MEASUREMENT EQUIVALENCE OF THE PENNSYLVANIA SCHOOL CLIMATE SURVEY ACROSS RACE AND GENDER: A FOCUS ON BLACK GIRLS

A Dissertation
Submitted to
The Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Brittney M. Williams
August 2020

Examining Committee Members:

Laura Pendergast, Ph.D., Advisory Chair, Psychological Studies in Education
Edward Fergus, Ph.D., Policy, Organizational and Leadership Studies
Lia Sandilos, Ph.D., Psychological Studies in Education
Joseph DuCette, Ph.D., Psychological Studies in Education

ABSTRACT

School climate is a construct frequently explored in educational research (Lee, Cornell, Gregory, & Fan, 2011; Wang & Degol, 2016) and is associated positive several outcomes including with improved student-teacher relationships (Croninger & Lee, 2001) and reduced school dropout risk (Jia, Konold, & Cornell, 2016). Unfortunately, emerging research indicates that racial differences in schools' penal practices may be negatively coupled with Black students' experience of a supportive school climate (Bottiani, Bradshaw, & Mendelson, 2017). Black students are disproportionately subjected to exclusionary disciplinary practices (e.g., suspensions and expulsions) in the learning environment (Fenning & Rose, 2007) and Black girls disproportionally encounter more adverse disciplinary outcomes in their educational setting (National Black Women's Justice Institute, 2018) relative to their White peers (Blake, Butler, Lewis, & Darensbourg, 2011; Crenshaw, Ocen, & Nanda, 2015; Epstein, Blake, & González, 2017; Hines-Datiri & Carter Andrews, 2017).

Persistent use of punitive practices may hinder Black girls' racial identity, academic, and social-emotional development in the classroom (Chavous, Rivas-Drake, Smalls, Griffin, & Cogburn, 2008; Leath, Mathews, Harrison, & Chavous, 2019) and consequently detrimentally impact their school climate. Researchers have called for evidence-based and culturally relevant interventions that promote positive academic outcomes for Black girls (Jones et al., 2018). However, evidence-based interventions are dependent on the use of empirically supported assessments with diverse student groups (Pendergast et al., 2017).

Evidence-based assessments are necessary to identify students' needs in the educational setting and provide baseline data that allow for evaluation of intervention effectiveness. Moreover, school climate assessments for Black girls may facilitate the development of evidence-based interventions for Black girls - who are disproportionally disciplined and may be at risk for experiencing a negative school climate. This study investigated the structural validity and reliability of scores from a school climate measurement tool. The study examined measurement invariance of the Student Connection Survey, with a specific focus on Black girls' scores. Results indicated the Student Connection Survey is represented by five-latent factors and is equivalent between Black and White middle school girls. Implications, strengths, limitations, and future research directions are discussed.

ACKNOWLEDGEMENTS

To my advisor and dissertation committee chair, Dr. Laura Pendergast for her guidance, challenging me, and providing a platform that allowed me to explore a topic I am passionate about. With her expertise and feedback, I have completed the most arduous task of my academic career.

I offer my sincerest thanks to my dissertation committee members, Dr. Lia Sandilos, Dr. Edward Fergus, and Dr. Joseph DuCette for their expertise, time, and commitment during this process. Thank you Dr. Sandilos for your assistance, and continued support throughout.

To my friends, mentors, and family who have been patient, encouraging, and understanding through this academic journey. Thank you. To my foundation, my mother, Denise, Taylor, and McKenna who have continued to encourage me and uplift me when times were difficult and reminding me of my faith. I love and appreciate each one of you. Lastly, to my partner, Brian who encourages me everyday. Thank you for your unwavering support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURE

Figure Page

CHAPTER 1

INTRODUCTION

Black students are disproportionately subjected to exclusionary disciplinary practices (e.g., suspensions and expulsions) in the learning environment (Fenning & Rose, 2007). Additionally, systemic bias and excessive disciplinary procedures put Black children at an elevated risk for academic failure and early contact with the criminal legal system (Morris & Perry, 2017). Strikingly, according to the National Black Women's Justice Institute (2018), Black girls are seven times more likely to receive at least one out of school suspension, and they are four times more likely to receive at least one in school suspension in comparison to their White female counterparts. Specifically, Black girls are the only group of girls that are disproportionally represented across multiple disciplinary outcomes (e.g., suspensions, expulsions, referral to law enforcement; National Black Women's Justice Institute, 2018).

Previous lines of work (e.g., Blake, Butler, Lewis, & Darensbourg, 2011; Crenshaw, Ocen, & Nanda, 2015; Epstein, Blake, & González, 2017; Hines-Datiri & Carter Andrews, 2017) indicate that Black girls inequitably receive school-based sanctions relative to their White peers. (The National Women's Law Center, 2017). Research suggests that both race and gender have a unique role in the educational experiences of Black girls in schools nationwide (E. W. Morris & Perry, 2017; M. W. Morris, 2012; Wun, 2016). For example, Epstein, Blake and González (2017) found that Black girls were perceived to need less protection and encouragement within the educational system, which may contribute to the unfair punishment by educators and school administrators.

Exclusionary discipline practices are considered the least effective method to address problem behaviors (American Academy of Pediatrics Committee on School Health, 2003; Lamont, 2013); yet, exclusionary disciplinary tactics are utilized across our nation's schools (U.S Department of Education, 2015; 2016) and disproportionately affect Black girls. A growing body of qualitative research examines how race and gender are related to the criminalization of Black girls in schools (Hines-Datiri & Carter Andrews, 2017; Wun, 2016); still "very little research highlights the short- and long-term effects of over discipline and push out policies on girls of color" (Crenshaw et al., 2015, p.8). Studies have shown that persistent use of punitive practices can negatively impact the extent to which Black girls develop positive self-concepts around academics, as well as their racial identity, gender identity, and social emotional well-being in the classroom (Chavous, Rivas-Drake, Smalls, Griffin, & Cogburn, 2008; Leath, Mathews, Harrison, & Chavous, 2019). Therefore, empirical investigation of how Black girls, who are overdisciplined, perceive their learning environment and the possible influence of negative outcomes (Epstein et al., 2017) is required.

In the scientific literature, the term *school climate* is used to represent many different aspects of the school environment (Lee, Cornell, Gregory, & Fan, 2011; Wang & Degol, 2016), and emerging research suggests that racial differences in schools' penal practices may be negatively coupled with Black students' experience of a supportive school climate (Bottiani, Bradshaw, & Mendelson, 2017). Notably, few studies examine how punitive practices influence Black girls' school experience.

The National School Climate Center defines school climate as "the quality and character of school life based on patterns of students', parents' and school staffs'

experience of school life; it also reflects standards, goals, values, interpersonal relationships, teaching and learning practices and organizational structures" (NSCC, 2007; para 1). A positive school environment has been linked to promising outcomes such as improved student-teacher relationships (Croninger & Lee, 2001), higher academic achievement (Wang et al., 2014) and reduced school dropout risk (Jia, Konold, & Cornell, 2016). Given these favorable outcomes, the U.S. Department of Education invested approximately $70 million to guide state and local efforts to improve school climate and adhere to a data-driven strategy (U.S. Department of Education, 2014). This included garnering feedback from families, students, educators, and school personnel to prevent, identify, reduce, and eliminate discriminatory discipline exercises.

Following a data-driven approach, the American Institute for Research (AIR) partnered with the Chicago Public Schools to establish the Student Connection Survey. The Student Connection Survey was created to identify features of school climate that could be used to inform school-based interventions (Garibaldi, Ruddy, Kendziora, & Osher, 2015). Although efforts to disseminate the Student Connection Survey across school districts and state departments nationwide are ongoing, there is limited evidence available that address the psychometric properties (e.g., validity, reliability, and measurement equivalence) of the survey.

According to *the Standards for Educational and Psychological Testing* (2014) one of the foundational considerations in developing and evaluating measurement tools is evidence of validity. Validity refers to "to the degree to which evidence and theory support the interpretation of tests scores for proposed uses of tests" (Joint Committee on Standards for Educational and Psychological Testing, 2014, p.11). A possible threat to

validity is an issue of measurement "bias", which in statistics refers to systematic error in the estimation of a value (Reynolds & Ramsay, 2003). That bias is said to be present if a measure overestimates or underestimates scores on a variable for individuals who differ across race, ethnicity, or gender characteristics (Reynolds & Ramsay, 2003). Researchers (Reynolds & Ramsay, 2003; Zumbo, 2007) have identified several forms of bias that can emerge. For example, systematic error in the estimation of a value may be present at the structural level of a measure (i.e., construct bias), suggesting that the definition of the construct of interest (e.g., school climate) is not the same across groups. Bias may also be present at the item level (e.g., now referred to as differential item functioning DIF; Zumbo, 2007) which suggests some features of the test item are not relevant to the latent construct (Zumbo, 2007).

Historically, issues of measurement bias have centered around disparate mean scores by groups based on race or gender (Zumbo, 2007); however, examination exclusively of mean differences is no longer considered to be sufficient (Zumbo, 2007; Zumbo & Koh, 2005). Instead, researchers should examine the way that bias might impact multiple forms of validity (Zumbo, 2007). Statistical approaches to address measurement bias include factor analysis techniques. An extension of factor analysis that is used to explore evidence of validity across groups is multi-group confirmatory factor analysis (Pendergast, von der Embse, Kilgus, & Eklund, 2017; Zumbo & Koh, 2005). Multi-group confirmatory factor analysis (MG-CFA) allows researchers to investigate to what extent the measure is equivalent/invariant across race and gender characteristics (Zumbo & Koh, 2005).

Measurement invariance is a tool that is utilized to explore whether the relationship between responses to items and the construct are operating in the same way across cultural and gendered groups at different levels (Chen, 2007; Pendergast et al., 2017). Measurement invariance procedures are necessary because "when groups are compared based on tools that do not measure the same underlying construct, inference problems occur" (Chen, 2007; p.464). Conclusions that are drawn from tools that have not been shown to demonstrate evidence of validity for all groups may not generalize to other instruments measuring the same construct (van de Vijver & Tanzer, 2004). Hence, it is necessary to investigate the measurement equivalence of a school climate survey before it can be used to make comparisons across groups and evidence-based interventions.

Consequences of Non-Equivalence/Invariance

Measurement non-invariance can impact the conclusions that are drawn and the recommendations including evidenced-based interventions to address school climate. The result of not capturing features of school climate across racially and ethnically diverse students may result in not understanding how diverse groups perceive and experience their school environment. Understanding how Black girls view their school climate and the influence of exclusionary practices may address future efforts for targeted evidence-based interventions. Evidence-based interventions are dependent upon evidence-based measures (Pendergast et al., 2017), and given that evidence of validity and measurement equivalence of the Student Connection Survey is scant, it is unknown to what extent scores are equivalent across race and gender; specifically among Black girls. Some researchers (e.g., Jones, Lee, Matlack, & Zigarelli, 2018) have called for evidence-based,

culturally relevant interventions that promote positive outcomes to address student needs with specific racial and gendered backgrounds. Pendergast and colleagues (2017) noted that evidence-based interventions are contingent on the use of scientifically based assessments with diverse student groups. Consequently, capturing Black girls' perceptions of school climate using a measurement tool that demonstrates strong psychometric qualities enhances development of appropriate school-wide interventions that address their school climate.

Research demonstrates variability in school climate among Black students in comparison to their White counterparts (Bottiani, Bradshaw, & Mendelson, 2016; Bottiani et al., 2017; Koth, Bradshaw, & Leaf, 2008) and that exclusionary discipline policies impact school climate and lead to deleterious outcomes such as school dropout, lower academic achievement, and involvement with the criminal justice system (Blake et al., 2011; Crenshaw et al., 2015; Lee et al., 2011). Given that Black girls are disciplined at alarming rates (Blake et al., 2011; Crenshaw et al., 2015; Epstein et al., 2017; Hines-Datiri & Carter Andrews, 2017), it is likely that they experience a less positive school climate, which places them at greater risk for school push out, criminalization, and lower academic achievement. Thus, targeted empirically based interventions to protect Black girls from such negative outcomes are required; however, such evidence-based interventions are dependent upon instruments that are demonstrated to produce valid scores for Black girls, and it is unclear whether currently available instruments achieve that aim. Moreover, few school climate assessment tools have been used to study how Black girls view their school environment. Investigation of how Black girls observe their school culture using a tool that demonstrates strong psychometric properties of school

climate is an imperative next step to the existing school climate and Black girl scholarship.

The influence of school discipline and Black girls' perceptions of school climate have thus far not been thoroughly investigated and are a major focus of the current research. Accordingly, the current study serves several purposes. First, to examine the extent to which the Student Connection Survey is equivalent across race and gender, with a focus on Black girls' scores. Research that is available (Garibaldi et al., 2015; Osher, Kendziora, & Chinen, 2008) is incomplete and lacks investigation of its potential bias or invariance. Broadly, this study aims to examine the reliability and validity of the Student Connection Survey, which was recently used by the Pennsylvania Department of Education. The current study adds to the available evidence by examining school climate measurement in terms of race and gendered perceptions. More importantly, examination of Black girls' scores across school climate domains may help to demonstrate sets of items on the Student Connection Survey that are more salient for this group. Investigating Black girls' scores may help to identify particular areas in need of targeted school climate intervention.

In summary, research has shown that a positive school climate is a protective factor against out of school suspensions, school dropout and lower academic achievement. Therefore, an improved understanding of the factors of school climate is indispensable. Establishing a strong theoretical basis of school climate using measurement and understanding race and gendered perceptions may bridge this gap in the literature. Moreover, analysis that centers on Black girls who are disproportionately disciplined is paramount.

Existing literature that explores exclusionary practices has largely focused on Black males (e.g., Fergus, Noguera & Martin, 2014; Gregory, Skiba, & Noguera, 2010). The study of Black male discipline is not to be used as a proxy for exmaining the study of Black children. Moreover, the focus on Black males may obscure the way in which Black females and males experience the phenomenon together and differently (Morris, 2016). Although it is necessary and important to understand the ways in which Black boys are disproportionately disciplined, the experiences of Black girls are unique, are vastly understudied and the focus of the current study.

CHAPTER 2

LITERATURE REVIEW

Overview of School Climate

The National School Climate Center defines school climate as "the quality and character of school life based on patterns of students', parents' and school staffs' experience of school life; it also reflects standards, goals, values, interpersonal relationships, teaching and learning practices and organizational structures" (NSCC, 2018; para 1). A positive school climate is associated with decreased student absenteeism in middle and high school (Hendron & Kearney, 2016), better academic attainment (Brand, Felner, Shim, Seitsinger, & Dumas, 2003; Wang et al., 2014), lower rates of suspension (Lee et al., 2011), and positive teacher-student relationships (Slaughter-Defoe & Carlson, 1996). In 2014, the U.S. Department of Education directed monies to nearly 90 school districts that focused on building social, emotional, and behavioral supports for their students. Through the respective initiatives, schools were tasked with incorporating a data-driven approach into inform school-wide evidenced-based interventions designed to create a safer, "more nurturing" (White House, 2013; p.10) school climate and to adhere to best practices around issues such as school discipline.

Although the term school climate can be traced back over a century in the educational literature (Perry, 1908), a systematic investigation of the construct began when Halpin and Croft (1963) developed the 64-item instrument to measure organizational climate of schools. The Organizational Climate Descriptive Questionnaire (Halpin & Croft,1963) included specific dimensions along a continuum of how open, closed, energetic, and lively the organization (e.g., school) was moving toward a goal

(Thomas, 1976). The measure was the catalyst for later climate research and instrument development to investigate the relationship between school climate, dimensions of school climate, staff perceptions, and student outcomes (Brookover et al., 1978; Zullig, Koopman, Patton, & Ubbes, 2010).

For example, Brookover and colleagues (1978) evaluated school climate, defined as the norms and expectations held by individuals in a sample of elementary schools in the midwestern United States. Findings revealed that school climate was positively associated with student achievement when accounting for race, socioeconomic status and racial composition of the school. Further, students' sense of academic futility (e.g., feelings of not having control over successes, failure in the school social system, or lack of teacher support) accounted for the most variance in student academic achievement. Academic futility was found to be most predictive of achievement in schools with a majority Black student body compared to schools with a majority White student body with large effects (R^2=.481). Overall, this study underscored that school climate may impact student outcomes (i.e., achievement) and may be most salient in a majority Black student body (Brookover et al., 1978).

Currently, school climate has garnered increased national attention in psychological and educational scholarship (Wang & Degol, 2016), and research demonstrates an association between school climate and various outcomes including academic achievement (Brand et al., 2003;Wang et al., 2014), suspension rates (Lee et al., 2011) and student-teacher relationships (Slaughter-Defoe & Carlson, 1996). School climate has been defined as the quality and character of school life (National School Climate Center, 2007); however, there is no universal definition among researchers

(Wang & Degol, 2016) of the construct as it is being used to encapsulate different aspects of school life. An understanding of how scholars have conceptualized school climate is needed.

Conceptualizing School Climate

To date, there is no agreed upon definition of school climate (Wang & Degol, 2016), which encompasses several characteristics of school life (Cohen, McCabe, Michelli, & Pickeral, 2009; Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013). Cohen and colleagues (2009) argue that school climate "refers to the quality and character of school life based on patterns of people's experience of school life and reflects norms, goals, values, interpersonal relationship, teaching and learning practices, and organizational structures" (Cohen et al., 2009; p.182). Using this definition, Cohen et al. (2009) identified four noteworthy aspects of school climate that are routinely studied in the literature: safety, teaching and learning, relationships, and structural organizational considerations such as class size, school size, and student-teacher ratio.

Conversely, Haynes, Emmons, and Ben-Avie (2010) explain school climate as "the quality and consistency of interpersonal interactions within the school community that influence children's cognitive, social, and psychological development" (Haynes, Emmons, & Ben-Avie, 2010; p. 322). Together, the definitions illustrate the lack of consensus (Wang & Degol, 2016) and how adopted definitions represent several aspects of the school environment (Thapa et al., 2013). Researchers have used several theories to support their definitions of school climate (Wang & Degol, 2016). The frequently used theories of school climate research are discussed hereafter.

Bioecological Systems Theory

The study of school climate has been commonly guided by bioecological systems theory (Bronfenbrenner, 1979) and social cognitive theory (Bandura, 1986). The bio-ecological (Bronfenbrenner, 1979) theoretical framework has been considered "one of the pillars" of school climate research (Wang & Degol, 2016; p. 319) that asserts that child development occurs gradually through reciprocal interactions between a child and their proximate environment. The bioecological systems framework encompasses four nested systems. The first level is the microsystem which encompasses an individual's interpersonal relationships and immediate surroundings (Kohl, Recchia, & Steffgen, 2013). The mesosystem is the second level that involves various interactions between those in the microsystem and settings (family, school, and peers) in the mesosystem. The mesosystem incorporates associations between the mesosystem and microsystem (Back, Polk, Keys, & McMahon, 2016), for example, the relationship between one's family and school administrators. The third level is the ecosystem, and represents the setting containing the mesosystem (e.g., quality of family relationships, school) which may not impact an individual directly, but indirectly. The fourth level is the macrosystem which describes the culture, economics, and politics of a place (Kohl et al., 2013) and in an educational setting, "might refer to a school's climate" (Back et al., 2016; p. 398).

Social Cognitive Theory

Social cognitive theory (Bandura, 1986) posits that environmental factors influence how people think of themselves and their environments. Therefore, this framework seeks to explain how students' perceptions likely contribute to their

behavioral and academic experiences. With a social cognitive theory perspective, school climate impacts student development through the quality of supportive teacher-student relationships, scholastic expectations, and maintaining an environment where students feel emotionally and physically safe (Wang & Degol, 2016). Thus, when considering the school climate using a social-cognitive perspective, child development is likely influenced by the interpersonal relationships between students and teachers (Hughes, Luo, Kwok, & Loyd, 2008), the school curriculum (Ford, Harris, Tyson, & Trotman, 2002) and school discipline practices (American Academy of Pediatrics Committee on School Health, 2003).

Taken together, the bioecological systems theory (Bronfenbrenner, 1979) and social cognitive theory (Bandura, 1986) emphasize the significance of the environment, quality of relationships and student development and have been frequently used frameworks in the school climate literature. Although these theories have largely influenced school climate scholarship an understanding of the unique ecological conditions children of color may encounter is absent from traditional paradigms. Thus, attention to experiences children of color face is needed (Coll et al., 1996).

Conditions for Learning

Another school climate paradigm and one that serves as the framework for the current study is the Conditions for Learning model (Garibaldi et al., 2015). In 2005, the American Institute for Research (AIR) partnered with the Chicago Public Schools to identify significant measurable factors that schools should address to improve student attendance, achievement, and graduation rates (Garibaldi et al., 2015). As a result of this partnership, aspects described as social emotional learning, student support, high

expectation/rigor and safe and respectful school climate were considered the most proximal to the process of teaching and learning and overall school climate. The Conditions for Learning model is hypothesized to account for school and individual characteristics that are vital for understanding school experiences that contribute to school climate (Garibaldi et al., 2015). The conditions identified are provided in detail hereinafter.

Social emotional learning

Social emotional learning (SEL) is the ability to develop effective strategies to help recognize and manage emotions that can support individuals to establish the capability to handle challenging situations (Collaborative for Social and Emotional Learning, 2015). Students who demonstrate adaptive social and emotional skills are better equipped to manage stress and difficult situations which are associated with learning and school performance (Osher et al., 2008). Social emotional learning has been associated with positive outcomes (Domitrovich, Durlak, Staley, & Weissberg, 2017; Durlak et al., 2011; Payton et al., 2008). For example, social emotional learning programs that integrated activities that foster personal and social skill development had positive effects on improving academic test scores (Payton et al., 2008).

Similarly, authors (Durlak et al., 2011) found students' behavior and academic skills improved when they were in a social emotional learning program compared to students who did not receive the social emotional learning program. In addition, Domitrovich et al. (2017) found that social and emotional skills could be reinforced through positive teacher-student relationships and were associated with lower rates of

depression and behavioral problems that may therefore influence a child's school climate (Domitrovich et al., 2017)

Student support

Student support is defined as students feel cared about and believe that the adults in students' lives work collaboratively to encourage and support them (Garibaldi et al., 2015; Osher et al., 2008). Researchers have examined how students feel supported and connected in their school setting (Croninger & Lee, 2001; Gregory et al., 2010; Pena-Shaff, Bessette-Symons, Tate, & Fingehut, 2019; Wu, Pink, Crain, & Moles, 1982). Specifically, research has shown that teacher support reduced the likelihood of a student dropping out of secondary school (Croninger & Lee, 2001) and was a protective factor against risk of suspension (Wu et al., 1982).

Similarly, Gregory et al. (2010) examined 7,300 ninth-grade students' beliefs of school structure and support (i.e., two corresponding aspects of school climate) and found that school climate was positive for students if they believed their teachers were supportive of them. Gregory and colleagues' (2010) findings demonstrate that students who attended schools without supportive relationships were more likely to experience punitive disciplinary actions, victimization, and reduced academic achievement (Gregory et al., 2010).

Thus, it is not surprising that students who feel supported by their teachers are at a lower risk of being disciplined. Particularly, students who view their teachers as having high expectations of them demonstrated increased motivation (Goodenow, 1993), classroom participation (Voelkl, 1995) and may be more likely to respond to teacher encouragement when they feel their teachers are concerned for their well-being (Jia et al.,

2016). More importantly, research indicates that positive student-teacher relationships are an aspect of school climate that can influence academic and disciplinary outcomes (Gregory & Ripski, 2008).Yet, researchers (Ferguson, 2003; McKown & Weinstein, 2008; Neal, McCray, Webb-Johnson, & Bridgest, 2003) suggest that when teacher perceptions of students are examined, race and ethnicity generate differential outcomes, particularly for Black students. For example, Neal, McCray, Webb-Johnson and Bridgest (2003) found that Black male students who exhibited a culturally specific movement style were perceived by their teachers to be more aggressive, have poorer academic achievement, and need special education services. Results demonstrate that teachers' expectations, biases and misunderstanding of cultural factors impact student-teacher relationships that eventually influence their students' school experience (Allen, Scott, & Lewis, 2013; Tenenbaum & Ruck, 2007).

High expectation/rigor

Osher et al. (2008) stressed that high expectations for student achievement influenced their school climate, and students are prone to perform better academically and behaviorally when they feel that teachers and other educators hold high expectations of them (Catalano, Haggerty, Oesterle, Fleming, & Hawkins, 2004). Studies demonstrate that students at schools with higher average achievement are significantly less likely to experience expulsions (Jia et al., 2009; Skiba et al., 2014). In turn, the available evidence shows that schools that build encouragement, community, and set high academic expectations can promote a climate that is safe and of mutual respect (Osher et al., 2008).

Safety and respectful school climate

Research has demonstrated that students feel physically and emotionally safe when they experience a safe and respectful school climate (Osher, 2008; Thapa et al., 2013). Under the Conditions for Learning framework, perceptions of safety underscore the physical and emotional security of the school setting and school-related activities (Garibaldi et al., 2015). Studies have shown that the lack of a safe or a respectful environment negatively influences student outcomes (Crenshaw, Ocen, Nanda, 2015). Specifically, research demonstrates that school practices (e.g., discipline) can influence students' feelings of safety (Ruck & Wortley, 2002). Notably, Voight, Hanson, O'Malley, and Adekanye (2015) found that when compared to White students, Black and Hispanic students had less favorable experiences of safety and connectedness with their educators. Voight et al. (2015) suggest that students can experience the same setting differently based on their racial group membership.

In summary school climate has been influenced by several theoretical frameworks (e.g., Bronfenbrenner, Social Cognitive Theory, Conditions for Learning). Each framework previously discussed proffers a conceptualization for how school climate may affect student, teachers, and school personnel (Back et al., 2016; Hughes et al., 2008). The Conditions for Learning framework (Garibaldi et al., 2015) which serves as the rationale for the Student Connection Survey and is represented by social emotional learning, student support, high expectation/rigor, safety and respectful school climate domains. Previous studies (Croninger & Lee, 2001; Gregory et al., 2010; Pena-Shaff, Bessette-Symons, Tate, & Fingehut, 2019; Wu, Pink, Crain, & Moles, 1982) highlight the relationship between aspects of school climate and how students' racial and ethnic

identity may influence their school experiences which suggests that there is an association between racial and ethnic differences and features of school climate.

Racial Differences in School Climate

Studies (Bottiani et al., 2017; Bryson & Childs, 2018; Fan, Williams, & Marie Corkin, 2011; Gregory, Cornell, & Fan, 2011; Haynes, Emmons, & Ben-Avie, 2010; Koth, Bradshaw, & Leaf, 2008; Mattison & Aber, 2007; Pena-Shaff et al., 2019; Wu et al., 1982) have shown that positive school climate differs among racially and ethnically diverse students across elementary and secondary school settings. For example, among a sample of over 2,000 5th graders, Koth et al. (2008) found that student race and gender were significant predictors of a negative school climate whereby Black male students exhibited more negative beliefs relative to their White classmates. The Koth et al. (2008) study was limited, however, as only two subscales (i.e.,, order and discipline, and academic motivation) from the School Development Survey (Haynes, Emmons, & Ben-Avie, 2001) were explored. The authors (Koth et al., 2018) recommended incorporating additional factors such as social relationships and academic abilities of students' school climate.

Conversely, a recent study suggests that Black high school students reported positive views of student-teacher relationships, fairness of rules, and student connectedness when compared to White and Hispanic students (Bryson & Childs, 2018). Bryson and Childs (2018) claimed that one of the explanations for why they found Black students as having more positive views of school climate was attributed to differences in school location (e.g., urban setting) and a higher percentage of racially and ethnically diverse students (50-55%) in the school (Bryson & Childs, 2018). Although this may be

one possible explanation for their findings, it should be noted that statistical and methodological limitations are observed throughout. For instance, the Bryon and Childs (2018) measure of school connectedness (e.g., feature of school climate) demonstrated relatively poor internal consistency (Cronbach alpha = .60) and best practice approaches to demonstrate evidence of internal/structural validity (e.g., factor analysis, factor loadings, promax rotation; Fabrigar, Wegener, MacCallum, & Strahan, 1999) were not utilized.

Conversely, in a separate sample of high school students, Bottiani, Bradshaw and Mendelson (2016) found student perceptions across features of school climate such as caring, equity, and high expectations were more negative for Black students relative to White students in more diverse, lower income schools. Results are also consistent with extant literature that has shown that race remains a significant predictor of negative outcomes (e.g., suspensions) even after controlling for SES (Skiba et al., 2014; Wallace, Goodkind, Wallace, & Bachman, 2008; Wu et al., 1982; Wun, 2016). Bottiani et al.'s (2016) results were further strengthened by empirical support of valid and reliable measures and best factor analytic practices. Findings demonstrate that consideration of student race is critical to improving our understanding of the processes that contribute to racial disparities in school outcomes and call for additional measurement to capture students' sense of school climate.

Bottiani et al. (2016) suggest that approaches to help reduce racial disparities in school climate and suspensions may include improving teacher-student relationships through school-level interventions for racially and ethnically diverse students. Additional research examining other aspects of school climate and analyses of Black students'

perceived academic environment may also broaden our understanding of racial differences of school climate.

Research indicates that frequent use of suspensions gives students the impression they are targets of discrimination (Haynes et al., 2010), and schools that recurrently rely on suspension practices maintain a learning environment that is viewed as punitive and not accepting of their students (Christle, Jolivette, & Nelson, 2007; Lee et al., 2011). Previous findings indicate that negative student perceptions of school climate are associated with more suspensions (Gregory et al., 2011; Lee et al., 2011) and disciplinary practices, such as suspensions, disproportionately affect Black male and female students (American Psychological Association Zero Tolerance Task Force, 2008; Blake et al., 2011; Carter, Skiba, Arredondo, & Pollock, 2016; Crenshaw et al., 2015; Gregory & Mosely, 2004; Morris & Perry, 2017; Skiba, 2000).Therefore, an understanding of Black students' perceptions of school climate (e.g., a group that is disproportionally suspended) is warranted.

Wu, Pink, Crain and Moles (1982) found that Black students' attitudes were impacted by school level characteristics. Particularly, school level features such as the prevalence of suspensions of other students in the schools, the inconsistency of when rules were enforced, and the belief that teachers were disinterested in them were significantly related to a student's likelihood of being suspended. Wu et al. (1982) concluded "although student suspensions are a matter of student misbehavior, it is more of a matter of how the school treats its students" (p. 370). Wu and colleagues (1982) found that schools that function in specific ways (e.g., enforced rules inconsistently,

prevalence of suspensions, teachers were disinterested in their students) were more likely to have students who held negative perceptions of school climate.

Similarly, Fan, Williams, and Corkin (2011) revealed significant variability in student beliefs of school climate across many individual and school level variables. Several aspects of school climate (i.e., discipline, order, safety, fairness and clarity of school rules and student-teacher relationships) were used. Specifically, Fan and colleagues (2011) found that male students perceived school rules to be less fair, (regression coefficient = 2.78, $p<.001$), and teacher-student relationships (regression coefficient = 2.93, $p<.001$) to be less supportive in relation to female students. Their findings provide empirical support for gender and racial differences across features of school climate. The Fan et al. (2011) results support earlier work (Haynes et al., 2010; Koth, Bradshaw, & Leaf, 2008; Mattison & Aber, 2007) indicating that different student groups perceive aspects of school climate in dissimilar ways and underscore the importance of investigating varied features of school climate among racially diverse groups separately.

In 2017, Bottiani, Bradshaw and Mendelson expanded school climate scholarship further by investigating discipline as a contextual factor associated with students' perceived school connectedness. The focus on early adolescence is consistent with research indicating that school suspensions peak during the middle school and early high school years (Raffaele Mendez et al., 2002; Theriot, Craun, & Dupper, 2010). Bottiani et al. (2017) conducted a confirmatory factor analysis and found a three-factor model of school climate (e.g., equity, school belonging, and adjustment problems) and examined measurement equivalence demonstrated by a series of models (Meredith, 1993) between

Black and White high school students. Measurement invariance was found by race at three levels (e.g., configural, metric, scalar). For example, configural invariance refers to whether the factor structure is the same across groups. Metric invariance is referring to the consistency of the magnitudes of relationships between items and their factors across groups. Scalar invariance refers to equal item intercepts across groups and is tested when metric invariance is supported (Dimitrov, 2010). Overall, findings of the study further support that racial differences in a schools' use of out of school suspensions may be negatively associated with Black adolescent students' experience of a positive school climate (Bottiani et al., 2016, 2017).

Despite confirmatory factor analysis findings, an explanation of the role of gender differences between racial groups was not addressed. A recent report from the U.S. Department of Education indicates that Black girls face a higher risk of out of school suspensions that begins as early as preschool and continues throughout Black females' educational experiences (U.S. Department of Education, 2016), and suggests that further investigation is needed to understand suspensions as a contextual factor associated with Black girls' educational experiences.

Taken together, studies have shown racial differences in school climate where Black students hold more negative views of their school environment (Bottiani et al., 2016, 2017; Fan et al., 2011; Koth et al., 2008; Pena-Shaff et al., 2019). Several characteristics of school climate (e.g., teacher-student relationships, safety, fairness) have been explored (Bryson & Childs, 2018; Fan et al., 2011) that contribute to Black students' observed negative school experience. Literature suggests that Black students perceive a negative school climate relative to their White classmates, and a negative

school experience may influence the risk for out of school suspension and school dropout. Investigating Black students' perceptions of their school climate across additional domains is necessary.

Thus, further measurement approaches (e.g., measurement equivalence/invariance techniques) to examine Black students' school climate are necessary. Equally important, adequately measuring Black students' perceived school experience may identify indicators on a measure of school climate that can inform evidence-based interventions. Researchers (e.g., Koth et al., 2008; Skiba, Arredondo, & Williams, 2014) recommend interventions to address school climate may be effective if they target the most vulnerable students (i.e., students with more negative perceptions of the school environment). Further, investigating Black student scores may reveal how different student groups perceive dimensions of school climate in various ways (Fan et al., 2011). Lastly, an investigation of Black girls' perceived school climate adds to the literature base as a growing body of work indicates Black girls are more likely to be disciplined relative to their White and Black male classmates (Blake et al., 2011; Morris & Perry, 2017) that may place them at greater risk for a negative school climate.

Black Girls, Discipline & School Climate

There is a dearth of research regarding the educational experiences for Black girls and illustrates the need to advance the scholarship focused on their experiences across a range of educational contexts (Patton, Crenshaw, & Haynes, 2016). Emerging evidence (Blake et al., 2011; Crenshaw et al., 2015; Morris & Perry, 2017) has shown that Black females face unfair practices similar to their Black male peers (Skiba, Michael, Nardo, & Peterson, 2002) and disproportionate discipline begins in elementary and extends through

high school (Blake et al., 2011). Blake et al. (2011) found that Black girls were two times more likely to receive in school and out of school suspensions than all female students in an urban Midwestern school district. Findings suggest that Black girls were penalized for more subjective (e.g., defiance) behaviors that were in opposition of conventional standards of femininity (Blake et al., 2011;Collins, 2004). In many cases, implicit biases and stereotyping have been shown to contribute to teachers' perceptions of Black girls (Blake et al., 2011; Crenshaw et al., 2015). Particularly, Black girls may be subjected to harsher punishment by their educators because they are perceived to be loud and uncontrollable (Crenshaw et al., 2015).

Crenshaw et al. (2015) noted that when Black girls face harsher punishments, they may feel unsafe or unsupported within their learning environment which can lead to underachievement, school dropout and involvement with the criminal justice system (Crenshaw et al., 2015). Limited studies have examined how Black girls view their school climate, but preliminary studies (Slaughter-Defoe & Carlson, 1996; Watkins & Aber, 2009) have indicated that Black girls perceive school climate in more negative terms than their White and male counterparts. Empirical research should build upon the school climate scholarship to examine how Black girls identify their learning environment across school climate domains. Further investigation of Black females' perceptions of school climate is needed at this time.

As described, earlier studies (Gregory & Ripski, 2008; Slaughter-Defoe & Carlson, 1996) indicate that the discrepancy in a positive and negative school climate is most dominant for Black students and features of school climate (e.g., teacher-student relationships, fairness, and safety) may contribute to Black students' witnessed negative

school experience (Koth et al., 2008; Mattison & Aber, 2007). Researchers (Bottiani et al., 2017; Fan et al., 2011; Koth et al., 2008; Pena-Shaff et al., 2019) have employed various methodologies to examine the relationship between individual student level predictors (e.g., race, gender) and student outcomes (e.g., discipline, out of school suspension, or student-teacher relationships) that impact students' overall school climate.

Notably, researchers examined several features of school climate to explore whether school discipline influenced one's school environment (Bottiani et al., 2016, 2017; Fan et al., 2011; Koth et al., 2008); however, additional research is needed to investigate racial and gender differences in school climate. Current school climate studies have been hindered by a paucity of measures that have been shown to produce reliable and valid scores among Black youth (e.g., Bryson & Childs, 2018) and underscore the need to examine currently available tools among Black students. Racial differences in a positive school climate may be explained by disparities in school disciplinary practices and empirical support is needed to explore the extent to which punitive policies influence Black girls' perceived learning environment. Therefore, examination of how individuals view school life using reliable and valid measures is necessary. Nearly 15% of the school climate studies have primarily focused on validating and developing measures of school climate (Bear, Gaskins, Blank, & Chen, 2011; Brand et al., 2003; Wang & Degol, 2016), and a common method used to capture student views of school climate are self-report surveys. A review of some recently employed surveys is discussed next.

Measurement of School Climate

Approximately 92% of studies assess school climate through self-report surveys (Wang & Degol, 2016). Likert-type surveys ask students, parents, and teachers to reflect

on their experiences with the particular school. The information gathered from survey data is useful for reviewing aspects of school climate at the individual, classroom and school level (Koth et al., 2008). Collecting information from different groups may provide unique perspectives of school climate (Wang & Degol, 2016), and investigating diverse perspectives on school climate across informants aids researchers in determining to what extent reporters converge or diverge (Wang & Degol, 2016).

Historically, teacher and staff perceptions have been the focus of research, but recently, researchers have shown an interest in student perceptions (Brand et al., 2003; Koth et al., 2008; Van Horn, 2003). Haynes et al. (2010) argued that students' views about their school environment should be incorporated to improve and understand student outcomes, and previous school climate literature (Bear, Gaskins, Blank, & Chen, 2011; Fan et al., 2011; Gregory et al., 2011; Koth et al., 2008; Lee et al., 2011) has demonstrated that students' racial group membership may impact views of school climate.

Specifically, Bear, Gaskins, Blank, and Chen (2011) established a brief 29-item instrument to assess student perspectives of school climate. Development of the scale was intended to measure four characteristics of the school environment: caring relationships, fairness of rules, school safety, and liking of school. Yet, the DSC-Student Survey was limited by the scale's internal consistency and generalizability.

Internal consistency is the extent to which items on a survey measure the same construct (Henson, 2001), and the reliability of observed scores is "central to understanding the relationship between variables" (Henson, 2001; p.178). Taber (2018) noted that internal consistency reliability coefficients in social science research range

from .60 to above .90 where interpretation of the values (e.g., adequate, good, high) vary. Interestingly, Bear and colleagues (2011) found that when examining the reliability coefficients of the five subscales across subgroups (e.g., grade level, gender, race), reliability coefficients ranged between .67 to .87 for African American students and ranged between .73 to .89 for White students on the Fairness of Rules (e.g., perceptions of school rules and their fairness) subscale. Bear et al. (2011) also examined measurement equivalence across race, ethnicity and grade level. Measurement invariance was supported across three levels (configural, metric, and scalar invariance). Bear and colleagues (2011) further recommended research would benefit from the investigation of measurement invariance of other school climate tools in order to make comparisons across racially and ethnically diverse groups and inform evidence-based interventions.

Student Connection Survey (AIR)-Self Report

As previously mentioned, the Student Connection Survey is a recently employed survey that was developed in partnership with American Institutes for Research and the Chicago Public schools. The survey was created to help explore the association between student-level and school-level predictors on student academic achievement. The result of this partnership was an instrument that was not unique to Chicago schools but could be administered to schools in any educational system (Osher et al., 2008). According to the available information (Osher, 2008), survey development was influenced by consultation with experts and the tenets of the Conditions for Learning framework designed to "capture information from students about the degree to which their school promoted a nurturing and supportive environment for learning" (p. 5). What has been demonstrated in previous literature is that the theoretical framework of Conditions for Learning (i.e.,

safety, social emotional learning, student support) demonstrate robust connections between students' experience and outcomes (Garibaldi et al., 2015).

According to AIR and the Chicago Public School results (Osher et al., 2008), the Student Connection Survey demonstrates good internal consistency across subscales: (social emotional scale; Cronbach's alpha; 0.795); student support scale (Cronbach's alpha 0.767); high expectations (Cronbach's alpha 0.738); and, safe and respectful school climate (Cronbach's alpha 0.772). Additionally, their report claims adequate fit based on Master's partial credit model and Rasch modeling (Osher et al., 2008; Chicago Public School; AIR Survey). However, further evaluation of evidence of validity and reliability of the Student Connection Survey (Osher et al., 2008) was not reported. As a result of the work conducted with the Chicago Public School and AIR, dissemination of the survey to other school districts and state departments of education is ongoing; thus evaluating the psychometric properties of a widely used measure of school climate is needed.

According to Garibaldi and colleagues (2015) "accurate measurement is the foundation of understanding needs, identifying targets for improvement, and monitoring progress over time" (Garibaldi et al., 2015; p. 352). Garibaldi and collaborators (2015) concluded that multiple considerations are needed when selecting a school climate survey and an important factor is that the tool is conceptually and psychometrically sound (Garibaldi et al., 2015). Investigation of the reliability and validity of widely used school climate survey is required (Garibaldi et al., 2015). Additionally, focusing on the extent to which the Student Connection Survey is equivalent across race and gender is warranted, as this level of analysis has not been investigated. Garibaldi et al. (2015) suggest that one

of the issues of school climate survey development is ensuring the survey is measuring what it is intended to measure.

According to *the Standards for Educational and Psychological Testing* (2014) one of the preliminary considerations in developing and evaluating measurement tools is evidence of validity which is the degree evidence and theory support the interpretation of tests scores (Joint Committee on Standards for Educational and Psychological Testing, 2014). There are several types of validity that researchers can investigate. For example, construct validity is the extent to which a test may be said to measure a theoretical construct (Reynolds & Ramsay, 2003), and content validity refers to the extent to which a measure is a representative sample of the trait being measured (Reynolds & Ramsay, 2003). However, issues of measurement bias can impact the interpretation of test scores which may influence any form of validity evidence (e.g., construct, content, etc.). Measurement bias refers to systematic error in the estimation of a value and may be present if a measure overestimates or underestimates scores on a variable for individuals from a cultural group (Reynolds & Ramsay, 2003).

Historically, issues of measurement bias have centered around disparate mean scores by groups based on race or gender (Zumbo, 2007); however, exclusive examination of mean differences is no longer considered to be sufficient (Zumbo, 2007; Zumbo & Koh, 2005), and further examination of how a survey is used across groups (Zumbo, 2007) is recommended. Some researchers (Brown, Reynolds, & Whitaker, 1999; Reynolds & Ramsay, 2003; van de Vijver & Tanzer, 2004; Zumbo, 2007) have underscored instances where issues of systematic estimation of a value (i.e., measurement bias) may be present.

Measurement non-equivalence

A latent construct that is not corresponding across groups equally (Reynolds & Ramsay, 2003; van de Vijver & Tanzer, 2004) may represent construct bias. Specifically, construct bias may occur when a "measure is either more difficult, valid or reliable for one group than for another" (Reynolds & Ramsay, 2003, p.96). Construct bias could occur when the definition of school climate is not the same across student race, or gendered groups.

Item Bias and Differential Item Functioning

Item bias exists when the survey items hold different meanings across groups. (van de Vijver & Tanzer, 2004). Given that measures consist of items, some researchers (Zumbo, 2007; Zumbo & Koh, 2005) were concerned about which specific items of a measure might be the source of bias. The term item bias was replaced in the scientific literature by differential item functioning (DIF) and DIF methods allow for one to distinguish whether there are true differences between groups in the underlying construct being measured (Zumbo, 2007). DIF methods are employed so that "researchers can make group comparisons and rule out measurement artifacts for group differences" (Zumbo, 2007; p. 230).

Several statistical approaches to address measurement bias and differential item functioning (DIF) are available to researchers. One approach includes factor analytic techniques and an extension of factor analysis to explore evidence of validity across groups is referred to as multi-group confirmatory factor analysis (Pendergast, von der Embse, Kilgus, & Eklund, 2017; Zumbo & Koh, 2005). Multi-group confirmatory factor

analysis (MG-CFA) allows for investigating the extent to which the measure is equivalent/invariant across race and gender characteristics (Zumbo & Koh, 2005).

Measurement equivalence/invariance is a tool that is utilized to explore the relationship between responses to items and the construct to ensure they are operating in a comparable way across cultural and gendered groups at different levels (Chen, 2007; Pendergast et al., 2017). Measurement invariance may be investigated across several levels; the first step is examining configural invariance, which is the extent to which items are measuring the same factors across groups (Hong, Malik, & Lee, 2003). The next step includes examining metric invariance and it refers to the magnitude of the relationships between test items and the latent factors are similar across groups (Hong et al., 2003; Meredith, 1993; Pendergast et al., 2017). The third step includes scalar invariance which refers to equal item intercepts and/or thresholds across groups and is examined when metric invariance is supported (Dimitrov, 2010; Meredith, 1993).

Measurement equivalence techniques are necessary to ensure that assessment tools demonstrate evidence of validity for all groups. Researchers that do not address issues of measurement invariance may be stymied by the lack of generalizability of their findings. Measurement tools that have not been shown to demonstrate evidence of validity may not generalize to other instruments measuring the same construct (van de Vijver & Tanzer, 2004). Examining measurement equivalence strengthens and increases the confidence of mean score comparisons, demonstrates evidence in support of structural validity and is a prerequisite before group comparisons are made (Pendergast et al., 2017; Zumbo & Koh, 2005).

Consequences of Non-Equivalence

Measurement bias can manifest in differences in scores across the latent construct, item, and scale level (Chen, 2007; Pendergast et al., 2017; Zumbo & Koh, 2005). Therefore, it is essential to examine the measurement equivalence of a school climate survey before it can be used to make comparisons across heterogenous groups (Chen, 2007). Various methods are employed to explore the extent to which factor loadings and error variances are equal across groups (Zumbo & Koh, 2005). When non-equivalent tools are used, researchers may not capture aspects of school climate that are salient for particular student groups. For instance, using tools that have been empirically supported to be equivalent across race and gendered characteristics allows researchers to effectively evaluate targeted interventions for student groups. Pendergast and colleagues (2017) noted that evidence-based interventions are dependent on the use of scientifically based assessments with diverse student groups. Thus, exploring school climate across student racial and gender characteristics is a necessary step to inform culturally specific interventions and improve school-wide decision-making practices to address school climate.

Statistical Approaches

Few studies have attempted to investigate features of school climate using well established psychometrically sound instruments (Hung, Luebbe, & Flaspohler, 2015) and few have investigated the perspectives of children and adolescents from varied backgrounds. In particular, Zullig and colleagues (2010) were the first to integrate modern scale development techniques (e.g., factor analyses, structural equation modeling) to attempt to establish a reliable and valid measure of school climate among

their sample of 6th-12th grade students. However, Zullig et al. (2010) did not evaluate whether the measure was operating equally across student groups and did not include best practice analytic techniques defined by Fabrigar, Wegener, Maccallum and Strahn (1999). Fabrigar et al. (1999) suggest that researchers incorporate multiple methods of factor retention (e.g., scree test, parallel analysis and eigenvalues), and utilize promax rotation. Instead, Zullig and colleagues used principal components analysis (PCA) and varimax rotation which has been shown to not differentiate between common and unique variances and lead researchers to extract too many factors (Fabrigar et al., 1999). Hence, Zullig et al.'s (2010) school climate survey consisted of eight factors where factor loadings ranged from .42 to .87. Factor loadings were retained if they had high loadings (>.40), did not load on two or more different factors, and contained eigenvalues greater than 1 (Zullig et al., 2010).

Next, the authors expanded upon the limitations of their 2010 study (e.g., mainly White sample, lack of diverse sample; failed to measure against well-established measures of school safety) and examined the survey among a diverse sample of Arizona public high school students. However, the authors chose to focus on only four domains (i.e., positive teacher relationships, academic support, order and discipline, and school physical environment) and did not utilize methodological approaches to explore measurement equivalence before comparing across race, gender or grade level groups. The four domains of the school climate survey were selected because they aligned with the Arizona's grant priorities (Zullig et al., 2014) and explained 36% of the variance in the original study (Zullig et al., 2010).

Like the original study (Zullig et al., 2010) construct validity was examined, and latent and manifest variables were explored utilizing SEM procedures. Despite selecting only four school climate domains, the authors suggested a need to further validate other school climate characteristics (e.g., student support, student connectedness) among a diverse student group. Yet, Zullig et al. (2014) failed to examine the extent to which the survey was equivalent across groups. Methodological approaches such as MG-CFA and examination of measurement invariance are necessary before making comparisons across race and gendered groups. What is encouraging from the Zullig et al. (2014) findings are that the results emphasize the need to strengthen school climate measurement, as consistency in measurement will improve understanding the ways in which school climate influences student outcomes.

Summary

School climate has been influenced by several theoretical frameworks (e.g., Bronfenbrenner, Social Cognitive Theory, Conditions for Learning), and has been shown to influence student, teachers, and educators (Back et al., 2016; Hughes et al., 2008). The available literature indicates that student support, academic expectations, and academic rigor are important protective factors and support a positive school climate for students (Goodenow, 1993; Jia, Konold, & Cornell, 2015).Yet, the preponderance of studies indicate that Black students report more negative views of school climate (Croninger & Lee, 2001; Gregory et al., 2010; Pena-Shaff et al., 2019; Wu et al., 1982), and reveal that students' racial and ethnic identity may influence aspects of their school climate. Specifically, being Black and male has been shown to contribute to a negative school climate, and this relationship may be influenced by punitive and discriminatory practices

such as out of school suspensions (Gregory et al., 2011; Koth et al., 2008; Lee et al., 2011).

Existing literature has shown that punitive practices influence Black students' school experience; however, literature has mostly focused on Black males (Gregory et al., 2010; Skiba et al., 2016). Additional scholarship is required to examine the educational experiences for Black girls, and specifically their school climate. Student perceptions of school climate are often gathered through self-report surveys. However, few studies have attempted to investigate aspects of school climate using well established, psychometrically sound instruments (Hung et al., 2015), and the available school climate surveys are hindered by methodological and statistical limitations described above (e.g., low internal consistency reliabilities, lack of examination of measurement invariance; Bear et al., 2011; Zullig et al., 2014, 2010).

Although about 15% of school climate studies focused exclusively on validating and developing measures of school climate (Wang & Degol, 2016), few studies employ best practices in factor analyses (Fabrigar et al., 1999) and measurement invariance (e.g., examining whether the measure is equal across groups) testing (e.g., Chen, 2007; Pendergast et al., 2017; Reynolds & Ramsay, 2003; Zumbo & Koh, 2005). Investigating the extent to which theory supports the interpretation of scores from school climate measures (e.g., validity) and ensuring that school climate measures are equivalent across groups may help with accuracy and improvement of empirically based school climate assessment which is required for the development of evidence-based interventions.

The AIR and Chicago Public Schools conceptualized school climate as encompassing social emotional learning, student support, feelings of safety and high

expectations/rigor. The AIR and Chicago public schools developed the Student Connection Survey (Garibaldi et al., 2015; Osher et al., 2008) to represent a measure of school climate for targeted interventions. Yet, few studies address the psychometric properties including validity and measurement equivalence of the Student Connection Survey although efforts to extend its use across school districts nationwide continue. Therefore, exploration of construct validity and measurement equivalence of the Student Connection Survey is needed. Given that the psychometric qualities of the Student Connection Survey are rarely reported, it is unclear whether this measure of school climate is corresponding equally across groups (Reynolds & Ramsay, 2003; van de Vijver & Tanzer, 2004). Thus, it is uncertain to what extent the items of the School Climate Survey are similar across students with different backgrounds. Capturing the perceptions of students (i.e., scores) who differ across race, ethnicity and gender using a measurement tool that demonstrates reliability and validity across groups will build upon an area of research that is lacking.

As the U.S. population continues to grow, there is a need for developing evidence-based and culturally relevant interventions that promote positive outcomes to address student needs with specific racial and gendered backgrounds (Jones et al., 2018). Specifically, Jones et al. (2018) found evidence that targeted interventions for Black girls influenced racial development and school engagement among this group. These findings support the notion by Pendergast, von der Embse, Kilgus, and Eklund (2017) that an intermediary step in developing appropriate evidence-based interventions begins with an accurate appraisal of our measurement tools across racial groups.

Particularly, instruments must demonstrate equivalent measurement across groups in order to make accurate conclusions about the construct (Pendergast et al., 2017; Reynolds & Ramsay, 2003; van de Vijver & Tanzer, 2004) and inform evidence-based programs. Examining the observed scores of the Student Connection Survey, for example, (i.e., a measure of school climate) across groups may reveal how individuals from varied backgrounds view their learning environment.

Finally, capturing Black girls' views of school climate across domains (i.e., social emotional learning, support, safety, academic rigor) by examining their scores on a measurement tool that demonstrates reliability and validity with this group is needed. This would allow researchers to effectively evaluate targeted interventions for a group that is disproportionately impacted by school discipline policies and may address their school climate. Black girls encounter punitive discriminatory practices that threaten their positive school experience which may jeopardize successful development around their social emotional well-being, racial identity and school engagement (Chavous et al., 2008; Leath et al., 2019). Understanding Black girls' educational experience in the context of school climate is needed. Appropriately examining how Black girls perceive their school climate is a gap in the existing scientific literature and is required to further Black girl scholarship.

Present Study

The purpose of the current study is to examine Black girls' school climate scores on a widely used school climate survey. Adequately measuring school climate is a necessary step to strengthen the empirical research on school climate and Black girls' scholarship. Research indicates that Black students' perceptions of school climate vary in

relation to their White classmates, but insufficient evidence is available on how Black girls perceive school climate using psychometrically sound measurement tools. Emerging research demonstrates that Black girls are seven times more likely to receive at least one out of school suspension and four times more likely to receive at least one in-school suspension in comparison to their White female counterparts (National Black Women's Justices Institute, 2018).

The literature also notes that school discipline may influence a negative school experience. Development of evidence-based and culturally relevant interventions that promote positive outcomes to address student needs such as school climate with specific racial and gendered backgrounds is needed (Jones et al., 2018). Consequently, a prerequisite in developing appropriate evidence-based interventions begins with an accurate appraisal of measurement tools across racial groups.

The study aims were to (1) examine the underlying factor structure of the Student Connection Survey Student version utilizing exploratory (EFA) and confirmatory factor analysis (CFA), (2) examine the measurement equivalence of the Student Connection Survey across race and gender using multiple-group confirmatory factor analysis; with a specific focus on Black girls' scores.

Research Questions

1. Is the factor structure of the Student Connection Survey as defined as social emotional learning, student support, safety, and high expectation/ rigor supported by the data?

2. Is the factor structure of the school climate survey equivalent across gender and racial groups demonstrated by three levels (i.e., configural, metric, and scalar/threshold) of equivalence?

Hypotheses

1. It is hypothesized that the factor structure of the school climate survey will be supported by the data.

2. It is expected that the school climate survey scores will differ at the metric and scalar/threshold of equivalence as a function of race/ethnicity and gender, as evidence by changes in the model fit. Available research indicates that perceptions of racially and ethnically diverse students are mixed (Bryson & Childs, 2018) but studies (Bottiani et al., 2017; Pena-Shaff et al., 2019) have shown more negative perceptions of student-teacher relationships and school safety, without examining the extent to which measurement invariance is present. Given that student support, and safety are features of the current school climate survey, it is expected that this will influence Black students' and particularly Black girls' scores on the Student Connection Survey.

CHAPTER 3

METHODOLOGY

Descriptive Statistics

A total of 8,357 middle school girls were included in the sample. Students attended schools across 28 counties in the Commonwealth of Pennsylvania. Middle schools in the Commonwealth were given the option to participate in the larger Pennsylvania Department of Education school climate study, and these data represent students from the participating schools. A total of 330 middle schools participated in the climate study. The sample was approximately equal across grade levels for 6th to 8th grade students. One student identified themselves as a 9th grader and was not included in subsequent analyses. Therefore, the final study sample included 8,356 self-identified middle school girls. Students included in the current sample were racially and ethnically diverse: 65% White/Caucasian, 12% Black/African American, 11% Hispanic/Latinx, 8% Multiracial, 5%American Indian/Alaskan Native, 3% Asian, 1% Native Hawaiian or Pacific Islander. Demographics, including race and grade level for the present sample are reported in Table 3.1.Table 3.1 conitnues on page 41.

Table 3.1

Demographics for the Study Sample of Self-Identified Middle School Girls

Race/Ethnicity	(N)	(%)
White/Caucasian	5467	65
Black/African American	1012	12
Hispanic/Latinx	984	8
Multi-Racial	674	11
Unavailable/Unknown/Decline	503	6
American Indian/Alaskan Native	447	5

Table 3.1 (Continued)

Asian	327	3
Native Hawaiian or Pacific Islander	54	1
Grade	8356	
Sixth	2123	25
Seventh	3379	40
Eighth	2854	34

Note. Race and ethnicity represent the item selected by self-identified girls in the Student Connection Survey data collection instrument. Middle school students may have endorsed more than one race and/or ethnicity. In addition, the term "Hispanic" reflects the term used in the Student Connection Survey data collection instruments and census. Other terms that some members of this group may prefer include but are not limited to Latino/Latina/Latinx

Measure

The Student Connection Survey

The Student Connection Survey (Pennsylvania Department of Education and American Institute of Research, 2018) is a 55-item, 4-point, Likert-type scale designed to assess whether students strongly agree or disagree with the provided statement (e.g., *strongly disagree to strongly agree, not safe to very safe*). A copy of the original survey can be found in the Appendix. Responses from the scale in the current study were treated as ordinal (i.e., categorical) data. The measure is designed to assess four domains of school climate: Social Emotional Learning, Student Support, High Expectations/Academic rigor, Safe and Respectful School Climate (Garibaldi et al., 2015; Osher et al., 2008).

Scores are calculated by averaging across items with higher scores indicating greater endorsement of items for each scale and more positive school climate. Therefore, higher levels of social emotional learning, student support, high expectation, and safe and

respectful school environment suggest more positive school climate (Osher et al., 2008). According to Osher et al. (2008), The Student Connection Survey indicated internal consistency of .75 on the Social Emotional subscale, .76 on the Student Support subscale, .78 on the High Expectations subscale, and .77 on the Safe and Respectful School Climate subscales.

Data Analysis

Screening for univariate and multivariate outliers was organized and analyzed using SPSS (Version 23). Descriptive statistics including race, gender, and grade level are reported. Using a random number function in Microsoft Excel, the data were randomly subdivided to the Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) subgroups.

Measurement Model

Exploratory factor analysis

EFA Exploratory factor analysis (EFA) was first analyzed following procedures outlined in Fabrigar et al. (1999). EFA was investigated using principal axis factor analysis with promax rotation. Principal axis factor analysis with promax rotation has been recommended by researchers (Fabrigar et al., 1999) as promax rotation provides more accurate representation of how constructs are likely to be related to one another and permits correlations among factors (Fabrigar et al., 1999). Next, multiple methods were explored to determine the retention of factors including scree test (Cattell, 1966), minimum average partial (MAP), (Velicer, 1976), parallel analysis (Horn, 1965) and interpretability. *A priori* criteria for factor retention included having pattern coefficients > .40 item salience and a minimum of three salient items per factor.

Confirmatory factor analysis

The hypothesized factor structure of the school climate survey generated from the EFA was used for subsequent Confirmatory Factor Analysis (CFA). CFA was conducted in Mplus version 7.1 (Muthén & Muthén, 2015) using weighted least square mean and variance (WLSMV) estimation. The model that demonstrated adequate fit was selected for single group CFA analysis. To conduct the single-group CFA, a separate CFA was evaluated for each of the racial and gendered groups. Model fit was evaluated based on the Chi-Square Goodness of Fit, Root Means Square Error of Approximation (RMSEA; (Steiger & Lind, 1980), Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973)Tucker & Lewis, 1973) and Standardized Root Mean Square Residual (SRMR). *A priori* criteria for acceptable model fit includes: a non-statistically significant chi-square, RMSEA ≤ 0.08, CFI ≥ 0.95, TLI ≥ 0.95(Byrne, 2012; Hu & Bentler, 1999), with chi-square holding the least weight because it is impacted by a large sample size (Kline, 2011).

MG-CFA was used to examine the degree of measurement invariance (MI) across race and gendered groups. MI was examined across configural, metric, and scalar/threshold invariance. Nested models were compared by examining the changes in the model fit. Fit statistics were evaluated for goodness of fit based on the evaluation of changes in CFI values, chi-square, and RMSEA (i.e., Δ CFI $\geq$ -0.01, Δ chi-square $\geq$.05, ΔRMSEA < 0.015). Emphasis was placed on the results of CFI as scholars (e.g., Cheung & Rensvold, 2002; Meade Johnson, Braddy, 2008) have found that the CFI and RMSEA approximate fit index perform relatively with a large sample.

CHAPTER 4

RESULTS

Descriptive Statistics

A random subsample of 200 cases from the total sample (*N=8,356*) were selected for the EFA (Costello & Osborne, 2005).The remaining cases *(N=8,156)* were reserved for subsequent CFA analyses. The sample for measurement invariance testing included White/Caucasian (*N=5330*) and Black/African American (*N=922*) self-identified girls.

Measurement Model

Exploratory Factor Analysis.

Analyses for the EFA were completed in SPSS Version 25. Principal axis factoring with oblique (promax) rotation was used on the initial 53 items. An oblique rotation was chosen as this approach allows for factors to be correlated (Costello & Osborne, 2005). The Bartlett's Test of Sphericity was statistically significant ($p \leq .000$), and the Kaiser Meyer Olkin Measure of Sampling Adequacy was (.834), suggesting acceptability for correlation matrix for factor analysis (Kaiser, 1974).

Multiple factor retention methods including scree test (Cattell, 1966), parallel analysis (PA; Horn 1965), and MAP analysis, (Velicer, 1976) were employed. The use of multiple methods for factor retention safeguards against specifying too few or too many factors (Hayton, Allen, & Scarpello, 2004) and is consistent with best factor analytic practices (Costello & Osborne, 2005; Fabrigar et al., 1999; Hayton et al., 2004), First, findings from the scree test suggested five factors should be retained. Hayton and colleagues (2004) indicate that the scree test "is relatively straightforward to examine, but suffers from subjectivity and ambiguity" (pg., 193).

Likewise, findings from parallel analysis (PA; Horn, 1965) suggested that five factors should be retained. Parallel analysis, a more accurate method for determining the number of factors to retain (e.g., Hayton et al., 2004; Velicer et al., 2000), takes the correlation matrices from a random sample based on the sample size of the real data set and compares the eigenvalues from the random data to the existing data set. Factors corresponding to eigenvalues that are greater than the parallel (randomly generated) average eigenvalues should be retained (Hayton et al., 2004). Velicer's minimum average partial correlation (MAP) test, a method that calculates the average of squared partial correlations until the minimum squared partial correlations are met (Velicer, 1976) suggested a six-factor solution.

In summary, scree and parallel analysis supported a five-factor solution, and MAP supported a six-factor solution. The six-factor solution was rejected because it failed to meet *a priori* criteria including pattern coefficients greater than .40 and having a minimum of three salient items per factor. Pattern coefficients greater than .40 were selected as items less than .40 may not be related to other survey items (Costello & Osborne, 2005). Furthermore, Velicer's MAP approach tends to retain too many factors (Hayton et al., 2004) and theoretically, a six-factor solution was not supported in previous school climate studies (e.g., Bear et al., 2011; Garibaldi et al., 2015; Osher et al., 2008). Therefore, a six-factor solution was not explored in subsequent CFA analyses.

Following the findings from scree and parallel analysis, a five-factor solution was examined. Six items were initially removed due to failing to meet *a priori criteria* including examination of pattern coefficients (e.g., pattern coefficient >.40) and a minimum of three salient items per factor. Multiple iterations of this process were

conducted until all remaining items met criteria. Consequently, thirteen items were removed, and the five-factor solution was comprised of 40 items from the Student Connection Survey.

The five factors were named Student Support from Teachers, Positive Student Behavior, Negative Student Behavior/Bullying, Safety, and High Expectations/Rigor. The Student Support from Teachers factor accounted for 22.93% of the variance. Sample items on this factor included "teacher notices if I have trouble learning something; teachers help me improve my work if I do poorly on an assignment." The Student Support from Teachers factor was comprised of 13 items and demonstrated good (Cortina, 1993) internal consistency (α=.882). The Positive Student Behavior factor accounted for 10.9% of the variance and demonstrated good internal consistency (α=.867). Sample items on the Positive Student Behavior factor included "students try to do a good job on schoolwork even when it is not interesting; students treat others with respect." The Negative Student Behavior/Bullying factor accounted for 7.54% of the variance and was comprised of nine items (α=.862). Sample items on this factor included "students are often teased or picked on" and "students are often bullied because of certain characteristics." The Safety factor accounted for 5.04% of the variance with four items (e.g., "I sometimes stay home because I don't feel safe at school.") that also demonstrated high internal consistency (α=.772). The High Expectations/Rigor factor accounted for 5.04% of the variance and was composed of four items (e.g., write a paper in which you defended your own point of view or ideas) and indicated adequate internal consistency (α= .674). The overall 40 item five-factor model demonstrated good internal consistency (α= .906) and because the five-factor solution met all *a priori* criteria with the

aforementioned adaptations, it was examined in more depth using confirmatory factor analyses.

Confirmatory Factor Analysis

Three models were tested in the current study. These included the EFA derived five factor structure (i.e., baseline model) with and without post hoc modifications and an alternative structure identified by Pendergast et al. (2020) (the model is described in detail on page 41). CFA was conducted in MPLUS using WLSMV estimation among the subsample of 8,157 middle school girls. WLSMV estimation is recommended when utilizing non-normal ordinal data (Beauducel & Herzberg, 2006).

Next, the model was identified by setting one item on each factor to one. The items that were fixed to one (e.g., Item 32, "Teachers will help me improve my work if I do poorly on an assignment"; Item 2, "Safe in the hallways and bathrooms of school") were chosen because they had the highest factor loadings on their respective factors in the prior EFA analyses. Goodness of fit was evaluated across several model fit indices including: a non-statistically significant chi-square, RMSEA $\leq$,0.08 CFI $\geq$, 0.95 TLI $\geq$,0.95 and SRMR$\leq$ .08 (Byrne, 2012; Hu & Bentler, 1999). Preliminary CFA findings are reported in the Appendix A. The model fit indices for the final CFA structure is reported in Table 4.1 found on page 49 (including $\chi 2$ values and p values). Factor names and survey items are reported in Table 4.2 found on page 50.

Baseline Model

The five-factor structure did not meet *a priori* thresholds (See Table 2 in Appendix A). Although the RMSEA (.095) was high, scholars (Hu & Bentler, 1999; Kline, 2011) suggest a single index reflects only one aspect of model fit and a favorable

value is not necessarily indicative of model fit. To explore the potential variables interfering with model fit, post hoc modifications were employed to determine if the model fit would improve.

Post Hoc Modifications

Intercorrelations between the five factors were specified. Three items were removed, and nine variables were correlated based on inspection of modification indices and model fit statistics. Items removed included Items 31 ("Teachers notice if I have trouble learning something") 35 ("I worry about crime and violence in school") and 50 ("Adults in this school are often too busy to give extra help"). In addition to the removal of items, residuals were allowed to be correlated between items. Similarly worded items and items with shared variance on the Safety latent factor were correlated including items 1 ("Outside around the school) 2 ("In the hallways and bathrooms of the school") 3 ("In your classes") and 40 ("I sometimes stay home because I do not feel safe at school").

Next, Items 11 ("Teachers often connect what I am learning to life outside the classroom") and 12 ("Teachers encourage students to share their ideas about things we are studying in class") on the Student Support from Teachers domain were allowed to be correlated. Items (e.g., "Writing a research paper of 2 or more pages") on the High Expectation latent factor were correlated. Fit indices with post hoc modifications demonstrated improved model fit (see Table 4.1).

Alternative Model

Jackson, Gillaspy, and Purc-Stephenson (2009) recommend, when possible, researchers test alternative models. Consequently, a separate alternative structure was investigated. Preliminary research conducted by Pendergast, Estrada, and Schneider

(2020) found four main aspects of school climate, consistent with the four-factor model originally hypothesized by scale developers (Student Support, Safe & Respectful School Climate, Social Emotional Learning and High Expectations, Academic Rigor & Challenge). Although a four factor EFA model was attempted prior to running a CFA for the current sample, examination of pattern coefficients for the four-factor model did not meet *a priori* criteria (a pattern coefficient >.40 and a minimum of three salient items per factor). The factor structure from the Pendergast et al. (2020) study was tested; however, this model did not converge. Therefore, the five-factor model with post hoc modifications was the best-fitting model (see Figure 1) on page 51. The best-fitting model was applied for subsequent equivalence testing.

Table 4.1

Model Fit Indices for the Final Five Factor Model of the Student Connection Survey

	χ^2	*df*	*p*	RMSEA (90% CI)	CFI	TLI
Five Factor	36754.502	618	0.00	0.085 [0.084-0.085]	0.839	0.826

Note. χ^2= chi-square; df_M = degrees of freedom; RMSEA = root-mean-square error of approximation; CI = confidence interval; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index. Post hoc analyses; model fit improved when select items were removed. Items were removed due to cross loadings determined by modification indices. When items PASC31, PASC35, and PASC50 were removed.

Table 4.2
Student Connection Survey Factor Names and Survey Items

Factor 1. Student Support from Teachers
Item 32. Will help me improve my work if I do poorly on an assignment
Item 15. Help me make up work after an excused absence
Item 22. Adults in this school are usually willing to take the time to give students extra help
Item 16. Often assign homework that helps me learn
Item 18. Adults in the school apply the same rules to all students equally
Item 11. Often connect what i am learning to life outside the classroom
Item 12. Encourage students to share their ideas about things we are studying in class
Item 33. The topics we are studying are interesting and challenging
Item 14. Really care about me
Item 53. I am usually bored in this class
Item 52. Treat some students better than others
Factor 2. Positive Student Behavior
Item 7. Do their best even when their schoolwork is difficult
Item 10. Try to do a good job on schoolwork even when it is not interesting
Item 4. Treat each other with respect
Item 48. Say mean things to other students when they think the other students deserves it
Item 9. Try to work out their disagreements with other students by talking to them
Item 5. Stop and think before doing anything when they get angry
Item 47. Think it's OK to fight if someone insults them
Item 49. Think it's OK to cheat if other students are cheating
Item 6. Do their share of the work when we have group projects
Item 8. Do all their homework
Item 42. Like to put others down
Factor 3. Negative Student Behavior
Item 38. Students at this school are often teased or picked on
Item 36. Students at this school are often bullied
Item 39. Students at this school are often bullied because of certain characteristics (ex: race, religion, or weight).
Item 37. Students at this school are often threatened
Item 44. Just look out for themselves
Item 42. Like to put others down
Item 41. Don't really care about each other
Item 43. Don't get along together well.
Factor 4. Safety
Item 2. In the hallways and bathrooms of the school
Item 1. Outside around the school
Item 3. In your classes
Item 40. I sometimes stay home because I don't feel safe at school
Factor 5. High Expectations/Rigor
Item 24. Write a research paper of 2 or more pages
Item 25. Write a paper in which you defended your own point of view or ideas.
Item 26. Make a formal presentation to a class about something you read or researched.
Item 29. Talked to an adult at school about something outside of school that is important to you

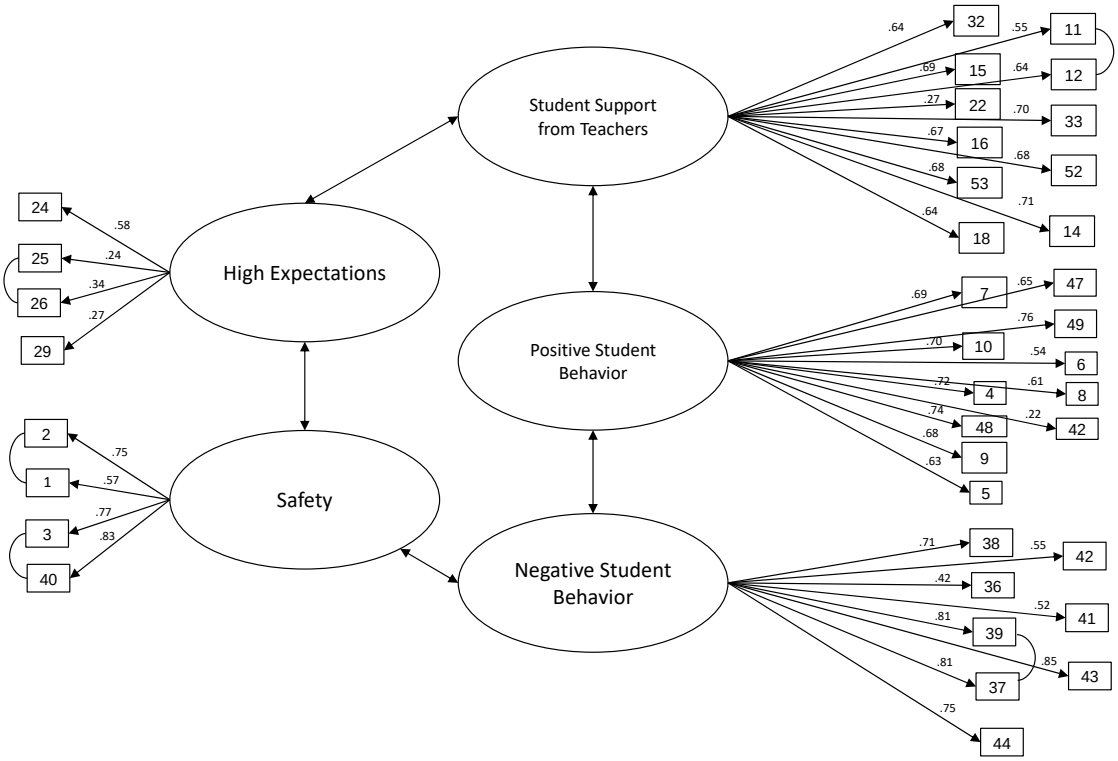

Figure 1. Five-factor structure for the Student Connection Survey among Middle School Self-Identified girls. Standardized pattern coefficients (WLSMV estimation) are presented.

Multigroup Confirmatory Factor Analysis

Participants from two racial groups (Black girls, *n* = 992, and White girls *n* =5330) were included in the measurement invariance analyses. Invariance testing was evaluated across a series of models.

The configural model provides a basis for comparison for subsequent model invariance testing (Dimitrov, 2010) and assumes that the same items load on the same factors across groups (Dimitrov, 2010; Pendergast et al., 2017). The next step in the

multigroup confirmatory factor analysis is examining the metric model. The metric model is more restrictive than the configural model, where factor pattern coefficients are constrained to be equal across groups. When metric invariance is supported, each item contributes to the latent construct to a similar degree across groups (Putnick & Bornstein, 2016). Lastly, scalar invariance assumes that group differences in the means of observed items should be the result of differences in the means of the underlying constructs. In order to test for scalar invariance, item intercepts should be constrained to remain the same across groups (Meredith, 1993; Putnick & Bornstein, 2016).

Detailed MG-CFA findings are presented in Table 4.3. In the first step, configural equivalence was supported. Fit indices for the configural model fell within specified ranges (CFI=0.916; RMSEA = 0.062) indicating that the same items on the Student Connection survey loaded on the same factors for both racial groups. Next, a metric model was tested and demonstrated adequate fit (CFI= 0.918; RMSEA= 0.061). Findings of the metric model indicate each item contributes similarly to the latent construct across racial groups.

Table 4.3

Fit Statistics for Measurement Invariance of the Student Connection Survey Between Black girls and White girls

	χ^2	df	*p*	CFI	ΔCFI	RMSEA	ΔRMSEA
Configural	16041.224	1225	<.001	.916	-	.062	-
Metric	15849.998	1258	<.001	.918	.002	.061	-.001
Scalar/Threshold	15909.408	1322	<.001	.918	.002	.059	-.003

Note. CFI= Comparative Fit Index; RMSEA=root-mean-square error of approximation

Subsequently, a scalar/threshold model was tested. Adequate fit was supported (CFI= 0.918; RMSEA 0.059) and indicates that Black and White girls have equal probability between categories of selecting agree, strongly agree and so forth, on the Student Connection Survey and demonstrate equal item intercepts across racial groups. Therefore, scalar invariance was established. Taken together, configural, metric and scalar equivalence were supported for both Black and White middle school girls and suggests the valid mean comparison of school climate scores can be made across groups.

CHAPTER 5

DISCUSSION

School climate is a construct frequently explored in educational research (Lee, Cornell, Gregory, & Fan, 2011; Wang & Degol, 2016). Prior findings suggest a positive school climate is associated with improved student-teacher relationships (Croninger & Lee, 2001), higher academic achievement (Wang et al., 2014) and reduced school dropout risk (Jia, Konold, & Cornell, 2016). Unfortunately, emerging research indicates that racial differences in schools' penal practices may be negatively coupled with Black students' experience of a supportive school climate (Bottiani, Bradshaw, & Mendelson, 2017).

Research has demonstrated that Black students are disproportionately subjected to exclusionary disciplinary practices (e.g., suspensions and expulsions) in the learning environment (Fenning & Rose, 2007). Specifically, Black girls disproportionally encounter more adverse disciplinary outcomes in their educational setting (e.g., suspensions, expulsions, referral to law enforcement; National Black Women's Justice Institute, 2018) relative to their White peers (Blake, Butler, Lewis, & Darensbourg, 2011; Crenshaw, Ocen, & Nanda, 2015; Epstein, Blake, & González, 2017; Hines-Datiri & Carter Andrews, 2017). Further, persistent use of punitive practices may hinder Black girls' racial identity, academic, and social-emotional development in the classroom (Chavous, Rivas-Drake, Smalls, Griffin, & Cogburn, 2008; Leath, Mathews, Harrison, & Chavous, 2019) and consequently detrimentally impact their school climate.

Researchers have called for evidence-based and culturally relevant interventions that promote positive academic outcomes for Black girls (Jones et al., 2018). However,

evidence-based interventions are dependent on the use of empirically supported assessments with diverse student groups (Pendergast et al., 2017). Evidence-based assessments are necessary to identify students' needs in the educational setting and provide baseline data that allow for evaluation of intervention effectiveness. Moreover, school climate assessments for Black girls may facilitate the development of evidence-based interventions for Black girls - who are disproportionally disciplined and may be at risk for experiencing a negative school climate.

The Student Connection Survey is a 55-item, 4-point, Likert type scale designed to assess students' school climate. Few studies have explored the evidence of the validity of scores from the Student Connection Survey, an instrument designed to inform school-based interventions and overall school climate (Garibaldi et al., 2015). In particular, few studies have investigated the extent to which there is evidence of validity for Black girls. In fact, to the author's knowledge, this was the first study to investigate the structural validity of any widely used school climate tool, with the focus on Black girls' scores.

The *Standards for Educational and Psychological Testing* (2014) underscore the importance of evidence of validity relative to the internal structure of a measure. Moreover, without evidence to support the validity of scores from a given tool, researchers, stakeholders and educators alike are hindered in their ability to provide evidence-based recommendations that are needed to address the needs of students with specific racial and gendered backgrounds.

The current study sought to examine the structural validity and reliability of scores from a school climate measurement tool. Additionally, the study sought to investigate the extent to which The Student Connection Survey is equivalent across racial

groups, with a specific focus on Black girls' scores. It is important to examine Black girls' scores as they represent a group that is disproportionately disciplined in the school setting, which may impact their positive school climate. Further, more research on the educational experiences of Black girls is needed.

Review of Findings

Hypothesis one stated that the original four factor structure (Garibaldi et al., 2015; Osher et al., 2008) of the Student Connection Survey would be supported by the data. This hypothesis was not supported. Instead, findings of the current study reveal a five-factor structure among middle school girls. Although findings from the current study are not consistent with the original factor model, findings of five latent domains of school climate have been largely supported (e.g., Lewno-Dumdie, Mason, Hajovsky, & Villeneuve, 2020) in the literature.

Original Four Factor Structure

The original four factor structure (e.g., Student Connection Survey) is grounded in the Conditions for Learning framework. Students that participated in the original study included over 63,000 male students, and about 68,000 female students. The sample consisted of 62,007 African American students, 50,919 Hispanic students, over 5,000 Asian/Pacific students, 201 Native Americans students and 13,033 students who identified as White (Osher et al., 2008). Further demographic breakdown of the sample, including the number of female students who identified as Black or African American was not provided. Garibaldi et al. (2015) posit that the measure is comprised of four subscales: Social Emotional Learning, Student Support, High Expectations/Rigor, Safe

and Respectful School Climate with internal consistency across subscales range between .72 and .76.

Five Factor Structure Among Middle School Girls

Findings from the current study support a five-factor model among middle school girls, with latent factors that differed from the original four factor model (Garibaldi, et al, 2015; Osher, 2008). Some items fell on different factors from the original scale. Nonetheless, the five specific factors in the current study are Student Support from Teachers, Positive Student Behavior, Negative Student Behavior/Bullying, Safety, High Expectations/Rigor factors. Internal consistency across factors range between .67 and .88.

Student Support from Teachers (Factor 1)

The Student Support from Teachers factor appeared to have a narrower focus compared to factor one in the original study (Garibaldi et al., 2015). Items on the current factor assess whether students feel supported by their teacher through encouragement, opportunities to share their ideas in class, and if their teacher notices when they have trouble learning something. Perceived teacher support has been linked to increased emotional adjustment for middle school youth, and emphasize the importance that teacher relationships may have on youth's adjustment to middle school and ultimately school climate (Reddy, Rhodes, & Mulhall, 2003; Way, Reddy, & Rhodes, 2007). Evidence of a student support factor is consistent with the original factor (Garibaldi et al., 2015; Osher et al., 2008) and previous research (e.g., Bear et al., 2011; Zullig et al., 2014) examining school climate among diverse adolescent youth.

In addition to the evidence of a Student Support from Teacher factor, the current factor demonstrated higher internal consistency (.88) relative to the .76 internal

consistency reported by scale developers (Garibaldi et al., 2015; Osher et al., 2008). Current findings support the reliability and consistency of a Student Support from Teachers factor among middle school self-identified girls with varied backgrounds, including Black girls. Booker and Lim (2018) found when teachers exhibit affirmative interpersonal relationships and created a sense of belonging, this strengthened Black adolescent girls' positive learning experience and continued performance in math instruction. Additional investigation using reliable survey items that ask female students about how they view teacher support or other student-teacher relationships in middle school, a period marked by increased demands, larger class sizes, and extensive involvement with teachers (Reddy et al., 2003), is warranted. Such exploration may help to establish evidence-based interventions to strengthen student-teacher support and school climate.

Positive Student Behavior (Factor 2)

Overall, the Social Emotional Learning (SEL) factor found in the original structure (Garibaldi et al., 2015; Osher et al., 2008) appeared to split into two factors in the current study. For instance, The Positive Student Behavior factor demonstrated higher internal consistency (.86) relative to the original SEL subscale (.75). Items on the Positive Student Behavior reflect qualities of social emotional learning, and the ability to identify ways in which students show empathy towards others.

Negative Student Behavior (Factor 3)

Separately, The Negative Student Behavior/Bullying factor represents opposite behaviors found on the Positive Student Behavior factor and includes items such as "Students at this school are often teased or picked on" and "Students at this school are

often bullied." It is possible that in middle school, a time when peer relationships become even more important and incidents of bullying increase (Werth et al., 2015), girls may be experiencing prosocial and negative behaviors at the same time. Although the inclusion of a Negative Student Behavior/Bullying factor was surprising, previous research (e.g., Van der Graaf 2014; Werth et al., 2015) revealed gender differences in forms of victimization, where girls experience more relational bullying and are more sensitive to recognizing the harm of bullying (Werth et al., 2015). It will be important for future studies to replicate the current findings and determine whether evidence of Negative Student Behavior/Bullying factor is reinforced among a separate sample of middle school girls. Results may help explain an important domain that could influence adolescent girls' school climate.

Safety (Factor 4)

The Safety factor was supported in the current model demonstrating similar internal reliability relative to the original four factor structure (.77). Evidence of Safety domain is consistent with previous school climate research (e.g., Bear et al., 2011; Thapa et al., 2013).

High Expectations (Factor 5)

The final High Expectations/Rigor factor was supported in the current sample of middle school girls with an internal consistency of .67. Internal consistency of the High Expectations factor is lower relative to factors in the current model and when comparing the High Expectation factor in the original scale (.78; Garibaldi et al., 2015). Lower internal consistency may be indicative of fewer survey items that measure the same construct (e.g., "Talk to an adult about something"). Research indicates high expectations

are often characterized by student engagement and when students are engaged in high level tasks (Buehler, Fletcher, Johnston, & Weymouth, 2015). Perhaps the current factor may be strengthened with additional related items that ask adolescent girls when they feel their teacher holds high expectations of them or when they feel they are engaged in a task. Additional research is necessary to explore sets of items representative of high expectations from adolescent girls' perspectives. Nonetheless, the current five-factor model underscores evidence of validity and reliability of scores among a diverse sample of adolescent girls.

Measurement Invariance of the Five Factor Structure

The next research question sought to answer whether the Student Connection Survey was equivalent across racial groups. That is, whether the measurement items are equivalent across different conditions (e.g., configural, metric, scalar/threshold models). Hypothesis two stated that the school climate survey scores would differ at the metric and scalar/threshold level as a function of race, demonstrated by changes in the model fit. Findings indicate configural, metric, and scalar/threshold invariance across racial-ethnic groups. Although these findings were unexpected, findings are consistent with previous research examining school climate scores among Black and White high school students (Bottiani et al., 2016). Current findings suggest that The Student Connection Survey is likely a psychometrically sound measurement tool and may be useful to evaluate targeted interventions for middle school self-identified girls. While the psychometrics of the scale are acceptable, (CFI; .839; TLI; .826), there is room for improvement and opportunity to strengthen this tool, particulary among its use with middle school girls.

Lastly, the Student Connection Survey may be a valuable tool for those who wish to further understand the shared perspectives of school climate across middle school girls. Black girls face academic and social barriers that are connected to structural forces such as racism, sexism and classism (Evans-Winters & Esposito, 2010), and there is a need to examine their resilience and needs in the context of these risks (Chavous & Cogburn, 2007). It is encouraging that evidence of structural validity and reliability of Black girls'scores is supported in the current study given that little is known about their academic needs (Young, Young, & Capraro, 2017). Specifically, current findings from the High Expectations factor suggest lower reliability among girls, but future applied research could be done to strengthen this factor among adolescent girls, with a focus on Black girls' perspectives. In fact, research has found that being academically challenged is important for Black girls and that Black girls are eager to see themselves fully integrated into the curriculum and feel that their voices are heard and respected (Mims & Kaler-jones, 2020; Morris, 2015). Therefore, scholars may wish to incorporate The Student Connection Survey that demonstrates reliable and valid scores among Black girls to further explore the high expectations/rigor domain. It will be important to build upon school climate measurement as well as Black girl scholarship in order to establish a foundation for evidence-based interventions that support them within the educational environment. One approach is examining the evidence-based assessment and exploring their perspectives that promote a positive school climate.

Strengths of the Present Study

Among the strengths of this study were the novel findings examining the structural validity of a commonly used tool, The Student Connection Survey. This is the

first study to examine the validity and reliability of the survey amongst a sample of middle school youth as well as evaluating the perceptions of over 8,000 middle school girls. Specifically, the data included a sample of over 1,000 girls who identified as Black or African American and expand the research for how Black girls perceive their learning environment. The current study bridges the gap of psychometric studies examining the scores of Black middle school girls, a group that is largely underrepresented in current psychometric school climate scholarship.

Limitations

The present study is limited by its focus solely on middle school student self-report and does not use other methods (e.g., teacher report, observational methods) to assess school climate. In light of demonstrating that the Pennsylvania Student Connection Survey is equivalent between racial groups, we know that penal practices impact the extent to which students experience a positive school climate. The current study did not examine whether scores from the Student Connection Survey predict out of school suspensions or other negative outcomes known to impact school climate, but it is likely to a be a useful tool for evaluating differences in school climate and may predict academic or disciplinary outcomes. Future studies may also include girls from additional varied ethnoracial backgrounds. Lastly, the current data do not fully account for nesting at the school level and does not address the configuration of middle schools. That is, whether additional grade levels (e.g., 5th grade) are also included in the school. This is in part due to the variables that were made available to the researcher for the current study.

Future Research

In addition to some of the recommendations described above, further investigation of the Student Connection Survey among all participants, including individuals who identified as Hispanic/Latinx or biracial, is warranted. Examining the measurement invariance across different racial groups from the current study would further reinforce its use among diverse groups within the educational setting. This approach may also highlight similar or new latent factors that are salient for middle school girls. Future research is also needed to clarify the High/Expectations/Academic Rigor factor. Targeted efforts to address the items and definition of the the High Expectations. Moreover, examining Black girls' understanding of their learning environment is imperative to understanding their unique experience in the school setting.

Black girls' scores have not been the focus of existing school climate or psychometric work and the current project contributes significantly to that area of scholarship. Future studies should explore whether Black girls with lower school climate scores experience higher levels of out of school suspension, or discipline sanctions at the school level. Previous literature uncovered the ways in which Black girls' school climate may be threatened and future research should explore how scores may predict out of school suspensions. This may be accomplished empirically by using structural equation modeling approaches. Doing so may help to identity points of intervention. Separately, researchers may want to investigate whether these school climate domains are supported by incorporating mixed methodology including semi-structured interviews, or focus groups to highlight Black girls' reports of support from their teachers, positive and negative student behavior, safety, and high expectations in their school setting.

As demonstrated in the current study, understanding Black girls' educational experience in the context of school climate is needed. The Student Connection Survey provides evidence of validity and reliability among Black girls, with item indicators ensuring that they feel supported by teachers, feel safe within and outside of their school community, and that they are academically challenged to promote a positive school learning environment. With accurate measurement of Black girls' perspectives of school climate, researchers, teachers, and stakeholders may make appropriate data-based decisions about selecting evidenced-based and culturally informed programs that bolster their learning community and promote an affirming school climate.

REFERENCES

Allen, A., Scott, L. M., & Lewis, C. W. (2013). Racial microaggressions and African American and Hispanic students in urban schools: A call for culturally affirming education. *Interdisciplinary Journal of Teaching & Learning*, *2*, 117–229.

American Academy of Pediatrics Committee on School Health. (2003). Out-of-school suspension and expulsion. American Academy of Pediatrics Committee on School Health. *Pediatrics*.

American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools? *American Psychological Association*. https://doi.org/10.1037/0003-066X.63.9.852

Back, L. T., Polk, E., Keys, C. B., & McMahon, S. D. (2016). Classroom management, school staff relations, school climate, and academic achievement: testing a model with urban high schools. *Learning Environments Research*, *19*(3), 397–410. https://doi.org/10.1007/s10984-016-9213-x

Bandura, A. (1986). *Social foundations of thought and action : a social cognitive theory. Englewood Cliffs, N.J: Prentice-Hall*.

Bear, G. G., Gaskins, C., Blank, J., & Chen, F. F. (2011). Delaware School Climate Survey—Student: Its factor structure, concurrent validity, and reliability. *Journal of School Psychology*, *49*(2), 157–174. https://doi.org/10.1016/j.jsp.2011.01.001

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*. https://doi.org/10.1207/s15328007sem1302_2

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological*

Bulletin, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Blake, J. J., Butler, B. R., Lewis, C. W., & Darensbourg, A. (2011). Unmasking the inequitable discipline experiences of urban Black girls: Implications for urban educational stakeholders. *The Urban Review*, *43*(1), 90–106. https://doi.org/10.1007/s11256-009-0148-8

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2016). Inequality in Black and White high school students' perceptions of school support: An examination of race in context. *Journal of Youth and Adolescence*, *45*(6), 1176–1191. https://doi.org/10.1007/s10964-015-0411-0

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. https://doi.org/10.1037/edu0000155

Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, Cultural Pluralism, and school safety. *Journal of Educational Psychology*, *95*(3), 570–588. https://doi.org/10.1037/0022-0663.95.3.570

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge: Harvard University Press.

Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. (1978). Elementary school social climate and school achievement. *American Educational Research Journal*, *15*(2), 301–318. https://doi.org/10.3102/00028312015002301

Brown, R. T., Reynolds, C. R., & Whitaker, J. S. (1999). Bias in mental testing since Bias in Mental Testing. *School Psychology Quarterly*, *14*(3), 208–238. https://doi.org/10.1037/h0089007

Bryson, S., & Childs, K. (2018). Racial and ethnic differences in the relationship between school climate and disorder. *School Psychology Review*, *47*(3), 258–274.

Buehler, C., Fletcher, A. C., Johnston, C., & Weymouth, B. B. (2015). *Perceptions of School Experiences During the First Semester of Middle School*. *School Community Journal* (Vol. 25). Retrieved from http://www.schoolcommunitynetwork.org/SCJ.aspx

Byrne, B. M. (2012). *Structural equation modeling with Mplus : basic concepts, applications, and programming*. *Multivariate applications series*.

Carter, P. L., Skiba, R., Arredondo, M. I., & Pollock, M. (2016). You cant fix what you dont look at: Acknowledging race in addressing racial discipline disparities. *Urban Education*, 1–29. https://doi.org/10.1177/0042085916660350

Catalano, R. F., Haggerty, K. P., Oesterle, S., Fleming, C. B., & Hawkins, J. D. (2004). Findings from the Social Development Research Group. *Journal of School Health*, *74*(7), 252–261.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chavous, T., & Cogburn, C. D. (2007). Superinvisible Women: Black Girls and Women in Education. *Black Women, Gender & Families*.

Chavous, T. M., Rivas-Drake, D., Smalls, C., Griffin, T., & Cogburn, C. (2008). Gender

matters, too: The influences of school racial discrimination and racial identity on academic engagement outcomes among African American adolescents. *Developmental Psychology*, *44*(3), 637–654. https://doi.org/http://dx.doi.org/10.1037/0012-1649.44.3.637

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Christle, C. A., Jolivette, K., & Nelson, C. M. (2007). School Characteristics Related to High School Dropout Rates. *Remedial and Special Education*, *28*(6), 325–339. https://doi.org/10.1177/07419325070280060201

Cohen, J., McCabe, E., Michelli, N., & Pickeral, T. (2009). School Climate: Research, policy, practice, and teacher education. *Teachers College Record*, *111*(January), 180–213.

Coll, C. G., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., & García, H. V. (1996). An Integrative Model for the Study of Developmental Competencies in Minority Children. *Child Development*. https://doi.org/10.1111/j.1467-8624.1996.tb01834.x

Collaborative for Social and Emotional Learning. (2015). *CASEL Guide: Effective social and emotional learning programs—middle and high school edition*. *Collaborative for Academic, Social and Emotional Learning. CASEL*. https://doi.org/10.1016/S1357-2725(03)00211-5

Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-

9010.78.1.98

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*.

Crenshaw, K., Ocen, P., & Nanda, J. (2015). Black girls matter: Pushed out, overpoliced and undeprotected. *African American Policy Forum*. https://doi.org/10.1002/ejoc.201201096

Croninger, R. G., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record*. https://doi.org/10.1111/0161-4681.00127

Dimitrov, D. M. (2010). Testing for Factorial Invariance in the Context of Construct Validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121–149. https://doi.org/10.1177/0748175610373459

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-Emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, *88*(2), 408–416. https://doi.org/10.1111/cdev.12739

Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Weissberg, R. P., Schellinger, K. B., Durlak, J. A., … Taylor, R. D. (2011). The impact of enhanding students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1007/s10972-014-9381-4

Epstein, R., Blake, J., & González, T. (2017). *Girlhood Interrupted: The Erasure of Black Girls' Childhood*. *SSRN*. https://doi.org/10.2139/ssrn.3000695

Evans-winters, V. E., & Esposito, J. (2010). Other people's daughters: Critical race feminism and black girls' education. *Educational Foundations*.

Fabrigar, L., Wegener, D., MacCallum, R., & Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Pyschological Methods*, *4*(3), 272–290.

Fan, W., Williams, C., & Marie Corkin, D. (2011). Group counseling in the school. *Hellenic Journal of Psychology*, *11*(3), 169–183. https://doi.org/10.1002/pits

Fenning, P., & Rose, J. (2007). Overrepresentation of in exclusionary discipline: The role of school policy. *Urban Education*. https://doi.org/10.1177/0042085907305039

Fergus, E., Noguera, P., Martin, M., & Harvard University, G. S. of E. (2014). *Schooling for Resilience: Improving the Life Trajectory of Black and Latino boys*. *Harvard Education Press*.

Ferguson, R. F. (2003). Teachers' perceptions and expectations and the Black-White test score gap. *Urban Education*, *38*(460), 1–49. https://doi.org/10.1177/0042085903254970

Ford, D. Y., Harris, J. J., Tyson, C. A., & Trotman, M. F. (2002). Beyond deficit thinking: Providing access for gifted African American students. *Roeper Review*, *24*(2), 52–58. https://doi.org/10.1080/02783190209554129

Garibaldi, M., Ruddy, S., Kendziora, K., & Osher, D. (2015). Assessment of climate and conditions for learning. *Handbook of Social and Emotional Learning: Research and Practice.*

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *The Journal of Early Adolescence*.

https://doi.org/10.1177/0272431693013001002

Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal*, *48*(4), 904–934. https://doi.org/10.3102/0002831211398531

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin? *Educational Researcher*, *39*(1), 59–68. https://doi.org/10.3102/0013189X09357621

Gregory, Anne, 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. https://doi.org/10.1037/a0018562

Gregory, Anne, & Mosely, P. M. (2004). The discipline gap: Teachers' views on the over-representation of african american students in the discipline system. *Equity and Excellence in Education*, *37*(1), 18–30. https://doi.org/10.1080/10665680490429280

Gregory, Anne, & Ripski, M. (2008). Adolescent trust in teachers: Implications for behavior in the high school classroom. *School Psychology Review*, *37*(3), 337–353. Retrieved from https://www.researchgate.net/profile/Anne_Gregory2/publication/228621819_Adolescent_trust_in_teachers_Implications_for_behavior_in_the_high_school_classroom/links/00463518004a1e4d27000000/Adolescent-trust-in-teachers-Implications-for-behavior-in-the-high-

Haynes, N., Emmons, C., & Ben-Avie, M. (2010). School Climate as a Factor in Student

Adjustment and Achievement. *Journal of Educational and Psychologial Consultation*, *8*(3), 321–329. https://doi.org/10.1207/s1532768xjepc0803_4

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organizational Research Methods*. https://doi.org/10.1177/1094428104263675

Hendron, M., & Kearney, C. A. (2016). School climate and student absenteeism and internalizing and externalizing behavioral problems. *Children and Schools*, *38*(2), 109–116. https://doi.org/10.1093/cs/cdw009

Henson, R. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*, *34*(3), 177–189.

Hines-Datiri, D., & Carter Andrews, D. J. (2017). The Effects of Zero Tolerance Policies on Black Girls. *Urban Education*, 1–22. https://doi.org/10.1177/0042085917690204

Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-western sample. *Educational and Psychological Measurement*. https://doi.org/10.1177/0013164403251332

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, J. N., Luo, W., Kwok, O. M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-Year longitudinal study. *Journal of Educational Psychology, 100*(1), 1–14. https://doi.org/10.1037/0022-0663.100.1.1

Hung, A. H., Luebbe, A. M., & Flaspohler, P. D. (2015). Measuring school climate: Factor analysis and relations to emotional problems, conduct problems, and victimization in middle school students. *School Mental Health, 7*(2), 105–119. https://doi.org/10.1007/s12310-014-9131-y

Jackson, D. D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods, 14*(1), 6–23. https://doi.org/10.1037/a0014694

Jia, Y, Way, N., Ling, G., Yoshikawa, H., Chen, X., Hughes, D., … Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: a comparison of chinese and american adolescents. *Child Dev, 80*(5), 1514–1530. https://doi.org/10.1111/j.1467-8624.2009.01348.x

Jia, Yuane, Konold, T. R., & Cornell, D. (2016). Authoritative school climate and high school dropout rates. *School Psychology Quarterly, 31*(2), 289–303. https://doi.org/10.1037/spq0000139

Jones, J. M., Lee, L. H., Matlack, A., & Zigarelli, J. (2018). Using sisterhood networks to cultivate ethnic identity and enhance school engagement. *Psychology in the Schools*, *55*(1), 20–35. https://doi.org/10.1002/pits.22087

Kline, R. B. (2011). *Principles and practice of structural equation modeling:New York: Guilford Press Google Scholar*. *Structural Equation Modeling*. https://doi.org/10.1038/156278a0

Kohl, D., Recchia, S., & Steffgen, G. (2013). Measuring school climate: An overview of measurement scales. *Educational Research*. https://doi.org/10.1080/00131881.2013.844944

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. https://doi.org/10.1037/0022-0663.100.1.96

Leath, S., Mathews, C., Harrison, A., & Chavous, T. (2019). Racial Identity, Racial Discrimination, and Classroom Engagement Outcomes Among Black Girls and Boys in Predominantly Black and Predominantly White School Districts. *American Educational Research Journal*, 1–33.

Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High Suspension Schools and Dropout Rates for Black and White Students. *Education and Treatnent of Children*, *34*(2), 167–192.

Lewno-Dumdie, B. M., Mason, B. A., Hajovsky, D. B., & Villeneuve, E. F. (2020). Student-Report Measures of School Climate: A Dimensional Review. *School Mental Health*. https://doi.org/10.1007/s12310-019-09340-2

Mattison, E., & Aber, M. S. (2007). Closing the achievement gap: The association of racial climate with achievement and behavioral outcomes. *American Journal of Community Psychology*, *40*(1–2), 1–12. https://doi.org/10.1007/s10464-007-9128-x

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*(3), 235–261. https://doi.org/10.1016/j.jsp.2007.05.001

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Mims, L. C., & Kaler-jones, C. (2020). Running , running the show : Supporting the leadership development of Black girls in middle school Running , running the show : Supporting the leadership development of Black girls in middle school. *Middle School Journal*, *51*(2), 16–24. https://doi.org/10.1080/00940771.2019.1707342

Morris, E. W., & Perry, B. L. (2017). Girls Behaving Badly? Race, Gender, and Subjective Evaluation in the Discipline of African American Girls. *Sociology of Education*, *90*(2), 127–148. https://doi.org/10.1177/0038040717694876

Morris, M.W. (2015). *Pushout: The criminalization of Black girls in schools.* New York: The New Press.

Morris, M. W. (2012). Race, Gender and the School-to-Prison Pipeline: Expanding Our Discussion to Include Black Girls. *African American Policy Forum*. https://doi.org/10.1146/annurev.soc.28.110601.141048

Neal, L. V. I., McCray, a. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *The Journal of Special Education*, *37*(1), 49–57. https://doi.org/10.1177/00224669030370010501

Osher, D., Kendziora, K., & Chinen, M. (2008). *Student Connection Research: Final Narrative Report to the Spencer Foundation*. Washington. Retrieved from https://www.air.org/sites/default/files/downloads/report/Spencer_final_report_3_31_08_0.pdf

Patton, L. D., Crenshaw, K., & Haynes, C. (2016). Why we can't wait: (Re)Examining the opportunities and challenges for black women and girls in education (Guest Editorial. *Journal of Negro Education*. https://doi.org/10.7709/jnegroeducation.85.3.0194

Payton, J., Weissberg, R., Durlak, J., Dymnicki, A., Taylor, R., Schellinger, K., & Pachan, M. (2008). *The positive impact of social and emotional learning for kindergarten to Eighth-Grade students*. *Collaborative for Academic, Social and Emotional Learning (CASEL)*. Retrieved from https://files.eric.ed.gov/fulltext/ED505370.pdf

Pena-Shaff, J., Bessette-Symons, B., Tate, M., & Fingehut, J. (2019). Race and ethnic differences in high school students' perceptions of school climate and discipline practices. *Race Ethnicity and Education*, *22*(2), 269–284.

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, *60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*. https://doi.org/10.1016/j.dr.2016.06.004

Reddy, R., Rhodes, J. E., & Mulhall, P. (2003). The influence of teacher support on student adjustment in the middle school years: A latent growth curve study. *Development and Psychopathology*. https://doi.org/10.1017/S0954579403000075

Reynolds, C. R., & Ramsay, M. C. (2003). Bias in Psychological Assessment: An

Empirical Review and Recommendations. In I. B. Weiner (Ed.), *Handbook of Psychology* (Second, pp. 82–113). John Wiley & Sons, Inc. https://doi.org/10.1002/0471264385.wei1004

Roth, P. L., & Switzer, F. S. (1999). Missing Data: Instrument-Level Heffalumps and Item-Level Woozles. *Research Methods Forum*, *4*(5).

Ruck, M. D., & Wortley, S. (2002). Racial and ethnic minority high school students' perceptions of school disciplinary practices: A look at some Canadian findings. *Journal of Youth and Adolescence*. https://doi.org/10.1023/A:1015081102189

Skiba, R. (2000). *Zero Tolerance, zero evidence: An analysis of school disciplinary practice*.

Skiba, R. J., Chung, C.-G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, *51*(4), 640–670. https://doi.org/10.3102/0002831214541670

Skiba, Russell J., Arredondo, M. I., & Williams, N. T. (2014). More Than a Metaphor: The Contribution of Exclusionary Discipline to a School-to-Prison Pipeline. *Equity and Excellence in Education*. https://doi.org/10.1080/10665684.2014.958965

Skiba, Russell J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, *34*(4), 317–342. https://doi.org/10.1023/A:1021320817372

Skiba, Russell J, Chung, C.-G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing Disciplinary Disproportionality. *American Educational Research*

Journal, *51*(4), 640–670. https://doi.org/10.3102/0002831214541670

Skiba, Russell J, Peterson, R. L., Williams, T., Education, S., Children, T., August, N., … Williams, T. (2016). Office Referrals and Suspension : Disciplinary Intervention in Middle Schools Stable URL : http://www.jstor.org/stable/42900491 Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use , available at Office Referrals and Su, *20*(3), 295–315.

Slaughter-Defoe, D. T., & Carlson, K. G. (1996). Young African American and Latino children in high-poverty urban schools: How they perceive school climate. *The Journal of Negro Education*, *65*(1), 60. https://doi.org/10.2307/2967368

Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. *Annual Meeting of the Psychometric Society*.

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, *99*(2), 253–273. https://doi.org/10.1037/0022-0663.99.2.253

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. https://doi.org/10.3102/0034654313483907

Theriot, M. T., Craun, S. W., & Dupper, D. R. (2010). Multilevel evaluation of factors predicting school exclusion among middle and high school students. *Children and*

Youth Services Review, *32*(1), 13–19. https://doi.org/10.1016/j.childyouth.2009.06.009

Thomas, A. R. (1976). The organizational climate of schools. *International Review of Education*. https://doi.org/10.1007/BF00598815

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *54*(2), 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Van Horn, M. L. (2003). Assessing the unit of measurement for school climate through psychometric and outcome analyses of the school climate survey. *Educational and Psychological Measurement*, *63*(6), 1002–1019. https://doi.org/10.1177/0013164403251317

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Voelkl, K. E. (1995). School warmth, student participation, and achievement. *Journal of Experimental Education*. https://doi.org/10.1080/00220973.1995.9943817

Waasdorp, T. E., Lindstrom Johnson, S., Shukla, K. D., & Bradshaw, C. P. (2020). Measuring School Climate: Invariance across Middle and High School Students. *Children & Schools*. https://doi.org/10.1093/cs/cdz026

Wallace, J. M., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, Ethnic, and Gender Differences in School Discipline among U.S. High School Students:

1991-2005. *The Negro Educational Review*.

Wang, M.-T., & Degol, J. L. (2016). *School climate: A review of the construct, measurement, and impact on student outcomes*. *Educational Psychology Review* (Vol. 28). https://doi.org/10.1007/s10648-015-9319-1

Wang, W., Vaillancourt, T., Brittain, H. L., McDougall, P., Krygsman, A., Smith, D., … Hymel, S. (2014). School climate, peer victimization, and academic achievement: Results from a multi-informant study. *School Psychology Quarterly*, *29*(3), 360–377. https://doi.org/10.1037/spq0000084

Watkins, N. D., & Aber, M. S. (2009). Exploring the relationships among race, class, gender, and middle school students' perceptions of school racial climate. *Equity and Excellence in Education*, *42*(4), 395–411. https://doi.org/10.1080/10665680903260218

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*. https://doi.org/10.1007/s10464-007-9143-y

Wu, S. C., Pink, W., Crain, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review*, *14*(4), 245–303. https://doi.org/10.1007/BF02171974

Wun, C. (2016). Unaccounted Foundations: Black Girls, Anti-Black Racism, and Punishment in Schools. *Critical Sociology*, *42*(4–5), 737–750. https://doi.org/10.1177/0896920514560444

Young, J. L., Young, J. R., & Capraro, M. M. (2017). Black Girls' Achievement in

Middle Grades Mathematics: How Can Socializing Agents Help? *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*. https://doi.org/10.1080/00098655.2016.1270657

Zullig, K. J., Collins, R., Ghani, N., Patton, J. M., Scott Huebner, E., & Ajamie, J. (2014). Psychometric Support of the School Climate Measure in a Large, Diverse Sample of Adolescents: A Replication and Extension. *Journal of School Health*, *84*(2), 82–90. https://doi.org/10.1111/josh.12124

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. https://doi.org/10.1177/0734282909344205

Zumbo, B. D. (2007). Three Generations of DIF Analyses: Considering Where It Has Been, Where It Is Now, and Where It Is Going. *Language Assessment Quarterly*, *4*(2), 223–233. https://doi.org/10.1080/15434300701375832

Zumbo, B. D., & Koh, K. H. (2005). Manifestation of differences in item-level characteristics in scale-level measurement invariance tests of multi-group confirmatory factor analyses. *Journal of Modern Applied Statistical Methods*, *4*(1), 275–282. https://doi.org/10.22237/jmasm/1114907040

APPENDIX A
PRELIMINARY MODEL FIT INDICES

	χ^2	*df*	*p*	RMSEA (90% CI)	CFI	TLI
Five Factor	54133.195	729	0.00	0.095 [0.094-0.095]	0.779	0.764

Note. This was a preliminary model. χ^2= chi-square; df_M = degrees of freedom; RMSEA = root-mean-square error of approximation; CI = confidence interval; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

APPENDIX B
ORIGINAL SURVEY

PA School Climate Survey

Student Grade 6 to 8 Climate Survey Printable Form

We have a police officer/school resource officer (SRO) in our school: (Choose one)

- ☐ Yes
- ☐ No
- ☐ Not Applicable

Rank how safe you believe your school is because of the school police officer. (1-just as safe, 2-somewhat safer, and 3-much safer; choose one)

- ☐ My school is just as safe as before the school police/SRO came to the school
- ☐ My school is somewhat safer as before the school police/SRO came to the school
- ☐ My school is much safer as before the school police/SRO came to the school

The school police officer helps make our middle school a safer place. (Choose one)

- ☐ Yes
- ☐ No
- ☐ Not Applicable

Which category best describes your Ethnicity/Race? (One or more categories may be marked)

- ☐ American Indian/Alaskan Native
- ☐ Black/African American
- ☐ Hispanic
- ☐ Multi-Racial
- ☐ White/Caucasian
- ☐ Asian
- ☐ Native Hawaiian or other Pacific Islander
- ☐ Unavailable/Unknown/Decline

What grade are you currently enrolled this school year?

_____

What is your Gender? (Choose one)

- ☐ Male
- ☐ Female

How much do you agree with the following statements about your school:

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. I worry about crime and violence in school.	○	○	○	○
b. Students at this school are often bullied.	○	○	○	○
c. Students at this school are often threatened.	○	○	○	○
d. Students at this school are often teased or picked on.	○	○	○	○
e. Students at this school are often bullied because of certain characteristics (ex: race, religion, or weight).	○	○	○	○
f. I sometimes stay home because I don't feel safe at school.	○	○	○	○