Psychology, with a specialization in trauma. Her educational background includes a Master of Science in School Psychology from Tulane University and a Bachelor of Arts in Psychology from Villanova University, with minors in Spanish, Cognitive Science, and Peace & Justice Studies. Before beginning her graduate studies, Elizabeth worked as a case manager and direct service provider in children's behavioral health at Catholic Community Services of Juneau, Alaska, a research assistant in the ADHD Clinic at Duke University Medical Center, and a research assistant at the Duke University Center for Child and Family Policy. During her graduate training, Elizabeth provided research support in Dr. Stacy Overstreet's lab relating to developing and implementing professional development and consultation for school-wide initiatives in social-emotional learning and trauma-informed practices. Currently, Elizabeth is completing her predoctoral internship as a doctoral intern in psychology with Virginia Beach City Public Schools. Following graduation, she hopes to continue her work with children, families, and child-serving systems through providing trauma-informed services in assessment, intervention, and consultation.