The Multilevel Effects of School Climate and Gender on Academic Achievement in Urban High School Students

by

REBECCA D. TAYLOR
B.A., Pomona College, 2002
M.A., University of Illinois at Chicago, 2008

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate College of the University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Roger P. Weissberg, Chair and Advisor
Mary C. Murphy
Stephanie Riger
Herbert J. Walberg, Education (Emeritus)
Kimberly Kendziora, American Institutes of Research
ACKNOWLEDGMENTS

I am indebted to many people for their assistance in the formation and completion of this dissertation. First, I would like to thank my advisor and chair, Roger Weissberg, for his input, support, and patience throughout this process. I also thank my other committee members—Mary Murphy, Stephanie Riger, Herb Walberg, & Kim Kendziora—for their support and assistance in developing and refining the ideas within this document. I would also like to acknowledge Rachel Gordon for her assistance in the proposal stage of my project.

I am incredibly grateful to my colleagues for their encouragement and assistance throughout this project and my graduate career. I would also like to thank the American Institutes of Research for developing and collecting these innovative measurements of student climate perceptions, and for allowing me access to such a rich database. Finally, I would like to acknowledge the financial support that the Graduate College of the University of Illinois at Chicago provided to me during the completion of this project in the form of a University Fellowship.

RDT
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>A. The Problem of Inequality</td>
<td>3</td>
</tr>
<tr>
<td>B. School Climate Perceptions</td>
<td>4</td>
</tr>
<tr>
<td>C. Stereotype Threat and its Impact on Academic Achievement</td>
<td>10</td>
</tr>
<tr>
<td>D. The Potential Role of Climate in Stereotype Threat Responses</td>
<td>12</td>
</tr>
<tr>
<td>E. Covariates Predicting Individual Student Achievement</td>
<td>13</td>
</tr>
<tr>
<td>F. Covariates Predicting School Aggregate Achievement</td>
<td>14</td>
</tr>
<tr>
<td>G. Conditions of the Current Study</td>
<td>15</td>
</tr>
<tr>
<td>II. METHOD</td>
<td>17</td>
</tr>
<tr>
<td>A. Sample and Participants</td>
<td>17</td>
</tr>
<tr>
<td>B. Student Measures</td>
<td>18</td>
</tr>
<tr>
<td>C. School-level Covariates</td>
<td>23</td>
</tr>
<tr>
<td>D. Procedures: Collection of Measures</td>
<td>24</td>
</tr>
<tr>
<td>E. Procedures: Data Analysis</td>
<td>25</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>29</td>
</tr>
<tr>
<td>A. Correlations and Mean Differences</td>
<td>30</td>
</tr>
<tr>
<td>B. Modeling Student-Level Relationships with Covariates</td>
<td>31</td>
</tr>
<tr>
<td>C. Modeling School-Level Relationships with Covariates</td>
<td>31</td>
</tr>
<tr>
<td>D. Multilevel Latent Covariate Model</td>
<td>33</td>
</tr>
<tr>
<td>E. Alternative Models</td>
<td>42</td>
</tr>
<tr>
<td>F. Summary</td>
<td>43</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>44</td>
</tr>
<tr>
<td>A. The Relationship of Climate and Academic Achievement</td>
<td>47</td>
</tr>
<tr>
<td>B. Stereotype Threat in an Urban School District</td>
<td>49</td>
</tr>
<tr>
<td>C. Can Climate Impact Stereotype Threat?</td>
<td>50</td>
</tr>
<tr>
<td>D. Limitations and Delimitations</td>
<td>52</td>
</tr>
<tr>
<td>E. Conclusion</td>
<td>54</td>
</tr>
<tr>
<td>V. REFERENCES</td>
<td>55</td>
</tr>
<tr>
<td>VI. TABLES</td>
<td>66</td>
</tr>
<tr>
<td>VII. FIGURES</td>
<td>75</td>
</tr>
<tr>
<td>VIII. APPENDICES</td>
<td>89</td>
</tr>
<tr>
<td>A. Appendix A</td>
<td>90</td>
</tr>
<tr>
<td>B. Appendix B</td>
<td>94</td>
</tr>
<tr>
<td>C. Appendix C</td>
<td>103</td>
</tr>
<tr>
<td>IX. VITA</td>
<td>135</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. BIVARIATE CORRELATIONS BETWEEN ALL VARIABLES USED IN THE MULTILEVEL MODELS</td>
<td>67</td>
</tr>
<tr>
<td>II. MEAN SCORES ON THE PLAN TESTS AND PERCEPTIONS OF CLIMATE BY GENDER AND RACE</td>
<td>68</td>
</tr>
<tr>
<td>III. UNSTANDARDIZED REGRESSION WEIGHTS FOR COVARIATES AT THE STUDENT LEVEL</td>
<td>69</td>
</tr>
<tr>
<td>IV. INTRACLUSTER (I.E., INTRASCHOOL) CORRELATIONS AND DESIGN EFFECTS</td>
<td>70</td>
</tr>
<tr>
<td>V. UNSTANDARDIZED REGRESSION WEIGHTS FOR COVARIATES AT THE SCHOOL-LEVEL</td>
<td>71</td>
</tr>
<tr>
<td>VI. PARAMETER ESTIMATES FOR SCHOOL-LEVEL COVARIATES WITHIN EACH SUBJECT BY CLIMATE MODEL</td>
<td>72</td>
</tr>
<tr>
<td>VII. MAIN EFFECTS OF STUDENT LEVEL CLIMATE WITHIN THE MULTILEVEL COVARIATE MODEL</td>
<td>73</td>
</tr>
<tr>
<td>VIII. MAIN EFFECTS OF SCHOOL-LEVEL CLIMATE WITHIN THE MULTILEVEL COVARIATE MODEL</td>
<td>74</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student and school-level components of the multilevel covariate model</td>
<td>76</td>
</tr>
<tr>
<td>2. Multilevel latent covariate model for PLAN Math scores</td>
<td>77</td>
</tr>
<tr>
<td>3. Multilevel latent covariate model for PLAN Science scores</td>
<td>78</td>
</tr>
<tr>
<td>4. Multilevel latent covariate model for PLAN Reading scores</td>
<td>79</td>
</tr>
<tr>
<td>5. Multilevel latent covariate model for PLAN English scores</td>
<td>80</td>
</tr>
<tr>
<td>6. Simple slopes for interaction effects between individual-level climate and gender</td>
<td>81</td>
</tr>
<tr>
<td>7. Simple slopes for interaction effects between race and gender on PLAN math tests</td>
<td>83</td>
</tr>
<tr>
<td>8. Simple slopes for interaction effects between race and gender on PLAN science tests</td>
<td>84</td>
</tr>
<tr>
<td>9. Simple slopes for interaction effects between school-level climate and gender on PLAN science tests</td>
<td>85</td>
</tr>
<tr>
<td>10. Simple slopes for interaction effects between school-level climate and gender on PLAN math tests</td>
<td>87</td>
</tr>
</tbody>
</table>
SUMMARY

This study investigates the relationship between students’ perceptions of their school’s psychosocial climate and their performance on academic achievement tests. These climate perceptions are taken from a survey of 27,203 students in the 10th and 11th grades from 109 high schools within an urban school district. The tests are a national standardized academic achievement exam with subscales in math, science, English, and reading. A multilevel latent covariate modeling approach was used to test the hypothesis that school-level and individual-level perceptions of positive psychosocial climate uniquely predict better academic outcomes. I found support for this hypothesis at the school-level for all climate perceptions and all subject tests, but individual-level climate perceptions showed mixed relationships with test scores. Additionally, based on the predictions of stereotype threat theory, I expected that female students would underperform in the evaluative tests of their math and science abilities. Holding student grade point average and other characteristics constant, female students scored significantly lower than males on Math and Science tests and higher on the English and Reading tests. I also hypothesized that the benefits of a positive school climate might narrow these gender-based achievement gaps, but the gender-by-climate interactions did not support this conclusion. The results indicate the importance of considering climate measurements at the school-level, the persistence of stereotype threat in a natural school environment, and the need for further exploration of the impact psychosocial climates might have on achievement gaps.
The Multilevel Effects of School Climate and Gender on Academic Achievement in Urban High School Students

There is a great diversity in the education experience of high school students and in the academic success that they achieve. Some of this variety can be accounted for by the characteristics of schools themselves. Schools that have features such as lower teacher turnover, strong principal leadership, and more clear and consistent disciplinary procedures provide very different environments, learning opportunities, and outcomes than those that lack these features (Guin, 2004; Hallinger, Bickman, & Davis, 1996; Ma & Klinger, 2000). However, two students within the same school often have disparate schooling experiences as well. Students who, for example, come from a lower socio-economic status or have learning disabilities may either be treated differently or interpret the school environment differently than their peers (Dodge, Pettit, & Bates, 1994; Kagan, 1990). Thus, a student’s educational experience and its impact upon their academic achievement is a result of the combination of school features and student features.

The combination of these school and individual features can produce an effect that is not only additive, but also interactional, such that some school features have a stronger impact on some types of students. For example, a classroom environment that is supportive and fair may boost learning for all students, but students who are high in self-criticism may benefit even more than those who are low in self-criticism (Kuperminc, Leadbeater, & Blatt, 2001). This person-environment fit approach suggests that students may be more satisfied, engaged, and productive in contexts that are closer to their preferred environment (Fraser & Fisher, 1983) or that respond to their age-specific developmental needs (Eccles et al., 1993). Thus, a key part of
understanding the gaps in students’ achievement in school and adjustment at different ages is the 
ability to understand and characterize their school environments.

There is a long tradition of psychological researchers in the United States directing the 
attention of educators to the psychosocial climate within schools. At the turn of the 20th century, 
reformers like John Dewey emphasized the interactive process of learning within the constructed 
social institution of the school (1900) and the role of schools in not just teaching academic skills, 
but also developing the full potential of the child (1897). Mid-century theorists built upon this 
view of the classroom as a social system (Thelen & Geltz, 1957), and researchers began to 
quantify the environmental press towards certain goals and behaviors found in different college 
and cultural norms also inspired those studying primary and secondary schools to investigate the 
prominent features characterizing those environments (Herr, 1965; Trickett & Moos, 1973; 
Walberg, 1969). Major dimensions of school climate emerged from this research to shape the 
assessment of school and classroom climate in the past few decades (see reviews in Anderson, 
1982; Fraser, 1987; Taylor, 2011).

In the current fields of positive youth development (PYD) and social and emotional 
learning (SEL), school climate has become an important part of the theoretical framework 
underlying these person-environment models (Durlak et al., 2007). This framework positions the 
school, family, and community systems as contexts for the development of personal 
competencies and healthy behaviors (Collaborative for Academic, Social, and Emotional 
Learning, 2003; Catalano, Berglund, Ryan, Lonczack, & Hawkins, 2002; Pittman, Irby, Tolman, 
Yohalem, & Ferber, 2003). Furthermore, PYD and SEL research has shown that interventions to 
promote healthy and supportive contexts can enact significant change in these systems and
benefit students’ social, emotional, and academic growth (Durlak et al., 2007; Tseng et al., 2002; Weissberg, Kumpfer, & Seligman, 2003). Students who possess strong prosocial skills and encounter positive school climates experience improved well-being and academic performance, both during the intervention and into the future (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Taylor, Weissberg, & Durlak, 2011).

This body of research on the personal competencies and contextual factors that contribute to the growth of the whole child has led to a new wave of climate-directed policy efforts throughout the country. Ohio has developed specific School Climate Guidelines, school personnel in California fill out a yearly School Climate Survey, and Wisconsin’s Standards of the Heart are designed to help create supportive learning environments. The U.S. Department of Education has even included school settings as one of the top six funding priorities in Race to the Top (i.e., School-Level Conditions for Reform, Innovation, and Learning), with an emphasis on “creating school climates and cultures that remove obstacles to, and actively support, student engagement and achievement” (U.S. Department of Education, 2009).

**The Problem of Inequality**

One of the persistent problems in education that climate reform might have the potential to impact is the inequality of educational outcomes for marginalized or stigmatized groups, otherwise known as the achievement gap (Weinstein, 2002). These gaps are found along racial, socio-economic, and gender divisions, and favor the groups with higher social status. For example, gender differences favoring males in mathematics performance and attitudes have been found cross-culturally (Mullis et al., 1998). These patterns, however, suggest cultural influence instead of inborn ability differences: gender differences tend to grow with age and experience (Maccoby & Jacklin, 1974) and vary culturally in relation to such factors as mother’s
employment, homework expectations, and father’s scientific-related employment (Harnisch, Steinkamp, Tsai, & Walberg, 1986). Additionally, the gender differences commonly found in tests of mathematical aptitude appear to be due to something more than differences in accumulated experience and preparation. Researchers from the Educational Testing Service (Wainer & Steinberg, 1992) found that women’s performance in college math courses was underestimated by their SAT Mathematics scores; compared to their male counterparts who received the same grades in the same classes, women had underscored them on the SAT-M by an average of 33 points.

In order to better understand how climate and stigmatized identities might interact to predict student adjustment and achievement in high-risk samples, I first explore the relationship between adolescents’ experiences of school climate and their well-being and academic achievement. Findings from this area illuminate the potential of school climate to affect certain processes and fill certain needs for adolescents. Then, I discuss the needs of and challenges experienced by students with stigmatized identities, and whether the benefits of a positive school climate have the potential to address these specific needs and concerns. Finally, I present a few additional characteristics of students and schools that may impact academic achievement, and outline how they will be combined with climate perceptions and gender in the current model to predict student achievement test scores.

School Climate Perceptions

When we talk about a school’s psychosocial climate, there is often some confusion about the exact topic being discussed. It is helpful to think of the distinction drawn by theorists (e.g., Fraser & Walberg, 1981; Moos, 1979; Tagiuri, 1968; Walberg, 1976) between the observable environment and the experiential psychosocial climate. Although all these theorists draw this
conceptual distinction, the terms “environment” and “climate” are not consistently used as labels for these separate concepts. The choice of this terminology is based primarily on the model constructed by Tagiuri (1968). The environment contains all the observable characteristics of a setting, such as the number of monitors in a lunchroom or the frequency of student questions during a lesson. In contrast, the climate encompasses the perceived qualities of the environment as experienced by the members of that setting, such as a sense of safety or competition. For example, the number of students in a classroom is a feature of the environment, but the amount of personal attention that students feel they receive from their teacher is a climate perception.

**Multilevel climate.** Climate is also a complex notion because it includes both personal and situational components. While any one student’s climate perceptions are a reflection of his or her personal experience in the school setting, climate is generally conceptualized to be a quality of the school itself that exists and endures beyond any individual student. Thus, climate is inherently a multilevel variable, with the school climate as a whole represented by the aggregate of student-level perceptions. There are a number of ways to approach the formation of these aggregate school-level climate measures (Chan, 1998), but they all result in scores that are intended to reflect relatively stable, collective experiences of students that reflect the relationships, norms, and culture of that specific school.

There is a good deal of evidence to suggest that a collective climate does have some influential effect above and beyond individual’s perceptions. Haertel, Walberg, and Haertel’s (1981) meta-analytic review of the relationship between climate and learning outcomes found a significantly larger correlation between school-level measures of these variables than between student-level measures. Additionally, many studies using multilevel models have found a significant relationship between school-level climate and student outcomes (see review in
Sellström & Bremberg, 2006). This justifies the need to model the psychosocial climate at both the individual and the school-level to properly determine its impact on students.

**Climate and development.** Climate perceptions may also be particularly connected to academic achievement during certain periods of a student’s life. For example, the 10th and 11th grades represent a post-transition period in high school when many students will have developed informed impressions of their school’s culture. Students in these grade levels may be particularly susceptible to those climate factors that increase a sense of belonging and academic goal setting, as they may be facing specific post-transition risks. A decline in the use of effective study habits has been observed for high-school sophomores and juniors (Slate, Jones, & Dawson, 1993). Additionally, it is a time period when students are at risk of opting out of school entirely; the legal dropout age in many states is typically reached during 10th or 11th grade. The reasons for student dropout during these grades shift compared to the transition period of 9th grade; dropout rates due to disengagement from school and academic problems increase, while those due to disciplinary problems decrease (Stearns & Glennie, 2006).

**Multidimensional climate.** In addition to being inherently multilevel, school climate is also generally conceptualized as multidimensional. While one might speak casually of schools that have “good” or “bad” climates, school climate measurements typically create several subscales that reflect specific elements of the climate on which schools can vary rather than global assessments of quality (Trickett & Moos, 1973; Walberg & Anderson, 1968). For example, one major influence on the field in the conceptualization of these elements are the distinctions specified by Moos (1979) between three broad domains: relationship dimensions, personal development dimensions, and system change and system maintenance dimensions. Other lines of research, such as the exploration of perceived differential treatment using the
Teacher Treatment Inventory (Kuklinski & Weinstein, 2000; McKown & Weinstein, 2008), have focused on one very specific part of the classroom climate. However, both types of research see the classroom environment as a rich, multidimensional construct. The focus differs in whether they attempt to partial out one aspect of this or to distinguish multiple factors. Many different reviews have covered the diversity of domains specified in climate models (e.g., Anderson, 1982; Fraser, 1998; Taylor, 2011); what is detailed below is a brief summary of the way in which some of these domains have demonstrated a relationship to student wellbeing and academic achievement.

**Safety.** When students feel a sense of safety within their school, they feel physically, socially, and emotionally secure. An unsafe environment does not just include problems of physical violence, but also emotional bullying. Students in unsafe schools are more likely to experience depression, while students in safe schools are more likely to experience higher academic expectations and efficacy (Brand, Felner, Shim, Seitsinger, & Dumas, 2003). Additionally, an unsafe school environment negatively impacts all students, not just those involved in the bullying; students who witness bullying are also at risk for poor mental health outcomes (Rivers, Poteat, Noret, & Ashurst, 2009).

**Teacher expectations and interactions.** One of the personal development dimensions that is included in many school climate assessments is the degree to which the environmental press includes high expectations for academic engagement and effort from students. Student perceptions of their teachers’ expectations are negatively correlated with discipline problems, and positively correlated with their engagement in the classroom (Murdock, 1999). Furthermore, student perceptions of differential treatment and expectations in their classroom predicts the size of the achievement gap, with classrooms higher in perceived differential expectations having a
larger year-end achievement gap favoring White and Asian students over Black and Latino students (McKown & Weinstein, 2008).

Teacher-student interactions are not just limited to conveyed expectations; students also experience school and classroom climates that differ in the degree of warmth, personal attention, and assistance they receive from teachers. Students who experience a caring relationship with their teachers are likely to have a higher satisfaction with school (Baker, 1999).

**Supportive relationships and belonging.** Students who experience supportive relationships with their teachers and peers feel a greater sense of connection to their school community. Students’ perceptions of connectedness to school serve to protect them from a variety of risky behaviors throughout junior high and high school (Resnick et al., 1997). Furthermore, student perceptions of support and belonging in their classrooms increase their academic motivation, particularly among girls (Goodenow, 1993). This increased sense of belonging has been found to be a key mechanism mediating the relationship between positive teacher-student relationships and adolescent’s positive affect towards school (Roeser, Midgley, & Urdan, 1996). The perception of prosociality among and emotional support from their peers can also boost students’ own prosocial goals and behaviors, as well as their academic motivation and efficacy (Harter, 1996; Wentzel, 1994; Wentzel & Caldwell, 1997).

**Peer conflict and disengagement.** While a sense of connection with peers and the presence of a supportive climate can boost motivation and achievement, the perception of antisocial behaviors among schoolmates can be detrimental to students. At the individual-level, there is a strong relationship between student affiliation with deviant peers and school failure and future delinquency (Kasen, Cohen, & Brook, 1998). Additionally, perceptions of friction amongst ones classmates tend to be associated with greater conduct problems and depression
(Loukas & Robinson, 2004). A negative association has also been found between experiences of peer harassment and students’ connection to school and grade point average (Eisenberg, Neumark-Sztainer, & Perry, 2003). Additionally, students who perceive their classmates as disengaged from the classroom are less likely to earn good grades (Moos & Moos, 1978), and students who see their peers as resistant to academic values are more likely to have maladaptive learning goals (e.g., avoidance due to fear of failure; Nelson & DeBacker, 2008).

**Institutional norms.** In addition to the perceptions of teachers and schoolmates, a student’s sense of school climate can also reflect a more diffuse sense of the norms and expectations conveyed at the institutional level throughout the school. For example, the concept of *academic press* suggests that schools differ in the degree to which they push students academically. Several studies have found that academic press is a necessary component for translating social support into academic achievement within schools. Schools that have features such as a high enrollment of students in challenging classes (Phillips, 1997) or the perception that students are held to high academic standards (Lee & Smith, 1999) tend to have a higher achieving student population. Additionally, it appears that a strong norm of academic values and practices within the setting may be most important for low-SES schools (Shouse, 1996).

These examples are only a sampling of the evidence for the relationship between school climate perceptions and academic motivation and performance. However, there is less understanding of the mechanisms by which climate influences students and for whom the impact is most powerful. Incorporating research on the psychological mechanisms of academic motivation and academic performance into school climate models may help to illuminate the process by which a positive school climate can impact students’ well-being and achievement.
Stereotype Threat and its Impact on Academic Achievement

One of the internal processes proposed to impact students’ academic motivation and achievement is the pressure and anxiety resulting from stereotype threat. When students encounter situations that make salient a negative stereotype about their group, it can interfere with their academic performance in that domain. Stereotype threat theory (STT) suggests that the extra pressure of being judged against the negative stereotype and consciously trying to avoid confirming it undermines student performance (Steele, 1997). Studies have suggested that this process consumes precious executive functioning resources (Beilock, Rydell, & McConnell, 2007; Schmader & Johns, 2003) and diminishes creative problem solving (Seibt & Forster, 2004), which can leave students facing these stereotype threats at a cognitive disadvantage. This cognitive deficit is particularly relevant when we discuss the impact of stereotype threat on academic achievement assessments.

Steele and Aronson (1995) proposed STT as an explanation for the underperformance of stigmatized groups that was contextually based, as opposed to theories that propose innate differences or accumulated experiential differences as the underlying mechanisms. STT posits a person-environment interaction in which certain contextual cues make salient a negative stereotype that creates a potentially threatening situation, but that individuals respond differently to that situation based on factors such as their self-efficacy or identification with the topic being assessed. Many instances of stereotype threat influencing academic performance have since been demonstrated in the literature, primarily among female students, students of color, and low SES students. The current study will focus primarily on the effect of stereotype threat on female students, but will simultaneously model potential race-based stereotype threat effects and interactions with gender.
In the past two decades, research has clearly shown that female students are inclined to experience stereotype threat in situations that make salient the gender-based stereotype of women’s inferior math and science ability (e.g., Spencer, Steele, & Quinn, 1999), and women who most highly identify with math and science may experience the greatest deficits (Gresky, Ten Eyck, Lord, & McIntyre, 2005). Additionally, several studies suggest that actually invoking the stereotype of inferior math ability for females is unnecessary to induce threat. Simply placing female students in an evaluative testing condition designed to test their mathematical intelligence or ability (as opposed to a low-stakes problem solving exercise) induces poorer performance (e.g., Marx & Stapel, 2006; Schimel, Arndt, Banko, & Cook, 2004). Danaher and Crandall (2008) looked at data from real-world administrations of an Advanced Placement Calculus test and found evidence that collecting demographic information about students’ gender and race before the exam (the typical practice) was enough to significantly increase the gender gap compared to collection after the exam.

In addition to simply demonstrating the presence of stereotype threat within math and science contexts, the STT literature has presented several pathways by which female students’ performance drops as a result of decreased motivation and expectations. Women’s expectations for their own performance drop in the face of stereotype threat induction (Kray, Thompson, & Galinsky, 2001; Stangor, Carr, & Kiang, 1998). Women also seem to not only expect poorer results, but also to set lower goals for their performance under stereotype threat conditions (Kray, Galinsky, & Thompson, 2002). Moreover, stereotype threat effects seem to be particularly damaging and difficult to counter for those who have lower self-esteem (Rydell & Boucher, 2010).
While the pervasive and almost automatic nature of stereotype threat response has a disconcerting impact on female students, there is also plenty of evidence that suggests that there are routes to alleviating this threat. For example, Spencer, Steele, and Quinn (1999) found that, for students with a background in math, simply telling participants that their test had never shown gender differences eliminated women’s underperformance. This reduction of stereotype threat effects has also been demonstrated in the naturally occurring school environment with a general population of high school students (Keller & Dauenheimer, 2003). Additional strategies that seem to alleviate stereotype threat include focusing on a different social identity that is not negatively associated with performance (Gresky et al., 2005; Rydell & Boucher, 2010), directly challenging the stereotype to create an environment where students no longer feel threatened by their identity (Davies, Spencer, & Steele, 2005), affirming a sense of self-integrity (Cohen, Garcia, Apfel, & Master, 2006), and cultivating a malleable as opposed to a fixed view of ability and intelligence (Good, Aronson, & Inzlicht, 2003).

**The Potential Role of Climate in Stereotype Threat Responses**

Many of the deficits observed in stereotype threat situations and the personal characteristics associated with its impact have the potential to be affected by school climate. For example, given the finding that women tend to lower their individual expectations in response to stereotype threat (Kray et al., 2002), an environment of perceived high expectations might counter that tendency and thereby reduce female students’ underperformance. Also, there is reason to believe that safe and supportive climates could interfere with the pathways from negative self-evaluation to underperformance. Youth high in self-criticism typically experience many social and emotional problems, but that relationship is not observed among students who perceive a positive school climate (Kuperminc et al., 2001). Additionally, in the high school
sample examined by Keller and Dauenheimer (2003), the experience of dejection mediated the impact of stereotype threat experiences on test performance. This suggests that minimizing student experiences of dejection might counter the impact of stereotype threat. Perceptions of teacher support and high academic expectations have been shown to reduce feelings of depression and increase self-esteem in adolescents (LaRusso, Romer, & Selman, 2008; Way & Robinson, 2003).

Given what we know about the fairly automatic induction of gender-based stereotype threat in relevant situations, it is reasonable to assume that female students face this same impairment when confronted with standardized achievement tests of their math or science abilities. However, considering what we know about conditions that alleviate the pressure of stereotype threat and the potential power of school climate to impact attitudes and behaviors, it is also reasonable to assume that there are some climates that might help to alleviate this pressure. Thus, certain climates might provide an extra boost to female students above and beyond what it provides their male peers. However, we would not expect to observe this interaction in subjects such as reading comprehension or literature, where there are no gender-based stereotypes of lower performance from female students.

**Covariates Predicting Individual Student Achievement**

In addition to exploring the affect of gender on achievement and the interaction of gender and climate perceptions in stereotyped academic subjects, several other key student characteristics are worth considering. The first of these important covariates is race. There is ample evidence that a persistent gap in academic achievement exists between White and Asian students and Black and Latino students (Jenks & Phillips, 1998; Lee, 2006). Similarly large and persistent gaps are found for students from low socio-economic status (SES) families, making
poverty the second important covariate to be included. SES and race are interrelated, particularly among urban populations, and both poverty and Black or Latino identities are associated with lower rates of many of the important precursors and predictors of academic success, such as parent participation and safe schools (Barton, 2003). Additionally, stereotype threat theory (Steele, 1997) suggests that these gaps are partially the result of negative academic stereotypes for these populations and thus the threatened identities these students experience during academic evaluation, regardless of the resources available to them. The third covariate that is controlled for in predicting academic achievement test scores is student grade point average in the first semester. As a measure of academic performance, student grades should be highly related to test scores, and examining student performance on high-stakes exams relative to normal classroom performance is one way of identifying stereotype threat. In the current study, for example, it is expected that girls will underperform on Math and Science exams relative to boys when their grade point average is controlled for. Lastly, student grade level (10th or 11th) should be controlled as the it is expected that student scores on the PLAN test will increase as students’ academic experience increases.

**Covariates Predicting School Aggregate Achievement**

There are also many characteristics of schools other than the psychosocial climate that may predict the collective level of achievement. To begin with, schools with higher poverty rates typically have fewer resources and face more challenges, resulting in lower overall academic motivation and achievement (Battistich, Solomon, Kim, Watson, & Schaps, 1995). Students’ academic achievement is reduced in high poverty schools, regardless of a student’s own SES (Caldas & Bankston, 1997). Additionally, schools that have high student turnover rates suffer from many negative academic outcomes; high mobility rates can be a source of
disruption even for non-mobile students and that can slow the pace of instruction and lower test scores (Rumberger, 2003). Another school feature that historically and frequently appears when discussing achievement is school and classroom size. Links have been suggested between total enrollment in high-schools and drop-out rates (Felter, 1989), and larger class sizes in high-school math courses are related to greater teacher time spent on behavior management and less time on innovative and engaging instructional methods (Rice, 1999). Finally, in this study will also consider the racial diversity of the school body as a predictor of academic outcomes. Greater racial diversity has been found to have both positive and negative impacts on the school experience. For example, Lucas and Berends (2002) found that schools with greater ethnic diversity were more likely to exhibit “de facto tracking”, in which students tend to enroll in courses of similar difficulty even though no formally assigned system exists. Conversely, greater ethnic diversity in school has also been found to promote feelings of security and higher self-worth among Black and Latino students (Juvonen, Nishina, & Graham, 2006). In the current study, the impact of school diversity on academic achievement will be examined, as well its potential interaction with students’ racial identity.

**Conditions of the Current Study**

The current study is a test of these complex relationships within a diverse, urban school district with a sample of students in the 10th and 11th grades. I examine the role of (a) students’ perceptions of the school climate; (b) student gender; and (c) covariates at the school and the student levels in predicting (d) their individual math, science, reading comprehension, and literature scores from a national standardized achievement test. In order to explore differences in achievement between students within schools, and between schools within the district, I used a
multilevel modeling approach. Student climate perceptions are also examined as both an individual-level and school-level factor within this multilevel model.

Before examining the effect of climate perceptions in a complex model, it is necessary to establish the domains of climate that are under examination. A recent attempt to distinguish multiple elements of school climate was undertaken by the American Institutes of Research (AIR); the Collaborative for Academic, Social, and Emotional Learning (CASEL); and the Learning First Alliance (LFA) in partnership with Chicago Public Schools (CPS) with the goal of assessing the social and emotional conditions for learning (Osher & Kendziora, 2010). Drawing on the growing research literature on the features of school environments that promote social, emotional, and academic development (Osher et al., 2008), four domains were identified for assessment within the Conditions for Learning survey (CFL): safe environments, supportive relationships, high expectations, and the social and emotional skills of peers. However, further analyses of multiple student samples that had completed the CFL revealed that these theoretical factor structures did not hold (Rosenberg, Bohrnstedt, & Ahadi, 2010). The current study uses the CFL items to assess student climate perceptions, but in light of these findings, I did not use the original four domains. Instead, I use factor analyses of the items to both (a) illustrate the multidimensional nature of climate and (b) establish a revised domain structure using a reduced set of the original items. These revised climate scales are then used throughout my analyses.

The following two hypotheses are tested in this study:

1. Perceptions of climate in multiple domains will be positively related to achievement at both (a) individual and (b) school-levels; and

2. Female students will have a stronger individual-level relationship between climate and achievement than male students, but only for math and science tests.
Significant relationships between the various climate measures and academic achievement tests would confirm previous findings of climate’s importance, as well as potentially clarify which domains of climate have the greatest impact on academic achievement. Additionally, a significant difference between females and males on the impact of climate in math and science tests would indicate that climate might interfere with the stereotype threat process. A stronger relationship between climate and math and science scores for girls as compared to boys would support the idea that climate has the potential to alleviate the stereotype threat typically experienced in these domains.

METHOD

Sample and Participants

The database analyzed in this study comes from two sources. The student level data were collected by the American Institutes of Research (AIR) in collaboration with the Chicago Public Schools (CPS) during the 2006-2007 school year. These data include the Conditions for Learning (CFL) climate survey, the PLAN academic achievement tests, student demographic and school performance variables, and the school attended. The school-level data were collected from publicly available, nonconfidential databases of the characteristics of CPS schools during the 2006-2007 school year (CPS, n.d.).

These two data sources were merged to create a multilevel database of students nested within schools. The data in this project are from students attending 109 different high schools within the CPS district. The average cluster size (i.e., number of students sampled per school) is 250 ($SD=244$; Range = 11 to 1525). The data were cleaned to ensure that the school ID number matched for the students’ first and second semesters, as any difference could indicate a transfer in schools. Additionally, each student had to have replied to a sufficient number of the CFL
items (see below) to be included in the analyses. Almost 31,000 students from the 10th and 11th grade completed the CFL survey. The substantial majority of the students (N = 27,203) met both of the above qualifications.

A slightly higher percentage of the student participants in this sample come from the 10th grade cohort (55%) than the 11th grade cohort (45%). The sample also includes slightly more female students (55%) than male students (45%). The students come from diverse racial backgrounds (45% Black, 37% Latino, 11% White, 6% Asian, 1% Native American), and most (85%) of their families are poor enough to qualify for free lunch status (i.e., an income of $26,000 or less annually for a family of four).

**Student Measures**

In order to measure the criterion of student academic achievement in different domains, this study uses student-level scores on the subject subscales of a standardized academic achievement test. The primary predictors of interest were student gender and school psychosocial climate, which was based on students’ responses to items assessing their perceptions. Student grade point average (GPA), grade level, student race, and free lunch status were included in the model as covariates.

**Student academic achievement.** Students who completed the climate scale during the Spring 2007 semester were matched with the primary outcomes of interest in this study: student scores on science, math, reading, and English academic achievement tests. These outcomes are from the standardized PLAN test for 10th and 11th graders, which is a pre-ACT test. On a national norm sample, reliability for each of the subject tests was good (Math: $\alpha = .81$, Science: $\alpha = .83$, Reading: $\alpha = .80$, English: $\alpha = .86$). The items within each PLAN subject test are selected to reflect approximately half of the examinees answering the average item correctly.
(mean item difficulty ranges from .51 to .59 based on the subject), and item difficulty is
distributed from .20 to .89 in such a way as to differentiate between high and low ability
students. Additionally, each item must have a correlation of at least .20 with the total subject
scale score, and the scale is constructed to have a mean discrimination correlation of a medium
size (ranging from .51 to .56; ACT, Inc., 2007).

The scores students can receive on each PLAN subject test range from 1 to 32, and each
subject has its own benchmark score that indicates college readiness (Math = 19, Science = 21,
Reading = 17, English = 15). These benchmarks indicate the cut-off score at which students
have a 50% likelihood of getting a B or higher in a college course. Thus, the PLAN is a
psychometrically sound test, capable of capturing the diversity of student ability, and with
proven validity for student performance at specified levels. Within the current sample, the
average score for the PLAN tests was below the benchmark in all subjects except English (Math:

\[ M = 15.88, \text{SD} = 4.18; \text{Science: } M = 17.21, \text{SD} = 3.36; \text{Reading: } M = 15.91, \text{SD} = 4.50; \text{English: } M = 15.65, \text{SD} = 4.14. \]

**Student climate perceptions.** The CFL climate survey contains 55 items that allow for
an exploration of student climate perceptions of their teachers, fellow classmates, and the general
school environment. The response scales differ somewhat between items, depending on whether
the item asked for degree of agreement with a description or the frequency of a behavior.
However, all responses were recorded on a 4-point scale, with higher scores indicating greater
agreement or higher frequency. For certain items, responses have been reverse coded when
necessary such that a higher score always indicates a more positive climate. The CFL item
responses collected in this sample are identified such that the scale scores could be analyzed at
both the individual student and aggregate school-levels.
The CFL was originally designed to tap four school climate scales (School Safety, Student Support, High Expectations, and Social and Emotional Skills), and all of the items for each scale are included in tables in Appendix A. Each item in the tables is also marked to indicate the response scale used and whether it was reverse coded. However, subsequent analyses have indicated that this hypothetical factor structure does not hold (Rosenberg et al., 2010). The same conclusions were reached for the current sample; an initial confirmatory factor analysis (CFA) regressing the items on the theoretical factors showed somewhat poor model fit.

In an effort to construct more reliable factors from the climate data, an exploratory factor analysis was conducted using a random sample of half of the cases. This resulted in a reduced group of 36 items, which were hypothesized to load on 5 factors: School Safety, Antisocial Peers, Prosocial Peers, Teacher/Student Interaction, and Institutional Expectations. These factors are described below. A second CFA was conducted with the remaining random half of the sample to assess the fit of the newly identified factors, and the model fit was improved for all metrics, though still smaller than ideal on the CFI. The specific model results and fit statistics, as well as the complete items listed under their new factors and reliability for each of the new scales, are included in Appendix B. Throughout the remaining analyses, the mean score on each of these factors was used to represent students’ perceptions of school climate. For all climate variables, including Antisocial Peers, a larger value indicates a more favorable view of the climate, as all items evaluating negative climate remained reverse scored.

**School safety.** The 5-item safety subscale measures both a global sense of student safety at school, and their safety within specific contexts. Sample items include “I worry about crime and violence in school” and “How safe do you feel in your classes?” The Cronbach’s alpha for the scale is .86. The mean perceived school safety score in this sample was 2.87 (SD= 0.64).
Antisocial peers. This 11-item subscale asks students about their perceptions of the presence of antisocial behaviors, such as bullying and cheating, among their fellow classmates. The degree to which students perceive that most of their classmates display these behaviors tells us something about the perceived norms at the school, as well as the social climate created among peers. Sample items include “Students at this school are often teased or picked on” and “Most students in my school give up when they can't solve a problem easily.” The Cronbach’s alpha for the scale is .86, and the mean score in this sample was 2.36 (SD= 0.51).

Prosocial peers. This 7-item subscale asks students about their perceptions of the presence of prosocial and responsible behaviors, such as respect and hard-work, among their fellow classmates. This scale describes a climate where students are engaged in their schoolwork and take positive approaches to problem solving. Sample items include “Most students in my school do their best, even when their school work is difficult” and “Most students in my school stop and think before doing anything when they get angry.” The Cronbach’s alpha for the scale is .82, and the mean score in this sample was 2.36 (SD= 0.53).

Teacher/student interaction. This 9-item subscale assesses the degree to which students feel that their teachers support them and engage them in the classroom. Most of the items describe how teachers facilitate student learning. Sample items include “My teachers really care about me” and “My teachers often connect what I am learning to life outside the classroom.” The Cronbach’s alpha for the scale is .86.

Institutional expectations. The 4-item expectations subscale measures the degree to which students feel that their school expects them to complete challenging course loads. These items reflect students’ perceptions of an institutional norm at their school that encourages
students to maintain a rigorous academic schedule. A sample item is “Students in this school are expected to take four years of math.” The Cronbach’s alpha for the scale is .70.

**Gender-based stereotype threat.** One of the primary interests of the current study is to explore the difference between male and female students’ scores on the PLAN tests, ratings of climate using the CFI, and the interaction with CFI scores in predicting PLAN scores. Student gender was entered into the models as a dichotomous variable, coded as Men = 0 and Women = 1. Thus, any predictive relationships seen in the models indicate the change in women’s scores relative to the baseline of men’s scores. In accordance with stereotype threat theory and past research, it is expected that female students have a stereotype-threatened gender identity in math and science domains, but not in English and reading.

**Student level covariates.** In addition to the primary relationships of interest, several other student characteristics were included in the study as important covariates.

**Racial stereotype threat.** In addition to the gender-based stereotype threat described above, I also sought to account for students’ experiences of racial stereotype threat, and the possible interactions between race and gender in predicting student academic performance. Thus, I used student reported race to create a dichotomous variable that grouped together African American, Latino American, and Native American students together as academically threatened racial groups (coded as 1), and White and Asian American students as academically non-threatened racial groups (coded as 0). This variable is referred to as “race” throughout the rest of the analyses, and represents stereotype-threatened racial identity.

**Student grade point average.** Student GPA was reported at the end of the Fall 2006 semester, and reflects the cumulative performance across all subjects in the first semester. It is
measured on a 5-point scale, and the average student GPA in this sample is 2.53 ($SD= 0.98$). In the current study, student GPA is used as a proxy for a student’s general academic performance.

**Grade level.** All students included in this sample were either in the 10th or 11th grade cohorts. This dichotomous control variable was coded so that 10th grade = 0 and 11th grade = 1. Thus, any predictive relationships seen in the models indicate the change in scores for 11th grade students relative to the baseline of 10th grade students’ scores.

**Free lunch status (SES).** The final student covariate included in this model identifies a student as receiving free lunch (coded as 1) or not (coded as 0). This dichotomous variable is used as a proxy to represent low SES, as family income has to be close to the poverty line for students to qualify for the federal free lunch program.

**School-level Covariates**

In addition to the control variables included at the student level, several features of the 109 schools were also used as covariates in the school-level and multilevel models.

**School diversity.** The racial heterogeneity of each school was quantified using the Racial Diversity Index (RDI; Lieberson, 1969), which reflects the percentage of students within a school who do not share a common racial identity. It is calculated by squaring the proportion of each racial group within the school, and subtracting the sum of those squared proportions from one. Thus, the RDI will be large if there is a fairly equal distribution of several ethnic groups, and small if there is one dominant racial group. This calculation allows for the inclusion of a single, continuous covariate to represent the ethnic diversity within a school. As a whole, schools in this sample tended to be fairly racially homogenous ($M = 0.39; SD = 0.26$).
School size. The school size was operationalized as the number of students enrolled in the school as a whole. As school size values are typically large, and the variance within school population is substantial, size was rescaled by dividing all values by 100 before being used in any analyses (This was done to meet the recommendations for analyses in Mplus that all variances be less than 10.). Schools in this sample ranged in size from 67 to 4,248 students enrolled ($M = 969; SD = 792$).

Mobility rate. A school’s mobility rate is calculated based on the proportion of the student population that transfers into or out of the school over the course of a year. Thus, a higher mobility rate indicates less stability in the members of the school community. Schools in the current study varied considerably in their mobility rates; the lowest rates were 0.01, and the highest was 0.62 ($M = 0.23, SD = 0.15$).

School free lunch (SES). Finally, in addition to the individual measure of a student’s personal free lunch status, the percentage of students receiving free lunches at that school was reported and is used as a school-level covariate. This variable is a proxy for the general SES of the school community as a whole, as it shows the proportion of students that come from low SES households. The average percentage is high; in most schools, the majority of students are receiving free lunch ($M = 0.79, SD = 0.21$).

Procedures: Collection of Measures

During the Fall 2006 semester, student demographic information was recorded in official school records. The measures of student gender, race, free lunch status, and grade level was taken from these records. Additionally, GPA was taken from each student’s official school record at the end of the semester.
During the Spring 2007 semester, the CFL was administered as part of a larger confidential, voluntary survey administered by the school district and designed along with the Consortium on Chicago School Research. In addition to the CFL items, this survey had many other questions asking students about their families, friends, teachers, aspirations, and personal behaviors. Later in the semester, students took the PLAN tests. Shortly before the academic subject tests were administered (typically the day or morning before the test administration), students filled out their personal information and interests, which included reporting their gender and racial/ethnic background. This should have primed these identities prior to taking the exams, and therefore makes the experience of stereotype threat more likely.

All of the school-level characteristics (i.e., the RDI, school size, mobility rates, and free lunch proportion) were collected and compiled by Chicago Public Schools, and the information was released in databases and published in school report cards after the conclusion of the school year.

**Procedures: Data Analysis**

**Selecting participants.** In order to select students who had at least a minimum investment in filling out the survey, only those who completed 50% or more of the items in the CFL as a whole and within each subscale were included in the data analysis.

**Verifying climate factors.** In order to establish the factor structure that was hypothesized within the CFL, I conducted a confirmatory factor analysis to test whether the theoretical structure of four climate scales holds for the data. Due to the poor fit of this model, I conducted an exploratory factor analysis on a random half of the cases, and used those results and guiding theoretical principles to form 5 new scales of climate. A second confirmatory factor
analysis was conducted on the remaining cases to verify the new factor structure. All of these analyses were conducted using Mplus 6, and are discussed in further detail in Appendix B.

**Analysis plan.** As school climate data are inherently nested perceptions of individual students within a larger school culture, the most appropriate method for testing the relationship between climate and achievement at individual and school-levels (i.e., Hypothesis #1) is to use a multilevel modeling approach to capture the variation within and between schools in one analysis. I also looked within this nested model for significant interactions between gender and climate on achievement test scores (i.e., Hypothesis #2).

**Specifying the model.** The first step of constructing and testing these multilevel models was to specify the hypothesized relationships with academic test scores at the school and student levels. In order to avoid problems of multicollinearity in the estimation of regression coefficients (Kline, 2005), the bivariate correlations were examined for any values over $r = .85$ (Table 1). There was a significant multicollinearity problem among the school-level aggregated PLAN scores ($r = .96-.98$), and therefore separate models had to be constructed to predict each of the four subject tests (PLAN_M=math; PLAN_S=science; PLAN_R=reading; PLAN_E=English). Additionally, an exploration of climate scales at both school and student levels revealed that, while the bivariate correlations were not unacceptably large, entering the scales together in one block resulted in suppression effects, resulting in significantly reduced or even reversed correlations with the PLAN scores for several climate scales. Therefore, separate models were conducted for each of the climate scales on each of the PLAN tests.

**Centering the variables.** One of the important decisions when conducting multilevel modeling is which of three metrics to use for the student level variables: raw scores, scores centered at the grand mean (CGM), or scores centered within each cluster (CWC). The type of
centering used in a model changes the interpretation of the parameter estimates (due to the change in the zero point of the predictor), but it can also result in an upward bias of school-level effects if individual scores are CGM and the research question looks across levels (Enders & Tofighi, 2007). Thus, for the following analyses, continuous student predictors have been centered within their respective schools (CWC), while school-level predictors have been centered around the grand school mean (CGM). Dichotomous variables in this dataset all have a meaningful zero value in their raw score metric, and thus the choice was made not to center them.

Model estimation. Each 2-level model of students nested within schools was estimated using Mplus 6, following the procedures for multilevel latent covariate (MLC) modeling (Lüdtke, Marsh, Robitzsch, Trautwein, Asparouhov, & Muthén, 2008). This method is a blend of approaches typically seen in hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) and multilevel structural equation modeling (MSEM; Preacher, Zyphur, & Zang, 2010). Like HLM, the MLC approach is able to examine variance at both student and school-levels simultaneously. This approach accounts for the similarity of students within their school-level context, and thus gives a less biased estimation of the impact of both individual-level and school-level regressions. However, when school-level variables are derived from the scores of students within the school (e.g., school-level climate is created from student level perceptions), HLM can be biased by its reliance on aggregated group means. These aggregate group scores assume perfect measurement of the variable at the school-level, which is problematic (Ludtke et al., 2008; Preacher et al., 2010). Using MLC, it is possible to model these group effects as latent variables that are derived from individual-level reports. Thus, MLC assumes that there will be
some error in this group level measurement, and will provide a more accurate estimation of the school-level relationships with a variable like school-level climate.

**Cross-level interactions.** Additionally, MLC allows for the regression of random slopes on school-level variables, so that school characteristics that predict the variation in relationship across schools can be modeled. This modeling of random slopes allows for the testing of cross-level interactions, such as the hypothesized impact of school-level climate on the student-level relationship between gender and PLAN scores (e.g., Hypothesis #2).

**Alternative models.** Finally, the model fit statistics that are provided by an MLC approach can be used to test alternative models that impose additional constraints on the relationships (e.g., setting a path coefficient equal to zero), and thus determine if the freed model is a significantly better fit of the data. For each model, the log likelihood value (a fit statistic) is multiplied by -2; this -2*log likelihood (-2LL) for the model with freed parameters is subtracted from the constrained model. The resulting difference in -2LL is distributed as a chi-square, and the difference in the degrees of freedom can be used to determine the significance. A significant difference in model fit using this test implies that the model which estimates additional parameters is a better fit for the data. If there is no significant difference in fit, then the parameters in question do not need to be estimated, and it is better to use the more constrained model.

The current study examines these hypothesized and alternative models using a series of analyses applying the same structural model but varying the key predictor and outcome. The MLC estimations were conducted for a total of 20 models, looking separately at the effects of each of the 5 climate scales on each of the 4 PLAN test scores. Due to the large number of models included in these analyses, the level of significance for all analyses was raised to $\alpha = .01,$
which will substantially reduce the likelihood of a Type 1 error. Each model estimates three random slopes at the student-level, which are modeled as latent variables at the school-level. PLAN test scores are also represented as latent variables in the MLC models. Thus, in MLC, all of the hypothetical relationships and interactions in the current study are tested within the same model.

RESULTS

These analyses are designed to explore the impact of climate perceptions on student achievement test scores, the relationship of gender to test performance, and the possible difference in climate’s impact for male and female students. These relationships are explored in a multilevel covariate (MLC) model that examines (a) student-level relationships while controlling for the school context, (b) school-level relationships while controlling for the student variation within each school, and (c) both within-level interactions and cross-level interactions. This model also incorporates covariates at the student-level (i.e., race, GPA, grade level, free lunch status) and the school-level (i.e., RDI, size, mobility rate, free lunch rate).

Before examining this complex model, the basic correlations between the variables included in the model are discussed, as well as the average level of climate perceptions and academic achievement in this sample. Then, the covariates, interactions, and relationships are described separately for the student-level and school-level parts of the model. Next, all the relevant variables and relationships are put into the MLC model, and the main effects of climate and gender on academic achievement scores are examined, as well as their within-level and cross-level interactions. Finally several alternative multilevel models are presented.
Correlations and Mean Differences

Due to the very large sample size (N= 27,203) examined in this database, it is especially important to consider magnitude as opposed to merely significance when looking at these correlations and all the following student-level analyses. Even a very small correlation may be significant in a sample this large, so only those of a substantial size are discussed below (although all bivariate correlations are presented in Table 1). Looking at the correlations, a few interesting patterns can be seen in the main relationships of interest. First, the strongest relationship for each of the four PLAN test score outcomes (aside from their relationships with each other) is with GPA (r from .51 to .56). Additionally, all of the subject test scores are lower for students who receive free lunch (r from -.31 to -.34) and those who belong to a stereotyped racial group (r from -.31 to -.39).

When looking at the bivariate relationships between climate measures and each of the PLAN tests, a consistent positive moderate relationship is seen for perceptions of Safety (r from .25 to .26). Small positive relationships are seen for perceptions of fewer Antisocial Peers (r from .15 to .20) and higher Expectations (r from .16 to .17). The relationships between Prosocial Peers and Teacher/Student perceptions and the PLAN tests, however, are all very small (r < .09).

Gender and race mean differences. Without controlling for any other covariates or nesting within schools, the simple mean differences between male and female students were significant on all of the PLAN tests (≤ 1.0), with females scoring slightly higher on all subjects but Math (see Table 2). The gender differences in climate scale scores were also significant, but very small (≤ 0.11). Students whose racial identity was stereotyped regarding academic achievement, however, showed a larger achievement gap in the simple mean differences. Black, Latino, and Native American students scored around 3.5 points lower than their non-stereotyped
peers on all PLAN tests. Racial identity was not related to any large differences in climate scale scores (all gaps ≤ 0.25), though all differences favored White and Asian students.

**Modeling Student-Level Relationships with Covariates**

The construction of an MLC model in Mplus involves a complicated specification of paths both within and across levels. Thus, relationships within the student-level were first specified to determine which student characteristics might significantly predict unique variation in each of the four academic outcomes. Simple linear regressions were run predicting each of the four PLAN subject tests from gender and all of the student-level covariates. The results are summarized in Table 3. When controlling for student GPA and other covariates, the expected pattern of gender differences in test performance is seen, with female gender predicting lower scores on the Math ($B = -1.00$) and Science exams ($B = -0.39$) and higher on Reading ($B = 0.30$) and English ($B = 0.20$). Additionally, all of the student level covariates predicted significant unique variance on each of the PLAN tests. They were all retained for use in the final models. The sample model shown in Figure 1a illustrates how these predictors were used at the student level, including the potential within-level interactions between gender and race and between gender and CWC climate perceptions. Additionally, the slopes predicting PLAN scores from gender (S1), CWC climate (S2), and race (S3) are labeled.

**Modeling School-Level Relationships with Covariates**

**Between school variance.** Before exploring the relationships at the school-level, it is necessary to verify that there is sufficient variation to be modeled between schools in both the climate factors and the PLAN test scores. Thus, the design effect for each model was calculated using the intra-cluster correlation (ICC) using the mixed models analysis in SPSS 18. The ICC reflects the percentage of total variance in a measure that is accounted for by the cluster; in this
case, it reflects intra-school correlations. The design effect is a measure of the cluster level variance based on the ICC that has been adjusted for cluster size, as larger cluster sizes indicate that even small ICC values should be examined (Muthén & Satorra, 1995). The ICC and design effect for each of the five climate factors and four PLAN tests are listed in Table 4. All of the design effects are well above 2, indicating the need to model the variable at the school-level, and the ICCs show that the percentage of variance accounted for at the school-level ranges from 7-19% for climate factors and 30-35% for PLAN variables. These values are substantial for school-level variance, and indicate that examining school climate measures or academic achievement tests without accounting for the nesting of students within schools may yield biased results.

**Significant predictors of between-school variance.** As with student level predictors, school-level predictors were specified to identify those with an impact on between-school PLAN scores. A separate simple regression was run to predict the mean school-level aggregate score on each of the four academic achievement tests, and significant school-level predictors are indicated in Table 5. Strong negative relationships are consistently seen between PLAN scores and both the mobility rate and the percentage of free lunch students. The school population size and its racial diversity have generally positive but smaller relationships. All the covariates predicted significant unique variation in all the aggregate academic outcomes and were retained for use in the MLC models. The sample model shown in Figure 1b illustrates how the school-level variables were used to predict the latent intercepts of PLAN scores, and their hypothesized role in cross-level interactions through their impact on the random slopes (S1-S3) of within-level relationships.
Multilevel Latent Covariate Model

Combining the separate levels depicted in Figure 1, each of the four PLAN test outcomes was modeled using each of the five school climate domains. This resulted in 20 separate MLC models being estimated, using the full information maximum likelihood (FIML) estimator in Mplus 6. The results of these models are depicted in Figures 2-5. Following the conventions established for models run in Mplus (Muthen & Muthen, 2010), any intercepts that were hypothesized to vary among schools and all slopes that were allowed to randomly vary according to between school factors (i.e., cross-level interactions) are marked with a solid dot. Each model in these analyses estimates one random intercept (i.e., the average PLAN score) and three random slopes (i.e., the regression of the PLAN score on individual climate perceptions, gender, and racial identity). These random components are then modeled as latent variables at the school-level, as indicated by their representation as open circles. Each PLAN test appears in a separate figure, and the pattern of results across models will be discussed below.

For each of the PLAN tests, all five climate models are portrayed within a single model diagram.¹ For the relationship between the individual-level covariates (e.g., grade level) and PLAN scores, the parameter values were very similar across all five climate models within each PLAN test; the main effects for these variables did not shift much depending on which climate measure was included in the model. Thus, for those parameters, a single value is labeled that represents all climate models. For example, in the regression of PLAN Math score on grade level (Figure 2), the parameter estimates from each of the climate models (Safety= 0.65, Antisocial= 0.66, Prosocial= 0.67, Teacher= 0.66, Expectations=0.67) were averaged into the

¹ In the model diagrams and in the summary results presented below, only the paths predicting the PLAN scores are shown at the individual level. However, paths were also calculated from the student level control variables to the climate measure in each model. For clarity of viewing and interpreting results, these parameter estimates will not depicted or discussed here, but can be found in all model outputs in Appendix C.
estimate of 0.66 that is depicted on the path from Grade to PLAN_M in Figure 2. This was done to allow for an easy visual interpretation of the parameters that did not differ substantially by climate domain model. In the summary of results below, these average results are indicated as approximate (=) parameter estimates.

However, other parameter estimates differed depending on the climate domain used in the model, and in those cases the results for each separate climate model are described in a table or figure. The relevant table or figure is indicated in Figures 2-5 instead of a single parameter estimate, and the relationships are also discussed in more detail throughout the text. Most of these varying relationships are at the school-level, which indicates that variations in the climate domain being measured have more of an impact on the relationships between covariates and academic achievement at the school-level than at the student-level.

**Interpreting parameter estimates.** Due to the inclusion of random slopes within the multilevel latent covariance models, it was not possible for Mplus to calculate standardized values for the parameter estimates. Thus, to aid in the interpretation of the magnitude of the relationships, any continuous variables were school-mean centered at the student level, and grand-mean centered at the school-level. At the student level, parameter estimates for continuous variables indicate the increase (or decrease) in an individual PLAN score for students at one standard deviation above the mean for their school on the predictor variable. Looking at Figure 2, the parameter estimate for GPA to PLAN_M is 1.59. This indicates that students who were one standard deviation in GPA above their schoolmates scored 1.59 points higher on the PLAN Math test than those with an average GPA. Thus, the score for students one SD above the mean on the PLAN_M, holding all other covariates constant, is the intercept (17.02) plus 1.59, or 18.61. For the dichotomous predictors, the parameter estimate indicates the change from the
intercept in PLAN scores for students who are scored as a 1 on the variable compared to those who are scored as a 0. Using our previous example from Figure 2, the parameter value of 0.66 for Grade to PLAN_M indicates that, controlling for all the other covariates, students in the 11th grade scored an average of 17.62 (i.e., 0.66 + 17.02) on the PLAN Math exam, whereas 10th graders scored an average of 17.02.

At the school-level, all variables are continuous, and parameter estimates predicting PLAN scores indicate the increase (or decrease) for a school that is one standard deviation above the mean for all schools on the predictor variable. [Parameter estimates predicting the random slopes (S1-S3) describe cross-level interactions, and their interpretation is discussed separately.] Using scores on the PLAN Math exam in Figure 2 as our example once again, this would mean that to find out how much the intercept of PLAN_M decreases for schools one SD above the mean in the percentage of Free Lunch students, we would look at the values in Table 6 within each climate model (e.g., in the Safety model, a decrease of -6.25 would lead to an intercept of 10.77). Paths that have non-significant results across all climate domain models are indicated with an NS along the parameter instead of an estimate or a Table reference.

**Covariates.** The relationships between the student and school covariates and the PLAN scores within the MLC models are presented below.

**Student level.** At the student level, first semester GPA was used as a proxy to control for students’ general academic performance in predicting PLAN scores. As expected, GPA was significantly positively related to Math ($B \approx 1.59$), Science ($B \approx 1.23$), Reading ($B \approx 1.47$), and English ($B \approx 1.45$) scores. Grade level was also positively related to all achievement test scores, with an approximate gain of 0.6 points from 10th to 11th grade [Math ($B \approx 0.66$), Science ($B \approx 0.53$), Reading ($B \approx 0.58$), & English ($B \approx 0.72$)]. Additionally, small but consistent negative
relationships between academic outcomes and free lunch status were observed [Math ($B \approx -0.29$), Science ($B \approx -0.31$), Reading ($B \approx -0.69$), & English ($B \approx -0.72$)].

**School-level.** At the school-level, the strongest predictor of a school’s aggregate performance across all the PLAN subject tests and climate models is its percentage of free lunch students ($B$ from -5.64 to -9.78; see Table 6). The higher proportion of free lunch students a school had compared to the other schools in the sample, the worse student scores were on all of the subject tests. A school’s mobility rate was also significantly negatively related to its average PLAN performance in all but 3 of the models; when a sense of Safety was included as the climate variable, mobility did not significantly explain any unique variance in math, science, or reading scores. School population size also had a significant but very small positive relationship to PLAN scores in 17 of the 20 models; after controlling for the other variables, the larger schools in this sample had just slightly higher test scores. However, the positive relationship that was initially observed between the RDI and PLAN scores was not significant in any of the full MLC models. Once school-level climate and individual-level variance on race and other variables was added to the model, racial diversity no longer contributed to a school’s academic performance.

**Main effects of climate impact.** The first hypothesis for the current study predicted a positive relationship between climate domains and academic outcomes at both the student and school-level. These relationships are explored in the context of the full MLC model below.

**Student level.** A positive main effect for perceived climate at the individual-level was predicted for all climate dimensions on all of the PLAN outcomes, but this hypothesis was not supported by the data. The main effects for student level climate are presented in Table 7. Contrary to the general hypothesis that individual-level climate would be consistently positively
related to student academic achievement, results were mixed according to the climate domain under assessment. Student level ratings of Teacher had no significant parameter estimates. Ratings of Safety ($B$ from 0.16 to 0.29) and Expectations ($B$ from .16 to .22; Math NS) were generally positively related to PLAN outcomes. Ratings of less Antisocial Peers ($B$ from -0.27 to -0.44) and more Prosocial Peers ($B$ from -0.44 to -0.73) were negatively related to PLAN outcomes.

**School-level.** A positive main effect for perceived climate was also predicted at the school-level for all climate dimensions on all PLAN outcomes. This hypothesis was largely supported by the data, as shown in Table 8. The magnitude of the effect varied, but all of the regressions of PLAN scores on the climate domains were positive ($B$’s ranged from 2.40 to 7.64); 16 of the relationships were significant, and the remaining four were marginal (using a conservative alpha of .01; all were $p<.05$). Thus, after controlling for important school features and individual student variation, schools with higher means in all of the climate domains tended to have higher average PLAN scores on all tests.

It should be noted that, unlike covariates, the interpretation of the parameter estimates for climate domains predicting PLAN scores are not additive because each of the climate domains was modeled independently. As an illustrative example, for the PLAN Math test, a school that is one SD above the mean in its rating of Safety would have a PLAN_M score of $4.97 + 17.02 = 21.99$ (i.e., slope from Table 8 plus Intercept from Figure 2). A school one SD above the mean in its rating of Prosocial behavior would have a PLAN_M score of $22.85$ (i.e., $5.83 + 17.02$). However, these parameter estimates cannot be combined to get the cumulative effect of being one SD above the mean on both measures of climate.
Main effects of stereotype threat. The following analyses explore the potential existence of student-level relationships between student demographics and test performance that would be consistent with stereotype threat theory. In order to properly observe the potential existence of stereotype threat, one must control for general ability or academic performance in low pressure situations when looking at performance on high-pressure, evaluative exams. Thus, these effects may differ from the simple mean differences described earlier in Table 2, because what is being examined here is the performance of different genders and races relative to what one would expect given their GPA, and after controlling for other covariates and nesting in different schools.

Gender. As expected, a pattern consistent with gender-based stereotype threat was seen in the main effect of gender across all of the models. After controlling for GPA (and other individual-level covariates), female students on average performed lower than male students on the PLAN Math ($B \approx -1.22$) and the PLAN Science ($B \approx -0.63$). They also performed slightly higher than males on the PLAN Reading ($B \approx 0.25$) and PLAN English ($B \approx 0.37$). Thus, females were underperforming relative to their GPA on math and science tests, and, to a lesser degree, males were underperforming in Reading and English.

Racial identity. The expected pattern of results for stereotyped racial identities was only partially reflected in the main effects of racial identity on PLAN scores. Once controlling for the covariates and nesting, there was no significant difference in the scores for Black and Latino\(^2\) students relative to White and Asian students in English and Reading exams. However, the predicted PLAN scores were lower for Black and Latino students in Math ($B \approx -1.00$) and Science ($B \approx -0.73$) than for White and Asian students. Thus the hypothesis that students with a

\(^2\) Native American students were also included in stereotyped group, but make up a very small portion of the population (<1%).
stereotyped racial identity would underperform relative to similar peers on all achievement tests was not supported, but they did underperform in Math and Science domains.

**Patterns of within-level student interactions.** The presence of a significant interaction between gender and climate would indicate that the relationship between a specific climate perception and a specific academic test differs for male and female students. The hypothesized reduction in stereotype threat effects for female students whose climate perceptions are positive would be seen in significant interactions for the Math and Science tests, with simple slopes relating climate to test scores that were stronger for female students than male students. Additionally, interactions might be expected for the Reading and English tests as well, but the simple slopes in those domains should instead show a stronger relationship for male students to support the hypothesis of reductions in stereotype threat through positive climate. However, neither of these hypothesized patterns of results were observed across the various models at the individual student level. The results that were observed are described below.

**Climate and gender.** A significant interaction between climate and gender was observed for four of the 20 models, but there was no consistent pattern within those results. The simple slopes for the four significant interactions that were observed are portrayed in Figure 6a-d. As seen in Figures 6a and 6b, perceptions of a less antisocial peer climate did decrease the gender gap in performance on Math and Science tests, but seemed to do so primarily through a negative relationship with male students’ scores. Figure 6c, which shows a sharper slope of the female regression line compared to the male regression line, indicates that Safety interacted with gender to increase the advantage for female students on Reading tests. The same interaction effect for Safety and gender can be seen for English tests in Figure 6d.
Race and gender. In order to explore the possible multiplicative effects for students who have both stereotyped gender and racial identities, I also tested for interactions between gender and race in their impact on PLAN scores. These interactions show a much clearer pattern of results. There are no significant race-by-gender interactions predicting reading or English scores. However, for science and math PLAN tests, there are significant positive interactions in every climate model, indicating that students with stereotyped racial identities experienced less of a gender gap in performance. The average simple slopes illustrating these interactive relationships are illustrated in Figures 7 and 8. These graphs show that Black and Latino students of both genders are underperforming on Math and Science exams relative to their White and Asian peers. However, this slope is more severe for male students than female students. Thus, students with a dual stereotyped identity (i.e., both female and Black/Latino in Math or Science domains) do perform the worst among all groups, but the relationship is not a simple additive disadvantage. Instead, it seems that female students, who would already theoretically be under a gender-based threat in these domains, are less impacted by the additional negative effects associated with racial identity than their male peers.

Patterns of cross-level interactions. Three types of cross level interactions were explored in the MLC models. First, a significant cross-level interaction between school-level climate and student gender provides an alternative test of the hypothesis of reductions of stereotype threat in positive climates. Instead of looking at the impact of an individual’s perception of the climate on their test scores, the cross-level interaction tests whether the relationship between gender and test scores within a given school changes based on its collective climate. (e.g., Does a school with a higher average Expectations score have a less negative relationship between gender and PLAN_M scores?) Again, the hypothesis here was that female
students in schools with more positive psychosocial climates would be protected from the anxiety of stereotype threat. Thus significant interactions were expected for Math and Science domains, with simple slopes showing a stronger relationship between climate and achievement for female students than male students.

Second, the impact of the collective school climate on the relationship between individual climate and test scores was examined (e.g., Is the relationship between individual perceptions of Expectations and PLAN_M scores stronger in a school with a higher average Expectations score?). Finally, the relationship between race and test scores was tested for interactions with collective school climate and with the racial diversity of the school. (e.g., Does a school with a higher average Expectations score have a less negative relationship between race and PLAN_M scores?; Is the relationship between race and PLAN_M scores weaker in a school with a higher average RDI score?)

There were no significant interactions between student race and school racial diversity. For the interactions between student race and school-level climate, only three significant differences were observed (in the models of Expectations on Math and English, and Teacher/Student Relationships on English). For the interactions between individual-level climate and school-level climate, significant differences were observed in just over half of the models (11/20), but none of the climate measures or PLAN outcomes were consistently part of these interactions. The simple slopes for these preceding interactions are not represented in the figures or described in the text because of the lack of a consistent pattern to the results. Only the first category of cross-level interactions (i.e., school-level climate and gender) showed a clear

---

3 The specific values for each of the interactions can be found in Appendix C under the Between Level model in the regression of S2 on each CGM climate score.
pattern of significant results, and it is discussed in detail in the following section and represented graphically.

**Aggregated school climate and student gender.** While there was no clear pattern of results for student-level interaction between climate and gender, the interactions between school-level climate and gender were more interpretable. For Reading and English tests, no cross-level interactions were observed. For the models predicting Science and Math scores, however, there were significant cross-level interactions for all climate domains except Teacher-Student interactions. Thus, for all the other domains, the collective climate in a school influences the relationship between student gender and Math and Science tests in that school. Math and science are the subjects in which I hypothesized significant interactions between student gender and collective climate based on stereotype threat theory. However, the simple slopes of these interactions (Figures 9 & 10) reveal that psychosocial climate does not operate in the hypothesized manner to reduce stereotype threat. Instead, at high levels of beneficial climate, the gender gap between males and females in math and science increased across the models. Students of both genders performed better in the schools with stronger climates, but male students benefitted more than female students. The exception to this pattern is found in the climate domain of Teacher-Student interactions; the gender gap is consistent at all levels of this climate measure. Thus, girls and boys benefit equally from schools where teachers are more caring, dedicated, and encouraging.

**Alternative Models**

In addition to the hypothesized models, a set of alternative models was tested in comparison. The model structure is nested within the hypothesized models, and thus can be directly compared for the fit impact of constraining additional parameters. For each of these
models, if the fit is not significantly improved by allowing certain parameters to vary, the more constrained model is preferred because it is more parsimonious.

**Random slope of gender.** Within the hypothesized models tested above, the impact of gender on test scores was allowed to vary between schools through a random slope. As an alternative model, this slope was constrained to be equal for all schools. Thus, this alternative model tests if allowing the relationship between gender and test scores to be a random slope added anything to model fit. If there is a significant fit improvement in the model with a random slope, it indicates that some schools do have greater gender gaps than others, when controlling for the included covariates.

The findings for this set of alternative models are divided by subject. For Math and Science tests, the $-2LL$ difference of fit tests were significant in all models (range from 15.01 to 53.38, all $p<.01$). Thus, the degree to which women underperform relative to men in Math and Science tests varies according to the school in which they are nested. However, for English and Reading tests, allowing the relationship between gender and PLAN scores to vary did not significantly improve model fit in any of the models ($-2LL$ range from 0.09 to 1.83, all $p>.05$). This shows that schools do vary in the impact of student gender on Math and Science testing, but not on English and Reading. Thus, some features of the school may be expected to impact the gender gap in Math and Science, but any gaps that exist in English and Reading are not affected by school-level differences in this sample.

**Summary**

Climate in this study was conceptualized and modeled as both an individual-level perception and a school-level characteristic, and the impact of climate on achievement was thus tested at both the student and school-levels. The hypothesis that a variety of school climate
domains would positively impact student achievement was confirmed at the collective school-level. At the individual-level, however, a mix of positive, negative, and null effects was seen across the different domains and different academic subjects. The second phenomenon explored in this study was the potential of school climate perceptions to impact the stereotype threat effect experienced by female students in math and science classes. The hypothesis that climate would have a stronger relationship with academic performance for females than males in these domains was not supported by these analyses. At the individual-level, only a few climate by gender interactions were significant, and the pattern of results was inconsistent. At the school-level, an interaction effect partially opposed to the hypothesis was observed: while both genders benefitted from positive climates, the gap between male and female students in math and science was largest in the schools with the highest scores on most climate dimensions. These results and their implications are discussed in more detail below.

**DISCUSSION**

The current study was designed to analyze an existing measure of school-level climate, and to evaluate the impact of psychosocial climate on the academic achievement test scores of 10th and 11th grade students. This analysis was complicated by some considerations regarding the nature of school climate and its measurement, which are discussed below. After forming a refined set of climate dimensions, my analyses revealed a strong relationship at the school level between all of the climate measurements and all subjects of the academic achievement tests. The details and implications of these findings are discussed below. At the student level, however, climate had a more complex and occasionally counterintuitive relationship to achievement, which raises questions about the nature of student’s individual perceptions. Finally, the hypothesized relationship between more positive climate perceptions and smaller achievement
gaps was not confirmed by the data. Possible future directions in researching these questions and alternative explanations are also presented.

**The Nature of Climate and Measuring Perceptions**

One of the major conclusions of this study is that the conceptualization and measurement of climate domains is incredibly complex. The original scales designed in the Conditions for Learning survey (Osher & Kendziora, 2010) did not hold. A subsequent exploratory factor analysis of the survey items confirmed the multidimensional nature of school climate: a model with only one central factor underlying the questions about climate was an extremely poor fit for the data, both with the full survey and the reduced set of items that was used in the subsequent analyses. A new factor structure was specified that had a more acceptable model fit, though it remained slightly less than ideal for the large sample size used here. This complexity in specifying climate domains affirms some important points for consideration in the measurement of school climate.

First, it is important that all items assessing a domain share a clear referent group, as that allows for clarity in the interpretation of the scale score (Chan, 1998). Empirically, this distinction is reflected in the factors that were drawn from the data, which were generally divided by reference to the self (Physical Safety), peers (Antisocial and Prosocial), teachers throughout the school (Teacher/Student Interactions), and the general school norms (Institutional Expectations). Furthermore, conceptually related items that assessed just one specific teacher instead of teachers throughout the school did not load on the Teacher/Student Interactions factor (see Appendix B). The use of a consistent referent is not only important for establishing the reliability of a scale, but it can also be important for deciding which level is appropriate for assessing the impact of the factor (Sinclair & Fraser, 2002; Taylor, 2011). In these analyses, the
Teacher/Student Interactions rating did not predict any individual level variation in achievement, but it did predict school level variance. One reason for this discrepancy might be that these questions did not typically ask individual students about their experiences with teachers, but rather referenced the level of interaction in the school in general. Safety assessments, on the other hand, asked about students’ own feelings of safety (not the feelings of their fellow classmates), which may contribute to their greater predictive power at the individual level. The choice of a group or individual referent has been shown to impact the level of agreement and the predictive power of climate measurements (e.g., Hoy, Smith, & Sweetland, 2002; Klein, Conn, Smith, & Sorra, 2001). Thus, the validity of using assessments that inquire about school-level trends in the climate to assess individual relationships is in question and requires further study.

A second methodological concern that was reflected in the factor analyses of this scale is the necessity to distinguish between the presence of positive factors and the absence of negative ones. The original conceptualization of SEL Skills collapsed both negative and positive assessments of student skills, but these items clearly loaded on different factors (Antisocial and Prosocial Peers) in the exploratory analysis reported in the current paper. A similar problem with the combination of positively and negatively worded items in a single scale has been demonstrated in other areas, such as the measurement of self-esteem (Dunbar, Hunt, Ford, & Der, 2000) and optimism/pessimism (Schwiezer & Rauch, 2008). More work needs to be done to explore whether the distinction between these factors is conceptually meaningful, or merely an artifact of question wording that can be controlled for by constructing a trait-method model (Höfling, Moosbrugger, Schermelleh-Engel, & Heidenreich, 2011).
The Relationship of Climate and Academic Achievement

The second major conclusion to be drawn from this study is that, at the school-level, the collective climate can be a powerful predictor of student performance. The relative strength of different climate domains among schools in these analyses was a consistent predictor of between-school variance in academic performance. Controlling for other key demographic factors and for within-school variation in responses, each climate domain predicted unique variance in schools’ performance on standardized achievement exams in Math, Science, English, and Reading. This conclusion supports the notion of the psychosocial climate within schools as a collective, contextual experience (Moos, 1979; Tagiuri, 1968). It also suggests that future studies of the impact of psychosocial climate perceptions on student achievement should, whenever possible, attempt to account for the nested nature of the data within settings, as climate appears to be quite powerful and consistent in its predictions of academic achievement at the aggregate level.

Furthermore, not only were the school-level results consistent, unique predictors, they were also uniformly positive. Even when controlling for such factors as a school’s poverty level and mobility rates, positive school climates were associated with substantial gains in school-level academic achievement. This was true across all of the subject tests, with the strongest relationships in Math. For a school that was at approximately the 68th percentile in Prosocial Peer Behavior (i.e., One standard deviation above the mean in a normal distribution), the predicted average Math score would be 22.85; for a school at the 50th percentile (i.e. average), that predicted score would only be 17.02. In practical terms, this means that the average student in the school with an above average climate would exceed the PLAN Math benchmark college readiness score (19), whereas the average student from the school with the average climate
would not. Thus, climate is not only an important variable for researchers in our ability to predict variance, but it is also a promising pathway for educators to advance student learning and potential for success.

The findings for the relationship of climate at the individual-level are less consistent, though some positive relationships did emerge in support of the hypothesis. Relative to the other students in their school, the more safe a student feels, the better they perform on academic achievement exams in all subject areas. This finding is in the expected direction of positive individual climate perceptions leading to greater academic success. It supports prior research showing that individual student’s feelings of safety impact their well-being in many areas, including academic motivation and risky behaviors (National Research Council, 1993; Brand et al., 2003). This reinforces the conceptualization of a sense of safety and security as a basic psychological need (Maslow, 1943), as the lack of it prevents students from achieving a sense of belonging in school and the academic motivation to succeed. A high level of institutional expectations was also positively related to students’ Science, Reading, and English test scores in this sample. This indicates that students who experience a greater academic press toward taking more challenging courses within their schools outperform their schoolmates on these subjects, which supports prior research findings regarding the importance of academic expectations and the pressure to achieve within schools (e.g., Lee & Smith, 1999; Marzano, 2000). It also aligns with findings on formal and informal tracking within schools—students are aware of the differential expectations that their school has for them as individuals within different tracks (Yonezawa & Jones, 2006), and tracking is one of the major mediators between students’ individual backgrounds and aspirations and their academic achievement (Lee & Bryk, 1998).
The consistently negative findings for the relationship between students’ ratings of their peers behaviors and their academic performance, however, is counterintuitive. Social-emotional learning and positive youth development theories would predict that the more prosocial the peer environment is, the more that facilitates positive development in the child. This theoretical relationship was supported when student perceptions were aggregated at the school-level. However, at the student-level, the effect of perceptions of peers does not uphold this theory. One possible interpretation of these contradictory findings is that, at the individual level, these questions about the behaviors of “most students” do not reflect the climate of the school so much as a student’s perception of their social group. Student social group membership is a reliable indicator of student grades over time (Wentzel & Caldwell, 1997), indicating that student social networks tend to include similarly achieving students. If we assume that a student’s ratings of “most students in my school” refers primarily to their social peer group, then low achieving individuals are primarily rating other low achieving students. A common effect in the group perception literature is that groups who are seen as less competent are rated as more warm, and vice versa (Fiske, Cuddy, Glick, & Xu, 2002). Thus, it is possible that being a low achieving student leads one to rate one’s similarly low competence peers as more warm, and that the opposite effect holds for high achieving students.

**Possible Stereotype Threat in an Urban School District.**

In contrast to the complex pattern of results between individual climate perceptions and academic achievement, a clearer pattern consistent with the predictions of stereotype threat theory emerged when looking at the relationship between scores on evaluative academic achievement tests and gender identity. As the theory would predict, female students underperformed in Math and Science domains, where a stereotype of poor performance is held
for their group. To a lesser degree, male students also underperformed in English and Reading exams. These results were apparent after controlling for prior grades, SES, and other student characteristics that might impact academic achievement. This pattern of results is consistent with the operation of gender stereotype threat in a naturalistic school setting. Very few studies have demonstrated academic performance patterns consistent with stereotype threat outside of a lab or manipulated conditions. This study thus offers evidence that normal school testing may lead to underperformance in stereotyped groups relative to their actual ability.

In future research, a more thorough examination of the phenomenon might control for grades within the specific subject corresponding to the relevant test subject, or perhaps compare the results of a practice subject test with a directed manipulation of stereotype threat to the official high stakes, stereotype threat inducing academic achievement exams. This would more clearly demonstrate that students facing stereotype threat are underperforming relative to their ability in a specific domain. Also, although a stereotyped racial identity did not predict the deficits across all subjects that were expected, a partial confirmation was seen in that, relative to their White and Asian peers, Black and Latino students also underperformed in Math and Science. It is possible that other distinctions might be found in English and Reading if races were examined individually, and future research should do more to examine these distinctions.

**Can Climate Impact Stereotype Threat?**

Unlike some previous research showing the impact of contextual factors on stereotype threat (Good et al., 2003; Murphy, Steele, & Gross, 2007) the individual perception of climate dimensions measured in the current study did not consistently impact stereotype threat. It could be that the more diffuse contextual climate aspects measured here are not targeted enough to impact the specific concerns underlying the effects of stereotype threat at the individual-level. A
general feeling of safety or a positive feeling about one’s teachers may not be as powerful in reducing anxiety regarding stereotype confirmation as the focused promotion of a malleable view of intelligence or the presence of examples of counter-stereotypical achievement in the context. Furthermore, it is likely that in order to influence achievement at the individual level within specific subjects, it is necessary to measure the climate of the specific classroom in which that subject is taught. Previous research has demonstrated that features of individual classrooms influence students’ perception of the general school climate (Koth, Bradshaw, & Leaf, 2008). This may be especially true for climate perceptions that focus on the teacher or fellow students, which shift between classes in the high-school setting. Thus, future studies that wish to examine the impact of stereotype threat are advised to measure climate within specific subject classrooms. Those immediate contexts may be most important in predicting students’ sense of threat or security, especially considering that gender stereotypes are subject specific.

At the school-level, the measures of school climate did impact the relationship between gender and school achievement, but the interaction did not operate in direction than was predicted. The stronger a school’s climate was in all of the measured dimensions other than Teacher/Student interactions, the larger the gender gap favoring boys was in Math and Science achievement test scores. There was no impact of any climate dimension on the gap favoring girls in English and Reading tests. This finding was unexpected, but is consistent with previous research by Lee and colleagues (Lee, Chen, & Smerdon, 1996) that showed that while positive school climates increased the effectiveness of schools, they did not promote equity. Thus, while a while a rising school climate lifts all test scores, it may also maintain or increase disparities.

However, despite results showing that these climate dimensions do not substantially alleviate the pressures of stereotype threat, the alternative models that were tested show that
schools do differ in the size of the gender gap in Math and Science. Models that forced all schools to have the same relationship between gender and achievement in these subjects did not fit as well as models that allowed the relationship to vary between schools. Thus, there are some schools that do have smaller gender gaps in these stereotyped academic domains than others, and an important area of future research will be to explore what features of the school environment or climate may be better able to explain this existing variance. This endeavor could be improved by incorporating the previous suggestions above on improving the measurement of climate and looking at classroom and subject specific climates. Also, when considering the potential impact of stereotype threat on students performance, it may be especially important to look for the relationship between underperformance on exams and teachers’ beliefs and expectations for student performance. The well documented relationship between teacher expectations and achievement gaps (e.g. McKown & Weinstein, 2008; Weinstein, 2002) and the literature looking at classroom goal orientations (e.g., Ames & Archer, 1988) suggest pathways by which achievement gaps are established and maintained in the classroom context.

**Limitations and Delimitations**

The biggest limitation on the conclusions to be drawn from these data is that the measure and conceptualization of climate would benefit from further revision and development. In the current study, I used both exploratory and confirmatory factor analyses to refine the measure, but some theoretically important features are not covered, and some psychometric concerns still exist. The concerns regarding referent groups and positive versus negative wording have already been suggested as areas that merit further investigation. In addition, the very construction of aggregate climate factors is also an area to be explored. In this study, as in most of the school and classroom climate research (Taylor, 2011), an average of the student perceptions is used to
describe the school climate, but other options also exist. Chan (1998) developed a typology of composition models within organizational psychology that could be readily adapted to school perceptions. In addition to three models that are variations on mean aggregation as school climate, Chan presents two interesting alternatives: the dispersion model, which focuses on the variability of climate ratings as a feature of the climate itself, and the process model, which conceptualizes parallel processes with corresponding important predictors and outcomes at the student and classroom level of analysis. Examining the level of agreement (i.e., lack of dispersion) on climate measures within a school or testing for equivalent structural relationships at different levels are promising future directions for investigating the impact of climate on academic achievement.

Another weakness within this study design is that the model used assumes that gender-based stereotype threat is elicited by math and science exams for female students even when they only report their gender before an exam without having the stereotype directly invoked. For ideal internal validity conditions, instead of using gender as a proxy for the experience of stereotype threat, I would have some measure or manipulation of the experience itself. However, while this is a threat to the internal validity of the presumed interaction, it does more accurately reflect the state of real world learning environments in which these tests take place. For example, students would not normally receive any special instructions regarding the gender neutrality or bias of their exams.

One major delimitation of this study is that it can only address the existence of relationships between the measured variables and not their causal direction. Most students in the 10th and 11th grade have already been immersed in their school climates for some time. Even for the results where the data did support the hypothesized positive relationship between climate and
achievement, it cannot be concluded that experiencing a positive climate will improve students’ academic achievement scores. Additionally, while the sample used in these analyses is very diverse ethnically and socioeconomically, it is within a single large urban school district. The relationships found in these analyses may not extend to students in rural schools, less diverse populations, or other areas of the country.

Conclusion

Despite the numerous complexities of these results, this study does show a clear relationship between aggregate school climate perceptions and average school performance on a variety of achievement tests, even after controlling for a school’s population size, the mobility rate of its students, racial diversity, and its percentage of low SES students. Thus, at the school-level, climate seems to be a robust predictor of a school’s academic excellence. This study has also served as a test of the implications of a central social psychological theory within a naturally occurring, multilevel learning environment. Evaluative academic achievement tests showed the expected gender gaps in stereotyped domains, even after controlling for other student characteristics that significantly predict achievement. Thus, this study serves as further confirmation of the significance of collective school climate, and it suggests that stereotype threat may play a role in the underperformance of girls on Math and Science exams. However, more research is needed to explore the interaction of these important predictors of academic success.
REFERENCES


environment fit on young adolescents' experiences in schools and in families. 


Tagiuri, R. (1968). The concept of organizational climate. In R. Tagiuri & G. W. Litwin (Eds.), Organizational climate: Explorations of a concept (pp.1-32). Boston: Division of Research, Graduate School of Business Administration, Harvard University.


Table 1

**Bivariate correlations (Pearson’s r) between all variables used in the multilevel models**

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Grade Level</th>
<th>Race</th>
<th>Lunch Status</th>
<th>GPA</th>
<th>Safety</th>
<th>Anti. Peers</th>
<th>Pro. /Student</th>
<th>Teacher Expect.</th>
<th>PLAN Math</th>
<th>PLAN Science</th>
<th>PLAN Reading</th>
<th>PLAN English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>.013*</td>
<td>.020**</td>
<td>.020**</td>
<td>.209**</td>
<td>.025**</td>
<td>.089**</td>
<td>-.034**</td>
<td>.084**</td>
<td>-.021**</td>
<td>.036**</td>
<td>.119**</td>
<td>.114**</td>
</tr>
<tr>
<td>Grade Level</td>
<td>.013*</td>
<td>1</td>
<td>-.032**</td>
<td>-.022**</td>
<td>.111**</td>
<td>.017**</td>
<td>.030**</td>
<td>.030**</td>
<td>.026**</td>
<td>-.058**</td>
<td>.117**</td>
<td>.116**</td>
<td>.100**</td>
</tr>
<tr>
<td>Race</td>
<td>.020**</td>
<td>-.032**</td>
<td>1</td>
<td>.326**</td>
<td>-.300**</td>
<td>-.142**</td>
<td>-.144**</td>
<td>-.067**</td>
<td>-.016**</td>
<td>-.067**</td>
<td>-.394**</td>
<td>-.371**</td>
<td>-.307**</td>
</tr>
<tr>
<td>Lunch Status</td>
<td>.020**</td>
<td>-.022**</td>
<td>.326**</td>
<td>1</td>
<td>-.194**</td>
<td>-.162**</td>
<td>-.111**</td>
<td>-.041**</td>
<td>.010</td>
<td>-.075**</td>
<td>-.313**</td>
<td>-.313**</td>
<td>-.310**</td>
</tr>
<tr>
<td>GPA</td>
<td>.209**</td>
<td>.111**</td>
<td>-.300**</td>
<td>-.194**</td>
<td>1</td>
<td>.166**</td>
<td>.127**</td>
<td>.064**</td>
<td>.199**</td>
<td>-.193**</td>
<td>.555**</td>
<td>.548**</td>
<td>.510**</td>
</tr>
<tr>
<td>Safety</td>
<td>-.025**</td>
<td>.017**</td>
<td>-.142**</td>
<td>-.162**</td>
<td>.166**</td>
<td>1</td>
<td>.412**</td>
<td>.259**</td>
<td>.207**</td>
<td>-.172**</td>
<td>.257**</td>
<td>.249**</td>
<td>.252**</td>
</tr>
<tr>
<td>Antisocial Peers</td>
<td>-.089**</td>
<td>.030**</td>
<td>-.144**</td>
<td>-.111**</td>
<td>.127**</td>
<td>.412**</td>
<td>1</td>
<td>.365**</td>
<td>.069**</td>
<td>.054**</td>
<td>.200**</td>
<td>.182**</td>
<td>.147**</td>
</tr>
<tr>
<td>Prosocial Peers</td>
<td>-.034**</td>
<td>.030**</td>
<td>-.067**</td>
<td>-.041**</td>
<td>.064**</td>
<td>.259**</td>
<td>.365**</td>
<td>1</td>
<td>.335**</td>
<td>.234**</td>
<td>.057**</td>
<td>.061**</td>
<td>.026**</td>
</tr>
<tr>
<td>Teacher/Student</td>
<td>.084**</td>
<td>.026**</td>
<td>-.016**</td>
<td>.010</td>
<td>.199**</td>
<td>.207**</td>
<td>.069**</td>
<td>.335**</td>
<td>1</td>
<td>.403**</td>
<td>.071**</td>
<td>.086**</td>
<td>.068**</td>
</tr>
<tr>
<td>Expectations</td>
<td>.055**</td>
<td>-.058**</td>
<td>-.067**</td>
<td>-.075**</td>
<td>.193**</td>
<td>.172**</td>
<td>.054**</td>
<td>.234**</td>
<td>.403**</td>
<td>1</td>
<td>.166**</td>
<td>.173**</td>
<td>.152**</td>
</tr>
<tr>
<td>PLAN Math</td>
<td>-.021**</td>
<td>.117**</td>
<td>-.394**</td>
<td>-.313**</td>
<td>.555**</td>
<td>.257**</td>
<td>.200**</td>
<td>.057**</td>
<td>.071**</td>
<td>1</td>
<td>.737**</td>
<td>.651**</td>
<td>.723**</td>
</tr>
<tr>
<td>PLAN Science</td>
<td>.036**</td>
<td>.116**</td>
<td>-.371**</td>
<td>-.313**</td>
<td>.548**</td>
<td>.249**</td>
<td>.182**</td>
<td>.061**</td>
<td>.086**</td>
<td>1</td>
<td>.737**</td>
<td>.651**</td>
<td>.694**</td>
</tr>
<tr>
<td>PLAN Reading</td>
<td>.119**</td>
<td>.100**</td>
<td>-.307**</td>
<td>-.310**</td>
<td>.510**</td>
<td>.252**</td>
<td>.147**</td>
<td>.026**</td>
<td>.068**</td>
<td>1</td>
<td>.651**</td>
<td>.694**</td>
<td>.745**</td>
</tr>
<tr>
<td>PLAN English</td>
<td>.114**</td>
<td>.125**</td>
<td>-.358**</td>
<td>-.339**</td>
<td>.548**</td>
<td>.261**</td>
<td>.150**</td>
<td>.023**</td>
<td>.068**</td>
<td>1</td>
<td>.723**</td>
<td>.713**</td>
<td>.745**</td>
</tr>
</tbody>
</table>

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Note: The n available for each pair of variables ranges to 22467 to 27203 depending on the correlation
Table 2

*Mean Scores on the PLAN Tests and Perceptions of Climate by Gender and Race*

<table>
<thead>
<tr>
<th>Gender</th>
<th>Race</th>
<th>M</th>
<th>(SD)</th>
<th>M</th>
<th>(SD)</th>
<th>M</th>
<th>(SD)</th>
<th>M</th>
<th>(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td></td>
<td>White/Asian</td>
<td></td>
<td>Black/Latino/Native</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PLAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Math</td>
<td>15.98</td>
<td>(4.51)</td>
<td>15.80</td>
<td>(4.51)</td>
<td>19.48</td>
<td>(5.08)</td>
<td>15.12</td>
<td>(3.53)</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>17.07</td>
<td>(3.58)</td>
<td>17.32</td>
<td>(3.58)</td>
<td>19.94</td>
<td>(4.16)</td>
<td>16.64</td>
<td>(2.86)</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>15.31</td>
<td>(4.51)</td>
<td>16.39</td>
<td>(4.43)</td>
<td>18.94</td>
<td>(4.98)</td>
<td>15.28</td>
<td>(4.12)</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>15.12</td>
<td>(4.13)</td>
<td>16.07</td>
<td>(4.09)</td>
<td>18.89</td>
<td>(4.75)</td>
<td>14.98</td>
<td>(3.65)</td>
</tr>
<tr>
<td></td>
<td>School Climate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Safety</td>
<td>2.89</td>
<td>(0.65)</td>
<td>2.85</td>
<td>(0.64)</td>
<td>3.07</td>
<td>(0.62)</td>
<td>2.83</td>
<td>(0.64)</td>
</tr>
<tr>
<td></td>
<td>Antisocial Peers</td>
<td>2.41</td>
<td>(0.50)</td>
<td>2.32</td>
<td>(0.52)</td>
<td>2.52</td>
<td>(0.51)</td>
<td>2.32</td>
<td>(0.51)</td>
</tr>
<tr>
<td></td>
<td>Prosocial Peers</td>
<td>2.38</td>
<td>(0.54)</td>
<td>2.34</td>
<td>(0.53)</td>
<td>2.44</td>
<td>(0.54)</td>
<td>2.34</td>
<td>(0.53)</td>
</tr>
<tr>
<td></td>
<td>Teacher/Student</td>
<td>2.70</td>
<td>(0.52)</td>
<td>2.78</td>
<td>(0.49)</td>
<td>2.76</td>
<td>(0.49)</td>
<td>2.74</td>
<td>(0.51)</td>
</tr>
<tr>
<td></td>
<td>Expectations</td>
<td>2.69</td>
<td>(0.60)</td>
<td>2.76</td>
<td>(0.58)</td>
<td>2.81</td>
<td>(0.64)</td>
<td>2.71</td>
<td>(0.58)</td>
</tr>
</tbody>
</table>

*Note: All differences are statistically significant, partially due to the large sample size.*
Table 3

*Unstandardized Regression Weights for Covariates at the Student Level*

<table>
<thead>
<tr>
<th>Simple Linear Regression</th>
<th>PLAN_Math</th>
<th>PLAN_Science</th>
<th>PLAN_Reading</th>
<th>PLAN_English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B$</td>
<td>$B$</td>
<td>$B$</td>
<td>$B$</td>
</tr>
<tr>
<td>Gender</td>
<td>-1.00*</td>
<td>-0.39*</td>
<td>0.30*</td>
<td>0.20*</td>
</tr>
<tr>
<td>Race</td>
<td>-2.13*</td>
<td>-1.53*</td>
<td>-1.36*</td>
<td>-1.69*</td>
</tr>
<tr>
<td>Lunch</td>
<td>-1.62*</td>
<td>-1.39*</td>
<td>-2.17*</td>
<td>-2.12*</td>
</tr>
<tr>
<td>GPA</td>
<td>2.13*</td>
<td>1.66*</td>
<td>2.01*</td>
<td>1.95*</td>
</tr>
<tr>
<td>Grade</td>
<td>0.53*</td>
<td>0.44*</td>
<td>0.47*</td>
<td>0.62*</td>
</tr>
</tbody>
</table>

* significant at $p<.01$
Table 4

*Intracluster (i.e., Intraschool) Correlations and Design Effects*

<table>
<thead>
<tr>
<th></th>
<th>ICC</th>
<th>Av. Clst Size</th>
<th>Design Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School Climate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Safety</td>
<td>0.19</td>
<td>249</td>
<td>47.72</td>
</tr>
<tr>
<td>Antisocial Peers</td>
<td>0.17</td>
<td>249</td>
<td>43.93</td>
</tr>
<tr>
<td>Prosocial Peers</td>
<td>0.09</td>
<td>249</td>
<td>24.51</td>
</tr>
<tr>
<td>Teacher-Student Interaction</td>
<td>0.07</td>
<td>249</td>
<td>18.11</td>
</tr>
<tr>
<td>Institutional Expectations</td>
<td>0.09</td>
<td>249</td>
<td>24.54</td>
</tr>
<tr>
<td><strong>Academic Achievement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_Math</td>
<td>0.35</td>
<td>249</td>
<td>87.51</td>
</tr>
<tr>
<td>PLAN_Science</td>
<td>0.32</td>
<td>249</td>
<td>81.67</td>
</tr>
<tr>
<td>PLAN_Reading</td>
<td>0.30</td>
<td>249</td>
<td>75.83</td>
</tr>
<tr>
<td>PLAN_English</td>
<td>0.34</td>
<td>249</td>
<td>86.29</td>
</tr>
</tbody>
</table>
Table 5

*Unstandardized Regression Weights for Covariates at the School-level*

<table>
<thead>
<tr>
<th></th>
<th>PLAN_Math</th>
<th>PLAN_Science</th>
<th>PLAN_Reading</th>
<th>PLAN_English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lunch</td>
<td>-4.94*</td>
<td>-4.14*</td>
<td>-5.93*</td>
<td>-5.31*</td>
</tr>
<tr>
<td>Mobility</td>
<td>-5.04*</td>
<td>-3.88*</td>
<td>-5.05*</td>
<td>-5.52*</td>
</tr>
<tr>
<td>RDI</td>
<td>2.31*</td>
<td>1.52*</td>
<td>1.58*</td>
<td>1.47*</td>
</tr>
<tr>
<td>Size</td>
<td>0.06*</td>
<td>0.04*</td>
<td>0.05*</td>
<td>0.06*</td>
</tr>
</tbody>
</table>

* *significant at p< .05 (the standard, larger alpha is chosen due to the much smaller number of school units)*

*Note: Population size was rescaled such that a one-unit increase in size represents 100 students.*
Table 6

Parameter estimates for school-level covariates within each subject by climate model

<table>
<thead>
<tr>
<th></th>
<th>PLAN_Math</th>
<th></th>
<th>PLAN_Science</th>
<th></th>
<th>PLAN_Reading</th>
<th></th>
<th>PLAN_English</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.06*</td>
<td>.00</td>
<td>0.03</td>
<td>.01</td>
<td>0.04*</td>
<td>.00</td>
<td>0.05*</td>
<td>.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>-1.82</td>
<td>.05</td>
<td>-1.26</td>
<td>.05</td>
<td>-2.01</td>
<td>.02</td>
<td>-2.41*</td>
<td>.00</td>
</tr>
<tr>
<td>Lunch</td>
<td>-6.25*</td>
<td>.00</td>
<td>-5.64*</td>
<td>.00</td>
<td>-8.02*</td>
<td>.00</td>
<td>-6.93*</td>
<td>.00</td>
</tr>
<tr>
<td>RDI</td>
<td>0.62</td>
<td>.63</td>
<td>0.13</td>
<td>.88</td>
<td>0.27</td>
<td>.81</td>
<td>0.98</td>
<td>.40</td>
</tr>
<tr>
<td><strong>Antisocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.05*</td>
<td>.00</td>
<td>0.02</td>
<td>.03</td>
<td>0.03</td>
<td>.02</td>
<td>0.04*</td>
<td>.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>-3.50*</td>
<td>.00</td>
<td>-2.34*</td>
<td>.00</td>
<td>-3.42*</td>
<td>.00</td>
<td>-4.20*</td>
<td>.00</td>
</tr>
<tr>
<td>Lunch</td>
<td>-6.95*</td>
<td>.00</td>
<td>-6.05*</td>
<td>.00</td>
<td>-8.48*</td>
<td>.00</td>
<td>-7.52*</td>
<td>.00</td>
</tr>
<tr>
<td>RDI</td>
<td>0.38</td>
<td>.77</td>
<td>-0.09</td>
<td>.92</td>
<td>0.02</td>
<td>.99</td>
<td>0.78</td>
<td>.51</td>
</tr>
<tr>
<td><strong>Prosocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.06*</td>
<td>.00</td>
<td>0.03*</td>
<td>.01</td>
<td>0.04*</td>
<td>.01</td>
<td>0.04*</td>
<td>.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>-4.31*</td>
<td>.00</td>
<td>-2.98*</td>
<td>.00</td>
<td>-4.02*</td>
<td>.00</td>
<td>-4.69*</td>
<td>.00</td>
</tr>
<tr>
<td>Lunch</td>
<td>-6.93*</td>
<td>.00</td>
<td>-6.02*</td>
<td>.00</td>
<td>-8.40*</td>
<td>.00</td>
<td>-7.46*</td>
<td>.00</td>
</tr>
<tr>
<td>RDI</td>
<td>0.79</td>
<td>.52</td>
<td>0.22</td>
<td>.79</td>
<td>0.14</td>
<td>.89</td>
<td>0.88</td>
<td>.43</td>
</tr>
<tr>
<td><strong>Teacher</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.07*</td>
<td>.00</td>
<td>0.04*</td>
<td>.00</td>
<td>0.05*</td>
<td>.00</td>
<td>0.05*</td>
<td>.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>-3.66*</td>
<td>.00</td>
<td>-2.54*</td>
<td>.00</td>
<td>-3.48*</td>
<td>.00</td>
<td>-4.01*</td>
<td>.00</td>
</tr>
<tr>
<td>Lunch</td>
<td>-8.89*</td>
<td>.00</td>
<td>-7.50*</td>
<td>.00</td>
<td>-9.78*</td>
<td>.00</td>
<td>-8.72*</td>
<td>.00</td>
</tr>
<tr>
<td>RDI</td>
<td>1.10</td>
<td>.33</td>
<td>0.45</td>
<td>.54</td>
<td>0.16</td>
<td>.86</td>
<td>0.95</td>
<td>.35</td>
</tr>
<tr>
<td><strong>Expectations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.06*</td>
<td>.00</td>
<td>0.03*</td>
<td>.00</td>
<td>0.04*</td>
<td>.00</td>
<td>0.05*</td>
<td>.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>-4.25*</td>
<td>.00</td>
<td>-3.00*</td>
<td>.00</td>
<td>-3.95*</td>
<td>.00</td>
<td>-4.39*</td>
<td>.00</td>
</tr>
<tr>
<td>Lunch</td>
<td>-7.02*</td>
<td>.00</td>
<td>-6.09*</td>
<td>.00</td>
<td>-8.31*</td>
<td>.00</td>
<td>-7.31*</td>
<td>.00</td>
</tr>
<tr>
<td>RDI</td>
<td>0.72</td>
<td>.48</td>
<td>0.23</td>
<td>.72</td>
<td>-0.02</td>
<td>.98</td>
<td>0.62</td>
<td>.51</td>
</tr>
</tbody>
</table>

* significant at p<.01
Table 7

Main Effects of Student Level Climate within the Multilevel Covariate Model

<table>
<thead>
<tr>
<th></th>
<th>PLAN_Math</th>
<th></th>
<th>PLAN_Science</th>
<th></th>
<th>PLAN_Reading</th>
<th></th>
<th>PLAN_English</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-level Climate</td>
<td>0.29*</td>
<td>.00</td>
<td>0.16*</td>
<td>.00</td>
<td>0.26*</td>
<td>.00</td>
<td>0.24*</td>
<td>.00</td>
</tr>
<tr>
<td><strong>Antisocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-level Climate</td>
<td>-0.27*</td>
<td>.00</td>
<td>-0.29*</td>
<td>.00</td>
<td>-0.42*</td>
<td>.00</td>
<td>-0.44*</td>
<td>.00</td>
</tr>
<tr>
<td><strong>Prosocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-level Climate</td>
<td>-0.64*</td>
<td>.00</td>
<td>-0.44*</td>
<td>.00</td>
<td>-0.73*</td>
<td>.00</td>
<td>-0.65*</td>
<td>.00</td>
</tr>
<tr>
<td><strong>Teacher</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-level Climate</td>
<td>-0.11</td>
<td>.08</td>
<td>-0.04</td>
<td>.57</td>
<td>-0.11</td>
<td>.09</td>
<td>-0.12</td>
<td>.03</td>
</tr>
<tr>
<td><strong>Expectations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-level Climate</td>
<td>0.11</td>
<td>.09</td>
<td>0.16*</td>
<td>.00</td>
<td>0.22*</td>
<td>.00</td>
<td>0.17*</td>
<td>.00</td>
</tr>
</tbody>
</table>

* significant at p<.01

Note: All climate measures are coded such that larger scores indicate a more positive climate.
Table 8

**Main Effects of School-level Climate within the Multilevel Covariate Model**

<table>
<thead>
<tr>
<th>安全</th>
<th>计划数学</th>
<th>计划科学</th>
<th>计划阅读</th>
<th>计划英语</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Sig.</td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
</tr>
<tr>
<td>安全</td>
<td>4.97</td>
<td>0.00</td>
<td>3.65</td>
<td>0.00</td>
</tr>
<tr>
<td>反社会</td>
<td>5.01</td>
<td>0.00</td>
<td>3.91</td>
<td>0.00</td>
</tr>
<tr>
<td>亲社会</td>
<td>5.83</td>
<td>0.00</td>
<td>4.62</td>
<td>0.00</td>
</tr>
<tr>
<td>教师期待</td>
<td>7.64</td>
<td>0.00</td>
<td>6.11</td>
<td>0.00</td>
</tr>
<tr>
<td>计划数学</td>
<td>5.69</td>
<td>0.00</td>
<td>4.38</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*significant at p<.01

*Note: All climate measures are coded such that larger scores indicate a more positive climate.*
Figure 1: Student and school level components of the multilevel covariate model

a: Student Level

b: School Level
Figure 2: Multilevel latent covariate model for PLAN Math scores

Note: Paths marked “**” are inconsistent; see Appendix C
Figure 3: Multilevel latent covariate model for PLAN Science scores

Note: Paths marked "**" are inconsistent; see Appendix C
Figure 4: Multilevel latent covariate model for PLAN Reading scores

Note: Paths marked "**" are inconsistent; see Appendix C
Figure 5: Multilevel latent covariate model for PLAN English scores

Note: Paths marked "**" are inconsistent; see Appendix C
Figure 6: Simple slopes for interaction effects between individual level climate and gender

6a: Gender*Antisocial Peers interaction for PLAN Math

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_M

6b: Gender*Antisocial Peers interaction for PLAN Science

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_S
Figure 6 (continued)

6c: Gender*Safety interaction for PLAN Reading

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_R

6d: Gender*Safety interaction for PLAN English

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_E
Figure 7: Simple slopes for interaction effects between race and gender on PLAN Math tests

Note: The graph below collapses across all climate models, as the interaction was near identical in each.

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_S
Figure 8: Simple slopes for interaction effects between race and gender on PLAN Science tests

Note: The graph below collapses across all climate models, as the interaction was near identical in each.

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN S
Figure 9: Simple slopes for interaction effects between school level climate and gender on PLAN Science tests

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_S
Figure 9 (continued):

**Gender*Prosocial for Science**

<table>
<thead>
<tr>
<th>Science Score</th>
<th>Low (1 SD Below)</th>
<th>Mean</th>
<th>High (1 SD Above)</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Males
- Females

**Gender*Expectations for Science**

<table>
<thead>
<tr>
<th>Science Score</th>
<th>Low (1 SD Below)</th>
<th>Mean</th>
<th>High (1 SD Above)</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Males
- Females
Figure 10: Simple slopes for interaction effects between school level climate and gender on PLAN Math tests

Y-axis values were chosen to reflect a range of approximately -1 SD to +1 SD on the PLAN_M
Figure 10 (continued):

Gender*Prosocial for Math

Gender*Expectations for Math
APPENDIX A (continued)

Original Item Classification from the Conditions for Learning Survey

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>▼ I worry about crime and violence in school. _R</td>
</tr>
<tr>
<td>2</td>
<td>▼ Students at this school are often teased or picked on. _R</td>
</tr>
<tr>
<td>3</td>
<td>▼ Students at this school are often threatened or bullied. _R</td>
</tr>
<tr>
<td>4</td>
<td>▼ I feel safe when security is present.</td>
</tr>
<tr>
<td>5</td>
<td>▼ I sometimes stay home because I don't feel safe at school. _R</td>
</tr>
<tr>
<td>6</td>
<td>* How safe do you feel outside around the school.</td>
</tr>
<tr>
<td>8</td>
<td>* How safe do you feel in the hallways and bathrooms of the school.</td>
</tr>
<tr>
<td>9</td>
<td>* How safe do you feel in your classes.</td>
</tr>
<tr>
<td>10</td>
<td>▼ Most students in my school don't really care about each other. _R</td>
</tr>
<tr>
<td>11</td>
<td>▼ Most students in my school like to put others down. _R</td>
</tr>
<tr>
<td>13</td>
<td>▼ Most students in my school don't get along together very well. _R</td>
</tr>
<tr>
<td>14</td>
<td>▼ Most students in my school just look out for themselves. _R</td>
</tr>
<tr>
<td>15</td>
<td>▼ Most students in my school treat each other with respect</td>
</tr>
</tbody>
</table>

Note. All items followed by a _R are reversed scored.

▼ = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree

* = Item scale: Not Safe, Somewhat Safe, Mostly Safe, Very Safe
## Support Scale Items

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>My teachers really care about me.</td>
</tr>
<tr>
<td>31</td>
<td>My teachers help me make up work after an excused absence.</td>
</tr>
<tr>
<td>32</td>
<td>My teachers give me feedback on my assignments that helps me improve my work.</td>
</tr>
<tr>
<td>35</td>
<td>Adults in this school are often too busy to give students extra help. _R</td>
</tr>
<tr>
<td>36</td>
<td>Adults in the school apply the same rules to all students equally.</td>
</tr>
<tr>
<td>37</td>
<td>I wish I went to a different school. _R</td>
</tr>
<tr>
<td>39</td>
<td>A counselor at this school has helped me plan for life after high school.</td>
</tr>
<tr>
<td>38</td>
<td>I can get extra help at school outside of my regular classes.</td>
</tr>
<tr>
<td>49</td>
<td>This school year, how often have you talked to a teacher about a problem you were having in class.</td>
</tr>
<tr>
<td>50</td>
<td>This school year, how often have you talked to an adult at school about something that was bothering you.</td>
</tr>
<tr>
<td>51</td>
<td>This school year, how often have you talked to an adult at school about something outside of school that is important to you.</td>
</tr>
<tr>
<td>52</td>
<td>This school year, how often have you talked to a counselor at school in depth about planning for college.</td>
</tr>
<tr>
<td>62</td>
<td>My teacher for my class closest to but before lunch notices if I have trouble learning something.</td>
</tr>
<tr>
<td>63</td>
<td>My teacher for my class closest to but before lunch will help me improve my work if I do poorly on an assignment.</td>
</tr>
</tbody>
</table>

**Note.** All items followed by a \_R are reversed scored.

\_ = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree

\_R = Item scale: Never, 1 or 2 Times, 3 or 4 Times, 5 or More Times
### Expectations Scale Items

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>My teachers often connect what I am learning to life outside the classroom.</td>
</tr>
<tr>
<td>28</td>
<td>My teachers encourage students to share their ideas about things we are studying in class.</td>
</tr>
<tr>
<td>29</td>
<td>My teachers often require me to explain my answers.</td>
</tr>
<tr>
<td>33</td>
<td>My teachers often assign homework that helps me learn.</td>
</tr>
<tr>
<td>34</td>
<td>My teachers think all students can do challenging work.</td>
</tr>
<tr>
<td>40</td>
<td>When students in this school already know the material that is being taught, the teacher gives them more-advanced assignments.</td>
</tr>
<tr>
<td>41</td>
<td>In my classes, we often discuss different interpretations of things we read.</td>
</tr>
<tr>
<td>42</td>
<td>Students in this school are expected to take four years of math.</td>
</tr>
<tr>
<td>43</td>
<td>Students in this school are expected to take four years of science.</td>
</tr>
<tr>
<td>44</td>
<td>Students in this school are expected to take more than two years of a foreign language.</td>
</tr>
<tr>
<td>45</td>
<td>Students in this school are encouraged to take advanced classes such as honors, Advanced Placement (AP), or International Baccalaureate (IB), or classes that lead to professional certification.</td>
</tr>
<tr>
<td>46</td>
<td>This school year, how often have your teachers given you an assignment to write a research paper of 5 or more pages using multiple sources of information?</td>
</tr>
<tr>
<td>47</td>
<td>This school year, how often have your teachers given you an assignment to write a paper in which you defended your own point of view or ideas?</td>
</tr>
<tr>
<td>48</td>
<td>This school year, how often have your teachers given you an assignment to make a formal presentation to a class about something you read or researched?</td>
</tr>
<tr>
<td>64</td>
<td>In this class the topics we are studying are interesting and challenging.</td>
</tr>
<tr>
<td>65</td>
<td>This class really makes me think.</td>
</tr>
<tr>
<td>66</td>
<td>I am usually bored in this class.</td>
</tr>
</tbody>
</table>

*Note. All items followed by a _R are reversed scored.*

- **V** = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree
- **⊥** = Item scale: Never, 1 or 2 Times, 3 or 4 Times, 5 or More Times
### SEL Scale Items

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Most students in my school stop and think before doing anything when they get angry.</td>
</tr>
<tr>
<td>17</td>
<td>Most students in my school do their share of work when we have group projects.</td>
</tr>
<tr>
<td>18</td>
<td>Most students in my school give up when they can't solve a problem easily.  _R</td>
</tr>
<tr>
<td>19</td>
<td>Most students in my school get into arguments when they disagree with people.  _R</td>
</tr>
<tr>
<td>20</td>
<td>Most students in my school do their best, even when their school work is difficult.</td>
</tr>
<tr>
<td>21</td>
<td>Most students in my school think it's OK to fight if someone insults them.  _R</td>
</tr>
<tr>
<td>22</td>
<td>Most students in my school do all their homework.</td>
</tr>
<tr>
<td>23</td>
<td>Most students in my school say mean things to other students when they think the other students deserve it.  _R</td>
</tr>
<tr>
<td>24</td>
<td>Most students in my school try to work out their disagreements with other students by talking to them.</td>
</tr>
<tr>
<td>25</td>
<td>Most students in my school think it's ok to cheat if other students are cheating.  _R</td>
</tr>
<tr>
<td>26</td>
<td>Most students in my school try to do a good job on school work even when it is not interesting.</td>
</tr>
</tbody>
</table>

*Note.* All items followed by a  _R are reversed scored.

∀ = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree
Appendix B: Recategorizing the Climate Survey Items Using Factor Analysis

**Method and results of factor analyses.** Due to the fact that analyses have indicated that the hypothetical factor structure of the Conditions for Learning Survey does not hold, an initial confirmatory factor analysis (CFA) was conducted that regressed the items on the theoretical factors as described in Appendix A. This fit of this model was not quite satisfactory ($\chi^2 = 200140.099$, $ns$; RMSEA = 0.07; CFI = 0.598; SRMR = 0.080). For large sample sizes and complex models, a strict criteria (i.e., CFI $\geq$ .95, RMSEA $\leq$ .06, and SRMR $\leq$ .08) is recommended when evaluating model fit statistics (Weston & Gore, 2006), and this model suggested poor fit on both the CFI and the RMSEA.

Therefore, in an effort to construct more reliable factors from the climate data, an EFA was conducted using a random sample of half of the cases. First, to confirm that the climate structure was multidimensional, the fit for a one-factor model was examined. The EFA indicated that a one-factor model was a consistently poor fit of the data ($\chi^2 = 196531.596$, $ns$; RMSEA = 0.089; CFI = 0.393; SRMR = 0.103), which allowed me to reject a unidimensional model of school climate. Next, I examined the SCREE plot, to determine the correct number of factors to explore, and it indicated that between 3 and 7 might exist. The factor loadings for a five factor model were examined, and are shown below, with values highlighted for the factor that each item loaded most strongly upon. These results showed that the poor fit of the CFA was not simply due to a few item misfits or multiple loadings, but that the proposed structure of the factors as a whole was not emerging. Items that were theorized to be of the same scale were separating, and other distinct factors were emerging. When items that were not loading strongly or did not fit well theoretically in the emerging factor were removed, a reduced group of 36 items remained. They load on five factors that I call Physical Safety, Antisocial Peers, Prosocial Peers, Teacher/Student Interaction, and Institutional Expectations.

I then conducted a second CFA with the remaining random half of the sample students, and it indicated that the reduced 5 factor model had a more acceptable fit (RMSEA = 0.053; CFI = 0.858; SRMR = 0.054), although the improved CFI was still lower than ideal. The model results, with the complete items listed under their new factors including reliability for each of the new scales, are shown in Tables B1-B5 below. Throughout the remaining analyses, the mean score on each of these factors was used to examine perceptions of school climate. For all climate variables, including Antisocial Peers, a larger value indicates a more favorable view of the climate, as all items evaluating negative climate remained reverse scored.

Shown below are the loadings of each item in the 5-factor solution. Bolded factor loadings show the strongest loading for an item. Italicized items indicate those that did not load well (i.e., above 0.35) on any factor. Items numbers are grouped by their original factor and item number. (e.g., LI16 is item #16, originally in the SEL subscale). Item names with an asterisk (*) were removed or included in a different factor than their loading indicates in order to maintain psychometrically consistent subscales.
NOTE: For the items listed below, item numbers that start with an “S” were originally designed to load on Safety, “L” on Social and Emotional Learning, “E” on High Expectations, and “T” on Student Support.

<table>
<thead>
<tr>
<th>Item</th>
<th>Antisocial</th>
<th>Safety</th>
<th>Teacher</th>
<th>Prosocial</th>
<th>Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI1_R</td>
<td>0.322</td>
<td></td>
<td>0.435</td>
<td>-0.101</td>
<td>-0.050</td>
</tr>
<tr>
<td>SI2_R</td>
<td>0.558</td>
<td>0.196</td>
<td></td>
<td>-0.047</td>
<td>0.015</td>
</tr>
<tr>
<td>SI3_R</td>
<td>0.554</td>
<td>0.290</td>
<td></td>
<td>-0.039</td>
<td>0.003</td>
</tr>
<tr>
<td>SI4</td>
<td>0.084</td>
<td>0.153</td>
<td>0.207</td>
<td>0.181</td>
<td></td>
</tr>
<tr>
<td>SI5_R</td>
<td>0.149</td>
<td>0.428</td>
<td>0.117</td>
<td>-0.177</td>
<td>-0.034</td>
</tr>
<tr>
<td>SI6</td>
<td>0.071</td>
<td>0.648</td>
<td></td>
<td>-0.033</td>
<td>0.089</td>
</tr>
<tr>
<td>SI8</td>
<td>-0.001</td>
<td>0.872</td>
<td></td>
<td>-0.012</td>
<td>0.055</td>
</tr>
<tr>
<td>SI9</td>
<td>-0.018</td>
<td>0.754</td>
<td>0.099</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>SI10_R</td>
<td>0.655</td>
<td>0.059</td>
<td>0.081</td>
<td></td>
<td>-0.014</td>
</tr>
<tr>
<td>SI11_R</td>
<td>0.723</td>
<td>0.051</td>
<td>0.025</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>SI13</td>
<td>0.625</td>
<td>0.112</td>
<td>0.058</td>
<td></td>
<td>-0.040</td>
</tr>
<tr>
<td>SI14_R</td>
<td>0.536</td>
<td>0.035</td>
<td>0.061</td>
<td></td>
<td>-0.032</td>
</tr>
<tr>
<td>SI15</td>
<td>0.215</td>
<td>0.100</td>
<td>0.040</td>
<td></td>
<td>0.506</td>
</tr>
<tr>
<td>LI16</td>
<td>0.094</td>
<td>0.003</td>
<td>-0.086</td>
<td></td>
<td>0.616</td>
</tr>
<tr>
<td>LI17</td>
<td>-0.008</td>
<td>0.048</td>
<td>0.106</td>
<td></td>
<td>0.502</td>
</tr>
<tr>
<td>LI18_R</td>
<td>0.493</td>
<td>0.005</td>
<td>0.066</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td>LI19_R</td>
<td>0.537</td>
<td>-0.079</td>
<td>-0.041</td>
<td>0.108</td>
<td>0.005</td>
</tr>
<tr>
<td>LI20</td>
<td>-0.030</td>
<td>0.031</td>
<td>0.121</td>
<td></td>
<td>0.593</td>
</tr>
<tr>
<td>LI21_R</td>
<td>0.556</td>
<td>-0.038</td>
<td>-0.042</td>
<td></td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>LI22</td>
<td>0.063</td>
<td>-0.031</td>
<td>-0.004</td>
<td><strong>0.644</strong></td>
<td>0.011</td>
</tr>
<tr>
<td>LI23_R</td>
<td><strong>0.567</strong></td>
<td>-0.096</td>
<td>-0.057</td>
<td>0.167</td>
<td>0.020</td>
</tr>
<tr>
<td>LI24</td>
<td>0.087</td>
<td>0.017</td>
<td>-0.030</td>
<td><strong>0.620</strong></td>
<td>0.042</td>
</tr>
<tr>
<td>LI25_R</td>
<td><strong>0.467</strong></td>
<td>-0.086</td>
<td>0.011</td>
<td>0.192</td>
<td>0.014</td>
</tr>
<tr>
<td>LI26</td>
<td>-0.016</td>
<td>0.028</td>
<td>0.083</td>
<td><strong>0.627</strong></td>
<td>-0.007</td>
</tr>
<tr>
<td>EI27</td>
<td>0.008</td>
<td>-0.009</td>
<td><strong>0.626</strong></td>
<td>0.003</td>
<td>0.017</td>
</tr>
<tr>
<td>EI28</td>
<td>-0.008</td>
<td>0.022</td>
<td><strong>0.707</strong></td>
<td>-0.001</td>
<td>-0.043</td>
</tr>
<tr>
<td>EI29</td>
<td>-0.085</td>
<td>0.091</td>
<td><strong>0.597</strong></td>
<td>-0.055</td>
<td>-0.024</td>
</tr>
<tr>
<td>EI33</td>
<td>0.016</td>
<td>-0.038</td>
<td><strong>0.698</strong></td>
<td>0.022</td>
<td>-0.026</td>
</tr>
<tr>
<td>EI34</td>
<td>-0.022</td>
<td>0.077</td>
<td><strong>0.555</strong></td>
<td>0.015</td>
<td>-0.029</td>
</tr>
<tr>
<td>EI40</td>
<td>-0.059</td>
<td>-0.059</td>
<td>0.297</td>
<td>0.204</td>
<td><strong>0.054</strong></td>
</tr>
<tr>
<td>EI41</td>
<td>-0.033</td>
<td>0.070</td>
<td><strong>0.465</strong></td>
<td>0.100</td>
<td>0.012</td>
</tr>
<tr>
<td>EI42</td>
<td>-0.083</td>
<td>0.088</td>
<td>0.251</td>
<td>0.154</td>
<td><strong>0.011</strong></td>
</tr>
<tr>
<td>EI43</td>
<td>-0.070</td>
<td>0.077</td>
<td>0.202</td>
<td>0.187</td>
<td><strong>0.051</strong></td>
</tr>
<tr>
<td>EI44</td>
<td>-0.078</td>
<td>0.056</td>
<td>0.201</td>
<td>0.185</td>
<td>0.028</td>
</tr>
<tr>
<td>EI45</td>
<td>-0.034</td>
<td>0.130</td>
<td><strong>0.375</strong></td>
<td>0.033</td>
<td>0.005</td>
</tr>
<tr>
<td>EI43_A</td>
<td>-0.044</td>
<td>0.004</td>
<td>0.096</td>
<td>0.084</td>
<td><strong>0.211</strong></td>
</tr>
<tr>
<td>EI44_A</td>
<td>0.014</td>
<td>0.107</td>
<td>0.234</td>
<td>-0.027</td>
<td>0.208</td>
</tr>
<tr>
<td>EI45_A</td>
<td>0.019</td>
<td>0.110</td>
<td>0.192</td>
<td>0.028</td>
<td>0.200</td>
</tr>
<tr>
<td>EI60</td>
<td>0.023</td>
<td>-0.006</td>
<td><strong>0.453</strong></td>
<td>0.026</td>
<td>0.056</td>
</tr>
<tr>
<td>EI61</td>
<td>0.006</td>
<td>-0.001</td>
<td><strong>0.451</strong></td>
<td>0.009</td>
<td>0.070</td>
</tr>
<tr>
<td>EI62_R</td>
<td>0.188</td>
<td>-0.082</td>
<td>0.261</td>
<td>-0.046</td>
<td><strong>0.015</strong></td>
</tr>
<tr>
<td>TI30</td>
<td>0.082</td>
<td>-0.037</td>
<td><strong>0.644</strong></td>
<td>0.024</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>TI31</td>
<td>0.020</td>
<td>-0.045</td>
<td><strong>0.641</strong></td>
<td>0.016</td>
<td>-0.008</td>
</tr>
<tr>
<td>TI32</td>
<td>0.022</td>
<td>-0.043</td>
<td><strong>0.731</strong></td>
<td>0.001</td>
<td>-0.012</td>
</tr>
<tr>
<td>TI35_R</td>
<td>0.328</td>
<td>0.030</td>
<td><strong>0.264</strong></td>
<td>-0.071</td>
<td><strong>0.003</strong></td>
</tr>
<tr>
<td>TI36*</td>
<td>0.077</td>
<td>-0.028</td>
<td><strong>0.301</strong></td>
<td>0.273</td>
<td>-0.012</td>
</tr>
<tr>
<td>TI37_R</td>
<td><strong>0.369</strong></td>
<td>0.131</td>
<td>0.161</td>
<td>0.001</td>
<td>-0.010</td>
</tr>
<tr>
<td>TI38</td>
<td>0.010</td>
<td>0.062</td>
<td><strong>0.385</strong></td>
<td>0.088</td>
<td>0.071</td>
</tr>
<tr>
<td>TI39</td>
<td><strong>0.008</strong></td>
<td>-0.031</td>
<td><strong>0.342</strong></td>
<td>0.104</td>
<td><strong>0.226</strong></td>
</tr>
<tr>
<td>TI46</td>
<td>-0.021</td>
<td>0.047</td>
<td>0.091</td>
<td>-0.030</td>
<td><strong>0.606</strong></td>
</tr>
<tr>
<td>TI47</td>
<td>-0.017</td>
<td>-0.010</td>
<td>-0.055</td>
<td>-0.017</td>
<td><strong>0.837</strong></td>
</tr>
<tr>
<td>TI48</td>
<td>0.030</td>
<td>0.007</td>
<td>-0.013</td>
<td>-0.012</td>
<td><strong>0.772</strong></td>
</tr>
<tr>
<td>TI52</td>
<td>-0.004</td>
<td>-0.048</td>
<td>0.040</td>
<td>0.088</td>
<td><strong>0.501</strong></td>
</tr>
<tr>
<td>TI57</td>
<td>-0.011</td>
<td>-0.002</td>
<td><strong>0.466</strong></td>
<td>-0.033</td>
<td>0.055</td>
</tr>
<tr>
<td>TI58</td>
<td>0.015</td>
<td>-0.016</td>
<td><strong>0.517</strong></td>
<td>-0.044</td>
<td>0.051</td>
</tr>
</tbody>
</table>
Table B1

Safety scale using the reduced set of climate items

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>▼ I worry about crime and violence in school. _R</td>
</tr>
<tr>
<td>5</td>
<td>▼ I sometimes stay home because I don't feel safe at school. _R</td>
</tr>
<tr>
<td>6</td>
<td>* How safe do you feel outside around the school.</td>
</tr>
<tr>
<td>8</td>
<td>* How safe do you feel in the hallways and bathrooms of the school.</td>
</tr>
<tr>
<td>9</td>
<td>* How safe do you feel in your classes.</td>
</tr>
</tbody>
</table>

Note. All items followed by a _R are reversed scored.

▼ = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree

* = Item scale: Not Safe, Somewhat Safe, Mostly Safe, Very Safe
Table B2

*Antisocial Peers scale using the reduced set of climate items*

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Students at this school are often teased or picked on. _R</td>
</tr>
<tr>
<td>3</td>
<td>Students at this school are often threatened or bullied. _R</td>
</tr>
<tr>
<td>10</td>
<td>Most students in my school don't really care about each other. _R</td>
</tr>
<tr>
<td>11</td>
<td>Most students in my school like to put others down. _R</td>
</tr>
<tr>
<td>13</td>
<td>Most students in my school don't get along together very well. _R</td>
</tr>
<tr>
<td>14</td>
<td>Most students in my school just look out for themselves. _R</td>
</tr>
<tr>
<td>18</td>
<td>Most students in my school give up when they can't solve a problem easily. _R</td>
</tr>
<tr>
<td>19</td>
<td>Most students in my school get into arguments when they disagree with people. _R</td>
</tr>
<tr>
<td>21</td>
<td>Most students in my school think its OK to fight if someone insults them. _R</td>
</tr>
<tr>
<td>23</td>
<td>Most students in my school say mean things to other students when they think the other students deserve it. _R</td>
</tr>
<tr>
<td>25</td>
<td>Most students in my school think its ok to cheat if other students are cheating. _R</td>
</tr>
</tbody>
</table>

*Note.* All items followed by a _R are reversed scored.

\(\triangledown\) = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree
Table B3

Prosocial Peers scale using the reduced set of climate items

Prosocial Peers Scale Items (alpha = .82)

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Most students in my school treat each other with respect</td>
</tr>
<tr>
<td>16</td>
<td>Most students in my school stop and think before doing anything when they get angry.</td>
</tr>
<tr>
<td>17</td>
<td>Most students in my school do their share of work when we have group projects.</td>
</tr>
<tr>
<td>20</td>
<td>Most students in my school do their best, even when their school work is difficult.</td>
</tr>
<tr>
<td>22</td>
<td>Most students in my school do all their homework.</td>
</tr>
<tr>
<td>24</td>
<td>Most students in my school try to work out their disagreements with other students by talking to them.</td>
</tr>
<tr>
<td>26</td>
<td>Most students in my school try to do a good job on school work even when it is not interesting.</td>
</tr>
</tbody>
</table>

*Note.* All items followed by a _R are reversed scored.

템 = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree
Table B4

*Teacher scale using the reduced set of climate items*

*Teacher-Student Interaction Scale Items (alpha = .86)*

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>My teachers often connect what I am learning to life outside the classroom.</td>
</tr>
<tr>
<td>28</td>
<td>My teachers encourage students to share their ideas about things we are studying in class.</td>
</tr>
<tr>
<td>29</td>
<td>My teachers often require me to explain my answers.</td>
</tr>
<tr>
<td>30</td>
<td>My teachers really care about me.</td>
</tr>
<tr>
<td>31</td>
<td>My teachers help me make up work after an excused absence.</td>
</tr>
<tr>
<td>32</td>
<td>My teachers give me feedback on my assignments that helps me improve my work.</td>
</tr>
<tr>
<td>33</td>
<td>My teachers often assign homework that helps me learn.</td>
</tr>
<tr>
<td>34</td>
<td>My teachers think all students can do challenging work.</td>
</tr>
<tr>
<td>41</td>
<td>In my classes, we often discuss different interpretations of things we read.</td>
</tr>
</tbody>
</table>

*Note.* All items followed by a _R are reversed scored.

∀ = Item scale: Strongly Disagree, Disagree, Agree, Strongly Agree
Table B5

*Expectations scale using the reduced set of climate items*

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>Students in this school are expected to take four years of math.</td>
</tr>
<tr>
<td>43</td>
<td>Students in this school are expected to take four years of science.</td>
</tr>
<tr>
<td>44</td>
<td>Students in this school are expected to take more than two years of a foreign language.</td>
</tr>
<tr>
<td>45</td>
<td>Students in this school are encouraged to take advanced classes such as honors, Advanced Placement (AP), or International Baccalaureate (IB), or classes that lead to professional certification.</td>
</tr>
</tbody>
</table>

*Note. All items followed by a _R are reversed scored]*
Appendix C: Model Outputs

PLAN_M: Models Predicting Mathematics Test Scores

1a. Climate measure: Safety

MODEL FIT INFORMATION

Number of Free Parameters 29

Loglikelihood

<table>
<thead>
<tr>
<th>H0 Value</th>
<th>-76313.364</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0 Scaling Correction Factor</td>
<td>2.824</td>
</tr>
<tr>
<td>for MLR</td>
<td></td>
</tr>
</tbody>
</table>

Information Criteria

<table>
<thead>
<tr>
<th>Akaike (AIC)</th>
<th>152684.728</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian (BIC)</td>
<td>152921.076</td>
</tr>
<tr>
<td>Sample-Size Adjusted BIC</td>
<td>152828.915</td>
</tr>
<tr>
<td>(n* = (n + 2) / 24)</td>
<td></td>
</tr>
</tbody>
</table>

MODEL RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two-Tailed Estimate</td>
<td>S.E.</td>
<td>Est./S.E.</td>
<td>P-Value</td>
</tr>
<tr>
<td>Within Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_M ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.652</td>
<td>0.093</td>
<td>6.999</td>
<td>0.000</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.290</td>
<td>0.090</td>
<td>-3.234</td>
<td>0.001</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.591</td>
<td>0.064</td>
<td>24.853</td>
<td>0.000</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.010</td>
<td>0.066</td>
<td>0.156</td>
<td>0.876</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.433</td>
<td>0.120</td>
<td>3.602</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTS ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.021</td>
<td>0.006</td>
<td>3.360</td>
<td>0.001</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.039</td>
<td>0.009</td>
<td>-4.590</td>
<td>0.000</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.057</td>
<td>0.015</td>
<td>-3.890</td>
<td>0.000</td>
</tr>
<tr>
<td>RERACE</td>
<td>0.031</td>
<td>0.010</td>
<td>3.155</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTS</td>
<td>0.038</td>
<td>0.011</td>
<td>3.390</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_M</td>
<td>8.193</td>
<td>0.351</td>
<td>23.361</td>
<td>0.000</td>
</tr>
<tr>
<td>CENTS</td>
<td>0.336</td>
<td>0.009</td>
<td>38.281</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
S1  ON
  CGM_S   -1.020  0.195  -5.243  0.000

S2  ON
  CGM_S   0.281  0.114   2.457  0.014

S3  ON
  CGM_S   -1.391  0.730  -1.906  0.057
  CGMRDI   0.390  1.332   0.293  0.770

PLAN_M  ON
  RECGMPOP  0.059  0.017   3.448  0.001
  CGMMOB   -1.815  0.944  -1.923  0.054
  CGMFRL   -6.247  1.549  -4.033  0.000
  CGMRDI   0.625  1.282   0.488  0.626
  CGM_S    4.947  1.165   4.246  0.000

Intercepts
  PLAN_M      17.030  0.286  59.453  0.000
  S1       -1.221  0.121  -10.066  0.000
  S2        0.286  0.049   5.804  0.000
  S3       -1.005  0.228   4.407  0.000

Residual Variances
  PLAN_M    0.895  0.216   4.134  0.000
  S1       0.039  0.019   2.030  0.042
  S2       0.005  0.017   0.297  0.767
  S3       0.400  0.204   1.959  0.050

1b. Climate measure: Antisocial Peers

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood
  H0 Value    -69663.489
  H0 Scaling Correction Factor 2.834
    for MLR

Information Criteria
  Akaike (AIC)  139392.979
  Bayesian (BIC)  139661.926
  Sample-Size Adjusted BIC  139557.053
    (n* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed Estimate  S.E. Est./S.E.  P-Value
Within Level

<table>
<thead>
<tr>
<th>PLAN_M ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GRADE</td>
<td>0.656</td>
<td>0.086</td>
<td>7.602</td>
<td>0.000</td>
</tr>
<tr>
<td>LEP</td>
<td>-1.463</td>
<td>0.139</td>
<td>-10.550</td>
<td>0.000</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.287</td>
<td>0.087</td>
<td>-26.192</td>
<td>0.000</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.313</td>
<td>0.085</td>
<td>-3.670</td>
<td>0.000</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.416</td>
<td>0.065</td>
<td>21.701</td>
<td>0.000</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.231</td>
<td>0.079</td>
<td>2.933</td>
<td>0.003</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.370</td>
<td>0.118</td>
<td>3.147</td>
<td>0.002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CENTA ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTGPA</td>
<td>0.002</td>
<td>0.006</td>
<td>0.410</td>
<td>0.682</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.002</td>
<td>0.010</td>
<td>0.228</td>
<td>0.820</td>
</tr>
<tr>
<td>LEP</td>
<td>0.118</td>
<td>0.015</td>
<td>7.800</td>
<td>0.000</td>
</tr>
<tr>
<td>SPED</td>
<td>0.084</td>
<td>0.012</td>
<td>7.216</td>
<td>0.000</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.093</td>
<td>0.013</td>
<td>-7.164</td>
<td>0.000</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.006</td>
<td>0.010</td>
<td>0.605</td>
<td>0.545</td>
</tr>
</tbody>
</table>

Intercepts

| CENTA     | 0.028 | 0.012 | 2.294 | 0.022 |

Residual Variances

<table>
<thead>
<tr>
<th>PLAN_M</th>
<th>7.629</th>
<th>0.371</th>
<th>20.591</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTA</td>
<td>0.212</td>
<td>0.004</td>
<td>57.215</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Between Level

<table>
<thead>
<tr>
<th>S1 ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>-1.254</td>
<td>0.241</td>
<td>-5.193</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2 ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>0.297</td>
<td>0.215</td>
<td>1.385</td>
<td>0.166</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3 ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>-1.558</td>
<td>0.896</td>
<td>-1.738</td>
<td>0.082</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>0.129</td>
<td>1.279</td>
<td>0.101</td>
<td>0.920</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PLAN_M ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RECGMPOP</td>
<td>0.049</td>
<td>0.017</td>
<td>2.861</td>
<td>0.004</td>
</tr>
<tr>
<td>CGMMOB</td>
<td>-3.162</td>
<td>1.010</td>
<td>-3.131</td>
<td>0.002</td>
</tr>
<tr>
<td>CGMFRL</td>
<td>-6.393</td>
<td>1.608</td>
<td>-3.975</td>
<td>0.000</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>0.322</td>
<td>1.273</td>
<td>0.253</td>
<td>0.801</td>
</tr>
<tr>
<td>CGM_A</td>
<td>5.013</td>
<td>1.467</td>
<td>3.418</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Intercepts

<table>
<thead>
<tr>
<th>PLAN_M</th>
<th>17.503</th>
<th>0.272</th>
<th>64.421</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>-1.312</td>
<td>0.117</td>
<td>-11.251</td>
<td>0.000</td>
</tr>
<tr>
<td>S2</td>
<td>-0.171</td>
<td>0.058</td>
<td>-2.926</td>
<td>0.003</td>
</tr>
<tr>
<td>S3</td>
<td>-1.080</td>
<td>0.236</td>
<td>-4.584</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Residual Variances

<table>
<thead>
<tr>
<th>PLAN_M</th>
<th>0.975</th>
<th>0.208</th>
<th>4.686</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.027</td>
<td>0.014</td>
<td>1.862</td>
<td>0.063</td>
</tr>
<tr>
<td>S2</td>
<td>0.005</td>
<td>0.043</td>
<td>0.106</td>
<td>0.916</td>
</tr>
<tr>
<td>S3</td>
<td>0.437</td>
<td>0.203</td>
<td>2.151</td>
<td>0.031</td>
</tr>
</tbody>
</table>
### 1c. Climate measure: Prosocial Peers

#### MODEL FIT INFORMATION

<table>
<thead>
<tr>
<th>Number of Free Parameters</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loglikelihood</td>
<td></td>
</tr>
<tr>
<td>H0 Value</td>
<td>-72043.243</td>
</tr>
<tr>
<td>H0 Scaling Correction Factor</td>
<td>2.767</td>
</tr>
<tr>
<td>Information Criteria</td>
<td></td>
</tr>
<tr>
<td>Akaike (AIC)</td>
<td>144152.486</td>
</tr>
<tr>
<td>Bayesian (BIC)</td>
<td>144421.433</td>
</tr>
<tr>
<td>Sample-Size Adjusted BIC</td>
<td>144316.560</td>
</tr>
<tr>
<td>((n^* = (n + 2) / 24))</td>
<td></td>
</tr>
</tbody>
</table>

#### MODEL RESULTS

<table>
<thead>
<tr>
<th>Two-Tailed</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.665</td>
<td>0.086</td>
<td>7.712</td>
<td>0.000</td>
</tr>
<tr>
<td>LEP</td>
<td>-1.388</td>
<td>0.141</td>
<td>-9.836</td>
<td>0.000</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.232</td>
<td>0.085</td>
<td>-26.176</td>
<td>0.000</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.290</td>
<td>0.084</td>
<td>-3.450</td>
<td>0.001</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.414</td>
<td>0.066</td>
<td>21.421</td>
<td>0.000</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.147</td>
<td>0.078</td>
<td>1.888</td>
<td>0.059</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.402</td>
<td>0.109</td>
<td>3.698</td>
<td>0.000</td>
</tr>
<tr>
<td>CENTP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>-0.013</td>
<td>0.006</td>
<td>-2.219</td>
<td>0.026</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.033</td>
<td>0.008</td>
<td>4.386</td>
<td>0.000</td>
</tr>
<tr>
<td>LEP</td>
<td>0.168</td>
<td>0.016</td>
<td>10.633</td>
<td>0.000</td>
</tr>
<tr>
<td>SPED</td>
<td>0.127</td>
<td>0.014</td>
<td>8.863</td>
<td>0.000</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.030</td>
<td>0.009</td>
<td>-3.198</td>
<td>0.001</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.008</td>
<td>0.009</td>
<td>0.958</td>
<td>0.338</td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTP</td>
<td>-0.042</td>
<td>0.009</td>
<td>-4.777</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Variances</td>
<td>PLAN_M</td>
<td>7.579</td>
<td>0.367</td>
<td>20.634</td>
</tr>
<tr>
<td>CENTP</td>
<td>0.257</td>
<td>0.006</td>
<td>44.825</td>
<td>0.000</td>
</tr>
<tr>
<td>Between Level</td>
<td>S1</td>
<td>ON</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2 ON CGM_P</td>
<td>0.410</td>
<td>0.289</td>
<td>1.422</td>
<td>0.155</td>
</tr>
<tr>
<td>S3 ON CGM_P</td>
<td>-2.022</td>
<td>0.942</td>
<td>-2.146</td>
<td>0.032</td>
</tr>
<tr>
<td>CGM_P</td>
<td>-1.330</td>
<td>0.335</td>
<td>-3.971</td>
<td>0.000</td>
</tr>
<tr>
<td>S3 ON CGM_P</td>
<td>0.086</td>
<td>1.207</td>
<td>0.071</td>
<td>0.943</td>
</tr>
<tr>
<td>S1</td>
<td>-1.329</td>
<td>0.112</td>
<td>-11.876</td>
<td>0.000</td>
</tr>
<tr>
<td>S2</td>
<td>-0.492</td>
<td>0.064</td>
<td>-7.688</td>
<td>0.000</td>
</tr>
<tr>
<td>S3</td>
<td>-1.133</td>
<td>0.230</td>
<td>-4.926</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**1d. Climate measure: Teacher-Student Interactions**

**MODEL FIT INFORMATION**

Number of Free Parameters 33

Loglikelihood

- H0 Value: -71003.558
- H0 Scaling Correction Factor: 2.585 for MLR

Information Criteria

- Akaike (AIC): 142073.115
- Bayesian (BIC): 142342.061
- Sample-Size Adjusted BIC: 142237.188
  \( n^* = (n + 2) / 24 \)

**MODEL RESULTS**

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
</table>

Within Level


PLANN M ON  
GRADE 0.656 0.086 7.595 0.000  
LEP -1.468 0.138 -10.639 0.000  
SPED -2.294 0.088 -26.042 0.000  
LUNCH -0.311 0.086 -3.625 0.000  
CENTGPA 1.427 0.066 21.480 0.000  
X1X2 0.070 0.076 0.919 0.358  
X1X3 0.483 0.112 4.323 0.000  

CENTTT ON  
CENTGPA 0.102 0.004 23.926 0.000  
LUNCH 0.021 0.008 2.645 0.008  
LEP 0.023 0.017 1.315 0.188  
SPED 0.015 0.009 1.735 0.083  
GENDER 0.044 0.007 6.389 0.000  
RACETHRT 0.020 0.008 2.655 0.008  

Intercepts  
CENTTT -0.062 0.008 -8.170 0.000  

Residual Variances  
PLAN_M 7.636 0.369 20.698 0.000  
CENTTT 0.236 0.005 43.572 0.000  

Between Level  
S1 ON  
CGM_T -0.710 0.484 -1.468 0.142  
S2 ON  
CGM_T 0.708 0.265 2.670 0.008  
S3 ON  
CGM_T -2.234 1.309 -1.707 0.088  
CGMRDI -0.131 1.051 -0.124 0.901  

PLAN_M ON  
RECGMPOP 0.071 0.015 4.753 0.000  
CGMPOP -3.258 0.897 -3.631 0.000  
CGMFRL -8.330 1.614 -5.161 0.000  
CGMRDI 1.066 1.120 0.952 0.341  
CGM_T 7.491 1.949 3.842 0.000  

Intercepts  
PLAN_M 17.639 0.259 68.181 0.000  
S1 -1.348 0.114 -11.805 0.000  
S2 -0.086 0.056 -1.537 0.124  
S3 -1.210 0.230 -5.263 0.000  

Residual Variances  
PLAN_M 0.816 0.196 4.158 0.000  
S1 0.071 0.033 2.124 0.034  
S2 0.003 0.024 0.105 0.916  
S3 0.409 0.185 2.214 0.027
### 1e. Climate measure: Expectations

**MODEL FIT INFORMATION**

Number of Free Parameters 33

Loglikelihood

- H0 Value: -73838.581
- H0 Scaling Correction Factor: 2.478 for MLR

Information Criteria

- Akaike (AIC): 147743.161
- Bayesian (BIC): 148011.825
- Sample-Size Adjusted BIC: 147906.952
  \[(n^* = (n + 2) / 24)\]

**MODEL RESULTS**

<table>
<thead>
<tr>
<th>Two-Tailed Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within Level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_M ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.664</td>
<td>0.088</td>
<td>7.564</td>
</tr>
<tr>
<td>LEP</td>
<td>-1.473</td>
<td>0.137</td>
<td>-10.733</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.280</td>
<td>0.088</td>
<td>-26.030</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.316</td>
<td>0.085</td>
<td>-3.711</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.414</td>
<td>0.065</td>
<td>21.895</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.077</td>
<td>0.069</td>
<td>1.108</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.444</td>
<td>0.119</td>
<td>3.722</td>
</tr>
<tr>
<td><strong>CENTE ON</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.066</td>
<td>0.006</td>
<td>10.981</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.000</td>
<td>0.010</td>
<td>-0.023</td>
</tr>
<tr>
<td>LEP</td>
<td>0.037</td>
<td>0.020</td>
<td>1.832</td>
</tr>
<tr>
<td>SPED</td>
<td>-0.058</td>
<td>0.012</td>
<td>-4.959</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.021</td>
<td>0.008</td>
<td>2.593</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.022</td>
<td>0.009</td>
<td>2.411</td>
</tr>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTE</td>
<td>-0.023</td>
<td>0.008</td>
<td>-2.811</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_M</td>
<td>7.618</td>
<td>0.368</td>
<td>20.715</td>
</tr>
<tr>
<td>CENTE</td>
<td>0.309</td>
<td>0.005</td>
<td>59.096</td>
</tr>
</tbody>
</table>

**Between Level**

<table>
<thead>
<tr>
<th>S1 ON</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Plan M</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_E</td>
<td>-0.965</td>
<td>0.213</td>
<td>-4.537</td>
<td>0.000</td>
</tr>
<tr>
<td>S2 ON</td>
<td>0.358</td>
<td>0.260</td>
<td>1.378</td>
<td>0.168</td>
</tr>
<tr>
<td>S3 ON</td>
<td>0.331</td>
<td>1.004</td>
<td>0.330</td>
<td>0.741</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plan M ON</th>
<th>Recgmpop</th>
<th>Cgmmob</th>
<th>Cgmfrl</th>
<th>Cgmrdi</th>
<th>Cgm_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_M</td>
<td>0.061</td>
<td>-3.855</td>
<td>-6.492</td>
<td>0.694</td>
<td>5.631</td>
</tr>
<tr>
<td>S1</td>
<td>-1.326</td>
<td>0.123</td>
<td>-10.811</td>
<td>-1.207</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>0.048</td>
<td>0.061</td>
<td>0.783</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>-1.207</td>
<td>0.218</td>
<td>-5.524</td>
<td>0.350</td>
<td></td>
</tr>
</tbody>
</table>

Intercepts

<table>
<thead>
<tr>
<th>Variable</th>
<th>Plan M</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_M</td>
<td>17.617</td>
<td>0.248</td>
<td>70.914</td>
<td>0.000</td>
</tr>
<tr>
<td>S1</td>
<td>-1.326</td>
<td>0.123</td>
<td>-10.811</td>
<td>0.000</td>
</tr>
<tr>
<td>S2</td>
<td>0.048</td>
<td>0.061</td>
<td>0.783</td>
<td>0.434</td>
</tr>
<tr>
<td>S3</td>
<td>-1.207</td>
<td>0.218</td>
<td>-5.524</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Residual Variances

<table>
<thead>
<tr>
<th>Variable</th>
<th>Plan M</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_M</td>
<td>0.738</td>
<td>0.233</td>
<td>3.161</td>
<td>0.002</td>
</tr>
<tr>
<td>S1</td>
<td>0.052</td>
<td>0.026</td>
<td>1.992</td>
<td>0.046</td>
</tr>
<tr>
<td>S2</td>
<td>0.048</td>
<td>0.024</td>
<td>1.966</td>
<td>0.049</td>
</tr>
<tr>
<td>S3</td>
<td>0.350</td>
<td>0.172</td>
<td>2.041</td>
<td>0.041</td>
</tr>
</tbody>
</table>
PLAN_S: Models Predicting Science Test Scores

2a. Climate measure: Safety

MODEL FIT INFORMATION

Number of Free Parameters                       33

Loglikelihood

   H0 Value             -71586.197
   H0 Scaling Correction Factor   3.155
      for MLR

Information Criteria

      Akaike (AIC)              143238.394
      Bayesian (BIC)            143507.342
      Sample-Size Adjusted BIC  143402.468
   (n* = (n + 2) / 24)

MODEL RESULTS

   Two-Tailed
   Estimate    S.E.   Est./S.E.    P-Value

Within Level

PLAN_S ON
   GRADE       0.522    0.069     7.613     0.000
   LEP         -1.141    0.099    -11.490    0.000
   SPED        -1.401    0.074    -18.969    0.000
   LUNCH       -0.311    0.074     -4.203    0.000
   CENTGPA     1.126    0.050     22.464    0.000
   X1X2        0.103    0.053      1.926    0.054
   X1X3        0.254    0.079      3.214    0.001

CENTS ON
   CENTGPA    0.017    0.006      2.608    0.009
   LUNCH      -0.025    0.008     -3.030    0.002
   LEP        -0.234    0.020    -11.490    0.000
   SPED       -0.060    0.012     -4.913    0.000
   GENDER     -0.063    0.015     -4.257    0.000
   RACETHRT    0.016    0.011      1.456    0.145

Intercepts
   CENTS      0.060    0.014      4.287    0.000

Residual Variances
   PLAN_S      5.415    0.325     16.640    0.000
   CENTS      0.334    0.009     38.081    0.000
Between Level

S1  ON
   CGM_S  -0.868  0.120  -7.240  0.000

S2  ON
   CGM_S  0.449  0.101  4.444  0.000

S3  ON
   CGM_S  -0.796  0.618  -1.288  0.198
   CGMRDI  0.370  0.880  0.420  0.674

PLAN_S  ON
   RECGMPOP  0.030  0.011  2.705  0.007
   CGMMOB  -0.996  0.638  -1.563  0.118
   CGMFRL  -5.288  0.920  -5.745  0.000
   CGMRDI  0.156  0.845  0.184  0.854
   CGM_S  3.615  0.861  4.199  0.000

Intercepts
   PLAN_S  18.280  0.185  98.889  0.000
   S1  -0.597  0.077  -7.748  0.000
   S2   0.106  0.042  2.514  0.012
   S3  -0.815  0.152  -5.381  0.000

Residual Variances
   PLAN_S  0.458  0.107  4.270  0.000
   S1   0.001  0.019  0.060  0.952
   S2   0.006  0.011  0.537  0.591
   S3   0.173  0.077  2.235  0.025

2b. Climate measure: Antisocial Peers

MODEL FIT INFORMATION

Number of Free Parameters  33

Loglikelihood
   H0 Value  -65818.555
   H0 Scaling Correction Factor  3.131
                     for MLR

Information Criteria
   Akaike (AIC)  131703.111
   Bayesian (BIC)  131972.058
   Sample-Size Adjusted BIC  131867.185
   (n* = (n + 2) / 24)

MODEL RESULTS
## Two-Tailed Estimate | S.E. | Est./S.E. | P-Value

### Within Level

**PLAN_S ON**
- **GRADE**: 0.529, 0.068, 7.798, 0.000
- **LEP**: -1.155, 0.098, -11.790, 0.000
- **SPED**: -1.399, 0.074, -18.924, 0.000
- **LUNCH**: -0.319, 0.072, -4.431, 0.000
- **CENTGPA**: 1.126, 0.051, 22.128, 0.000
- **X1X2**: 0.242, 0.075, 3.232, 0.001
- **X1X3**: 0.264, 0.083, 3.170, 0.002

**CENTA ON**
- **CENTGPA**: 0.002, 0.006, 0.410, 0.682
- **LUNCH**: 0.002, 0.010, 0.228, 0.820
- **LEP**: 0.118, 0.015, 7.800, 0.000
- **SPED**: 0.084, 0.012, 7.216, 0.000
- **GENDER**: -0.093, 0.013, -7.164, 0.000
- **RACETHRT**: 0.006, 0.010, 0.605, 0.545

**Intercepts**
- **CENTA**: 0.028, 0.012, 2.294, 0.022

### Residual Variances

**PLAN_S**: 5.420, 0.325, 16.685, 0.000
**CENTA**: 0.212, 0.004, 57.215, 0.000

### Between Level

**S1 ON**
- **CGM_A**: -0.942, 0.164, -5.735, 0.000

**S2 ON**
- **CGM_A**: 0.610, 0.150, 4.065, 0.000

**S3 ON**
- **CGM_A**: -0.930, 0.829, -1.122, 0.262
- **CGMRDI**: 0.058, 0.869, 0.067, 0.947

**PLAN_S ON**
- **RECGMPOP**: 0.023, 0.011, 2.096, 0.036
- **CGMMOB**: -2.123, 0.706, -3.006, 0.003
- **CGMFRL**: -5.708, 0.960, -5.943, 0.000
- **CGMRDI**: -0.100, 0.857, -0.117, 0.907
- **CGM_A**: 3.847, 1.143, 3.365, 0.001

**Intercepts**
- **PLAN_S**: 18.238, 0.194, 93.821, 0.000
- **S1**: -0.628, 0.081, -7.792, 0.000
- **S2**: -0.225, 0.058, -3.889, 0.000
- **S3**: -0.770, 0.157, -4.897, 0.000

### Residual Variances

**PLAN_S**: 0.491, 0.104, 4.703, 0.000
S1  0.002  0.018  0.098  0.922  
S2  0.009  0.016  0.540  0.589  
S3  0.194  0.081  2.389  0.017  

2c. Climate measure: Prosocial Peers

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood
H0 Value -68213.419
H0 Scaling Correction Factor 3.067 for MLR

Information Criteria
Akaike (AIC) 136492.838
Bayesian (BIC) 136761.785
Sample-Size Adjusted BIC 136656.912
(n* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed
Estimate  S.E.  Est./S.E.  P-Value

Within Level
PLAN_S ON
GRADE  0.534  0.068  7.841  0.000
LEP    -1.108  0.099 -11.223  0.000
SPED   -1.367  0.074 -18.583  0.000
LUNCH  -0.300  0.069 -4.369  0.000
CENTGPA 1.126  0.051  22.055  0.000
X1X2   0.099  0.062  1.591  0.112
X1X3   0.309  0.084  3.663  0.000

CENTP ON
CENTGPA -0.013  0.006 -2.219  0.026
LUNCH   0.033  0.008  4.386  0.000
LEP     0.168  0.016 10.633  0.000
SPED    0.127  0.014  8.863  0.000
GENDER -0.030  0.009 -3.198  0.001
RACETHRT 0.008  0.009  0.958  0.338

Intercepts
CENTP  -0.042  0.009 -4.777  0.000

Residual Variances
PLAN_S  5.393  0.322 16.770  0.000
CENTP   0.257  0.006 44.825  0.000
Between Level

S1  ON
  CGM_P  -1.205  0.212  -5.693  0.000

S2  ON
  CGM_P  0.770  0.250  3.075  0.002

S3  ON
  CGM_P  -1.189  0.909  -1.308  0.191
  CGMRDI  0.058  0.813  0.071  0.943

PLAN_S ON
  RECGMPOP  0.031  0.012  2.695  0.007
  CGMMOB  -2.782  0.734  -3.793  0.000
  CGMFRL  -5.694  0.939  -6.065  0.000
  CGMRDI  0.210  0.803  0.262  0.793
  CGM_P  4.507  1.319  3.417  0.001

Intercepts
  PLAN_S  18.256  0.181  100.628  0.000
  S1  -0.664  0.083  -7.989  0.000
  S2  -0.341  0.047  -7.231  0.000
  S3  -0.823  0.158  -5.222  0.000

Residual Variances
  PLAN_S  0.480  0.106  4.529  0.000
  S1    0.002  0.019  0.101  0.920
  S2    0.027  0.016  1.756  0.079
  S3    0.197  0.077  2.562  0.010

2d. Climate measure: Teacher-Student Interactions

MODEL FIT INFORMATION

Number of Free Parameters  33

Loglikelihood
  H0 Value  -67167.299
  H0 Scaling Correction Factor  2.903
    for MLR

Information Criteria
  Akaike (AIC)  134400.598
  Bayesian (BIC)  134669.544
  Sample-Size Adjusted BIC  134564.671
    (n* = (n + 2) / 24)
### MODEL RESULTS

#### Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
</table>

**Within Level**

| PLAN_S ON | | | |
|-----------| | | |
| GRADE     | 0.527 | 0.068 | 7.733 | 0.000 |
| LEP       | -1.169 | 0.097 | -12.100 | 0.000 |
| SPED      | -1.418 | 0.074 | -19.185 | 0.000 |
| LUNCH     | -0.322 | 0.073 | -4.385 | 0.000 |
| CENTGPA   | 1.128 | 0.052 | 21.674 | 0.000 |
| X1X2      | 0.132 | 0.080 | 1.654 | 0.098 |
| X1X3      | 0.434 | 0.095 | 4.556 | 0.000 |

| CENTT ON | | | |
|-----------| | | |
| CENTGPA   | 0.102 | 0.004 | 23.926 | 0.000 |
| LUNCH     | 0.021 | 0.008 | 2.645 | 0.008 |
| LEP       | 0.023 | 0.017 | 1.315 | 0.188 |
| SPED      | 0.015 | 0.009 | 1.735 | 0.083 |
| GENDER    | 0.044 | 0.007 | 6.389 | 0.000 |
| RACETHRT  | 0.020 | 0.008 | 2.655 | 0.008 |

**Intercepts**

| CENTT     | -0.062 | 0.008 | -8.170 | 0.000 |

**Residual Variances**

| PLAN_S     | 5.432 | 0.326 | 16.642 | 0.000 |
| CENTT      | 0.236 | 0.005 | 43.572 | 0.000 |

**Between Level**

| S1 ON | | | |
|-------| | | |
| CGM_T | -1.059 | 0.410 | -2.584 | 0.010 |

| S2 ON | | | |
|-------| | | |
| CGM_T | 0.609 | 0.281 | 2.166 | 0.030 |

| S3 ON | | | |
|-------| | | |
| CGM_T | -1.667 | 1.216 | -1.371 | 0.170 |
| CGMRDI | -0.053 | 0.716 | -0.075 | 0.941 |

| PLAN_S ON | | | |
|-----------| | | |
| RECGMPOP  | 0.040 | 0.010 | 3.971 | 0.000 |
| CGMOMB    | -2.290 | 0.620 | -3.694 | 0.000 |
| CGMFRL    | -7.146 | 0.913 | -7.828 | 0.000 |
| CGMRDI    | 0.445 | 0.720 | 0.617 | 0.537 |
| CGM_T     | 6.045 | 1.625 | 3.719 | 0.000 |

**Intercepts**

| PLAN_S | 18.378 | 0.181 | 101.541 | 0.000 |
| S1     | -0.746 | 0.098 | -7.618 | 0.000 |
| S2     | -0.024 | 0.059 | -0.409 | 0.682 |
| S3     | -0.912 | 0.157 | -5.812 | 0.000 |
APPENDIX C (continued)

Residual Variances

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_S</td>
<td>0.419</td>
<td>0.107</td>
<td>3.896</td>
<td>0.000</td>
</tr>
<tr>
<td>S1</td>
<td>0.012</td>
<td>0.009</td>
<td>1.325</td>
<td>0.185</td>
</tr>
<tr>
<td>S2</td>
<td>0.002</td>
<td>0.008</td>
<td>0.308</td>
<td>0.758</td>
</tr>
<tr>
<td>S3</td>
<td>0.185</td>
<td>0.076</td>
<td>2.452</td>
<td>0.014</td>
</tr>
</tbody>
</table>

2e. Climate measure: Expectations

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood

- H0 Value: -70024.694
- H0 Scaling Correction Factor: 2.834
  for MLR

Information Criteria

- Akaike (AIC): 140115.388
- Bayesian (BIC): 140384.052
- Sample-Size Adjusted BIC: 140279.179
  \( (n^* = (n + 2) / 24) \)

MODEL RESULTS

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_S ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.539</td>
<td>0.068</td>
<td>7.972</td>
</tr>
<tr>
<td>LEP</td>
<td>-1.168</td>
<td>0.098</td>
<td>-11.886</td>
</tr>
<tr>
<td>SPED</td>
<td>-1.411</td>
<td>0.073</td>
<td>-19.439</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.320</td>
<td>0.074</td>
<td>-4.313</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.123</td>
<td>0.051</td>
<td>22.056</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.005</td>
<td>0.059</td>
<td>0.081</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.376</td>
<td>0.102</td>
<td>3.674</td>
</tr>
</tbody>
</table>

| CENTE ON |
| CENTGPA  | 0.066 | 0.006    | 10.981  | 0.000   |
| LUNCH    | 0.000 | 0.010    | -0.023  | 0.982   |
| LEP      | 0.037 | 0.020    | 1.832   | 0.067   |
| SPED     | -0.058| 0.012    | -4.959  | 0.000   |
| GENDER   | 0.021 | 0.008    | 2.593   | 0.010   |
| RACETHRT | 0.022 | 0.009    | 2.411   | 0.016   |

Intercepts

| CENTE | -0.023 | 0.008 | -2.811 | 0.005 |

Residual Variances
### APPENDIX C (continued)

<table>
<thead>
<tr>
<th></th>
<th>5.416</th>
<th>0.326</th>
<th>16.636</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTE</td>
<td>0.309</td>
<td>0.005</td>
<td>59.096</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Between Level**

<table>
<thead>
<tr>
<th>Level</th>
<th>ON</th>
<th>CGM_E</th>
<th>0.999</th>
<th>0.192</th>
<th>-5.211</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td></td>
<td>-0.999</td>
<td>0.192</td>
<td>-5.211</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.460</td>
<td>0.296</td>
<td>1.552</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td></td>
<td>-1.171</td>
<td>0.645</td>
<td>-1.816</td>
<td>0.069</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.230</td>
<td>0.644</td>
<td>0.357</td>
<td>0.721</td>
<td></td>
</tr>
</tbody>
</table>

**Intercepts**

<table>
<thead>
<tr>
<th></th>
<th>18.338</th>
<th>0.163</th>
<th>112.809</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>-0.704</td>
<td>0.109</td>
<td>-6.441</td>
<td>0.000</td>
</tr>
<tr>
<td>S2</td>
<td>0.122</td>
<td>0.044</td>
<td>2.763</td>
<td>0.006</td>
</tr>
<tr>
<td>S3</td>
<td>-0.895</td>
<td>0.142</td>
<td>-6.289</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Residual Variances**

<table>
<thead>
<tr>
<th></th>
<th>0.395</th>
<th>0.132</th>
<th>2.995</th>
<th>0.003</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.002</td>
<td>0.011</td>
<td>0.156</td>
<td>0.876</td>
</tr>
<tr>
<td>S2</td>
<td>0.024</td>
<td>0.012</td>
<td>2.003</td>
<td>0.045</td>
</tr>
<tr>
<td>S3</td>
<td>0.166</td>
<td>0.073</td>
<td>2.295</td>
<td>0.022</td>
</tr>
</tbody>
</table>
PLAN_E: Models Predicting English Test Scores

3a. Climate measure: Safety

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood

H0 Value -75478.140
H0 Scaling Correction Factor 2.345 for MLR

Information Criteria

Akaike (AIC) 151022.279
Bayesian (BIC) 151291.226
Sample-Size Adjusted BIC 151186.353
(n* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.700</td>
<td>0.088</td>
<td>7.995</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.708</td>
<td>0.107</td>
<td>-25.232</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.238</td>
<td>0.087</td>
<td>-25.643</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.733</td>
<td>0.120</td>
<td>-6.119</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.264</td>
<td>0.053</td>
<td>24.023</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.261</td>
<td>0.075</td>
<td>3.504</td>
</tr>
<tr>
<td>X1X3</td>
<td>-0.153</td>
<td>0.112</td>
<td>-1.360</td>
</tr>
<tr>
<td>CENTS ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.017</td>
<td>0.006</td>
<td>2.608</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.025</td>
<td>0.008</td>
<td>-3.030</td>
</tr>
<tr>
<td>LEP</td>
<td>-0.234</td>
<td>0.020</td>
<td>-11.490</td>
</tr>
<tr>
<td>SPED</td>
<td>-0.060</td>
<td>0.012</td>
<td>-4.913</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.063</td>
<td>0.015</td>
<td>-4.257</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.016</td>
<td>0.011</td>
<td>1.456</td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTS</td>
<td>0.060</td>
<td>0.014</td>
<td>4.287</td>
</tr>
<tr>
<td>Residual Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>7.707</td>
<td>0.274</td>
<td>28.127</td>
</tr>
<tr>
<td>CENTS</td>
<td>0.334</td>
<td>0.009</td>
<td>38.081</td>
</tr>
</tbody>
</table>
## APPENDIX C (continued)

Between Level

<table>
<thead>
<tr>
<th>Level</th>
<th>ON</th>
<th>CGM_S</th>
<th>CGMRDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>ON</td>
<td>0.073</td>
<td>0.067</td>
</tr>
<tr>
<td>S2</td>
<td>ON</td>
<td>0.444</td>
<td>-0.067</td>
</tr>
<tr>
<td>S3</td>
<td>ON</td>
<td>-0.067</td>
<td>-0.811</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.161</td>
<td>0.595</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.857</td>
<td>-0.113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.452</td>
<td>-0.910</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.651</td>
<td>0.910</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>ON</th>
<th>RECGMPOP</th>
<th>CGMMOB</th>
<th>CGMFRL</th>
<th>CGMRDI</th>
<th>CGM_S</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E</td>
<td>ON</td>
<td>0.048</td>
<td>-1.961</td>
<td>-6.211</td>
<td>1.098</td>
<td>2.857</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.014</td>
<td>0.804</td>
<td>1.036</td>
<td>1.138</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.423</td>
<td>-2.437</td>
<td>-5.994</td>
<td>0.965</td>
<td>3.068</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.001</td>
<td>0.015</td>
<td>0.000</td>
<td>0.335</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Intercepts

<table>
<thead>
<tr>
<th>Level</th>
<th>PLAN_E</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16.601</td>
<td>0.294</td>
<td>0.143</td>
<td>-0.636</td>
</tr>
<tr>
<td></td>
<td>0.240</td>
<td>0.092</td>
<td>0.047</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>69.028</td>
<td>3.200</td>
<td>3.032</td>
<td>-2.731</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.001</td>
<td>0.002</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Residual Variances

<table>
<thead>
<tr>
<th>Level</th>
<th>PLAN_E</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.678</td>
<td>0.006</td>
<td>0.002</td>
<td>0.258</td>
</tr>
<tr>
<td></td>
<td>0.124</td>
<td>0.017</td>
<td>0.010</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>5.466</td>
<td>0.361</td>
<td>0.208</td>
<td>2.230</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.718</td>
<td>0.835</td>
<td>0.026</td>
</tr>
</tbody>
</table>

### 3b. Climate measure: Antisocial Peers

**MODEL FIT INFORMATION**

Number of Free Parameters: 33

Loglikelihood

- H0 Value: -69719.904
- H0 Scaling Correction Factor: 71.928 for MLR

Information Criteria

- Akaike (AIC): 139505.809
- Bayesian (BIC): 139774.756
- Sample-Size Adjusted BIC: 139669.883

\( n^* = (n + 2) / 24 \)

**MODEL RESULTS**
Two-Tailed Estimate       S.E.  Est./S.E.    P-Value

Within Level

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.711</td>
<td>0.075</td>
<td>9.424</td>
<td>0.000</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.232</td>
<td>0.219</td>
<td>-10.182</td>
<td>0.000</td>
</tr>
<tr>
<td>SPED</td>
<td>-0.742</td>
<td>0.528</td>
<td>-1.405</td>
<td>0.160</td>
</tr>
<tr>
<td>LUNCH</td>
<td>1.271</td>
<td>0.042</td>
<td>30.550</td>
<td>0.000</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.073</td>
<td>0.563</td>
<td>0.129</td>
<td>0.897</td>
</tr>
<tr>
<td>X1X3</td>
<td>-0.212</td>
<td>0.520</td>
<td>-0.407</td>
<td>0.684</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.002</td>
<td>1.419</td>
<td>0.002</td>
<td>0.999</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.002</td>
<td>3.221</td>
<td>0.001</td>
<td>0.999</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.118</td>
<td>0.145</td>
<td>0.812</td>
<td>0.417</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>-0.093</td>
<td>0.060</td>
<td>-1.557</td>
<td>0.119</td>
</tr>
</tbody>
</table>

Intercepts

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTA</td>
<td>0.028</td>
<td>0.176</td>
<td>0.158</td>
<td>0.875</td>
</tr>
</tbody>
</table>

Residual Variances

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E</td>
<td>7.716</td>
<td>0.098</td>
<td>78.521</td>
<td>0.000</td>
</tr>
<tr>
<td>CENTA</td>
<td>0.212</td>
<td>0.008</td>
<td>26.140</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Between Level

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 ON</td>
<td>CGM_A</td>
<td>0.141</td>
<td>1.047</td>
<td>0.135</td>
</tr>
<tr>
<td>S2 ON</td>
<td>CGM_A</td>
<td>0.462</td>
<td>1.751</td>
<td>0.264</td>
</tr>
<tr>
<td>S3 ON</td>
<td>CGM_A</td>
<td>-0.410</td>
<td>7.063</td>
<td>-0.058</td>
</tr>
<tr>
<td></td>
<td>CGMRDI</td>
<td>-0.896</td>
<td>4.625</td>
<td>-0.194</td>
</tr>
<tr>
<td>PLAN_E ON</td>
<td>RECGMPOP</td>
<td>0.037</td>
<td>0.160</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>CGMMOB</td>
<td>-3.847</td>
<td>1.446</td>
<td>-2.661</td>
</tr>
<tr>
<td></td>
<td>CGMFRDL</td>
<td>-6.818</td>
<td>0.310</td>
<td>-22.010</td>
</tr>
<tr>
<td></td>
<td>CGMRDI</td>
<td>0.836</td>
<td>3.834</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>CGM_A</td>
<td>2.240</td>
<td>5.704</td>
<td>0.393</td>
</tr>
</tbody>
</table>

Intercepts

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E</td>
<td>16.537</td>
<td>0.131</td>
<td>125.907</td>
<td>0.000</td>
</tr>
<tr>
<td>S1</td>
<td>0.304</td>
<td>0.552</td>
<td>0.551</td>
<td>0.582</td>
</tr>
<tr>
<td>S2</td>
<td>-0.329</td>
<td>0.278</td>
<td>-1.183</td>
<td>0.237</td>
</tr>
<tr>
<td>S3</td>
<td>-0.582</td>
<td>0.566</td>
<td>-1.029</td>
<td>0.304</td>
</tr>
</tbody>
</table>

Residual Variances

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_E</td>
<td>0.790</td>
<td>0.283</td>
<td>2.789</td>
<td>0.005</td>
</tr>
</tbody>
</table>
3c. **Climate measure: Prosocial Peers**

**MODEL FIT INFORMATION**

Number of Free Parameters 33

Loglikelihood

- H0 Value: -72051.940
- H0 Scaling Correction Factor: 2.169 for MLR

Information Criteria

- Akaike (AIC): 144169.881
- Bayesian (BIC): 144438.828
- Sample-Size Adjusted BIC: 144333.955 (n* = (n + 2) / 24)

**MODEL RESULTS**

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.720</td>
<td>0.087</td>
<td>8.314</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.649</td>
<td>0.110</td>
<td>-24.122</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.177</td>
<td>0.084</td>
<td>-25.967</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.709</td>
<td>0.116</td>
<td>-6.099</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.266</td>
<td>0.053</td>
<td>24.047</td>
</tr>
<tr>
<td>X1X2</td>
<td>-0.139</td>
<td>0.076</td>
<td>-1.828</td>
</tr>
<tr>
<td>X1X3</td>
<td>-0.194</td>
<td>0.104</td>
<td>-1.859</td>
</tr>
<tr>
<td>CENTP ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>-0.013</td>
<td>0.006</td>
<td>-2.219</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.033</td>
<td>0.008</td>
<td>4.386</td>
</tr>
<tr>
<td>LEP</td>
<td>0.168</td>
<td>0.016</td>
<td>10.633</td>
</tr>
<tr>
<td>SPED</td>
<td>0.127</td>
<td>0.014</td>
<td>8.863</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.030</td>
<td>0.009</td>
<td>-3.198</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.008</td>
<td>0.009</td>
<td>0.958</td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTP</td>
<td>-0.042</td>
<td>0.009</td>
<td>-4.777</td>
</tr>
<tr>
<td>Residual Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>7.626</td>
<td>0.269</td>
<td>28.312</td>
</tr>
<tr>
<td>CENTP</td>
<td>0.257</td>
<td>0.006</td>
<td>44.825</td>
</tr>
</tbody>
</table>
Between Level

S1 ON
CGM_P  0.206  0.193  1.067  0.286

S2 ON
CGM_P  0.582  0.273  2.128  0.033

S3 ON
CGM_P  -1.016  0.742 -1.370  0.171
CGMRDI -0.701  1.069 -0.656  0.512

PLAN_E ON
RECGMPOP  0.042  0.014  3.143  0.002
CGMMOB  -4.362  0.930  -4.691  0.000
CGMFRL  -6.782  1.060  -6.398  0.000
CGMRDI  0.949  1.067  0.889  0.374
CGM_P   2.862  1.267  2.258  0.024

Intercepts
PLAN_E  16.508  0.233  70.766  0.000
S1     0.299  0.081  3.670  0.000
S2   -0.483  0.067  -7.175  0.000
S3   -0.610  0.231  -2.643  0.008

Residual Variances
PLAN_E  0.748  0.150  4.995  0.000
S1     0.005  0.025  0.188  0.851
S2     0.071  0.028  2.563  0.010
S3     0.354  0.134  2.636  0.008

3d. Climate measure: Teacher-Student Interactions

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood
H0 Value -71066.718
H0 Scaling Correction Factor 2.047 for MLR

Information Criteria
Akaike (AIC)  142199.436
Bayesian (BIC)  142468.382
Sample-Size Adjusted BIC  142363.509
(n* = (n + 2) / 24)

MODEL RESULTS
### APPENDIX C (continued)

<table>
<thead>
<tr>
<th>Two-Tailed Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within Level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.707</td>
<td>0.087</td>
<td>8.089</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.756</td>
<td>0.107</td>
<td>-25.643</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.255</td>
<td>0.086</td>
<td>-26.108</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.744</td>
<td>0.120</td>
<td>-6.184</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.284</td>
<td>0.054</td>
<td>23.630</td>
</tr>
<tr>
<td>X1X2</td>
<td>-0.024</td>
<td>0.067</td>
<td>-0.356</td>
</tr>
<tr>
<td>X1X3</td>
<td>-0.202</td>
<td>0.099</td>
<td>-2.034</td>
</tr>
<tr>
<td>CENTT ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.102</td>
<td>0.004</td>
<td>23.926</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.021</td>
<td>0.008</td>
<td>2.645</td>
</tr>
<tr>
<td>LEP</td>
<td>0.023</td>
<td>0.017</td>
<td>1.315</td>
</tr>
<tr>
<td>SPED</td>
<td>0.015</td>
<td>0.009</td>
<td>1.735</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.044</td>
<td>0.007</td>
<td>6.389</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.020</td>
<td>0.008</td>
<td>2.655</td>
</tr>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTT</td>
<td>-0.062</td>
<td>0.008</td>
<td>-8.170</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>7.736</td>
<td>0.277</td>
<td>27.977</td>
</tr>
<tr>
<td>CENTT</td>
<td>0.236</td>
<td>0.005</td>
<td>43.572</td>
</tr>
<tr>
<td><strong>Between Level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGM_T</td>
<td>-0.154</td>
<td>0.309</td>
<td>-0.497</td>
</tr>
<tr>
<td>S2 ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGM_T</td>
<td>0.333</td>
<td>0.448</td>
<td>0.742</td>
</tr>
<tr>
<td>S3 ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGM_T</td>
<td>-1.845</td>
<td>0.681</td>
<td>-2.707</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>-0.745</td>
<td>0.956</td>
<td>-0.778</td>
</tr>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RECGMPOP</td>
<td>0.055</td>
<td>0.013</td>
<td>4.396</td>
</tr>
<tr>
<td>CGMMOB</td>
<td>-3.617</td>
<td>0.851</td>
<td>-4.252</td>
</tr>
<tr>
<td>CGMFRL</td>
<td>-8.000</td>
<td>0.968</td>
<td>-8.261</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>1.014</td>
<td>0.978</td>
<td>1.036</td>
</tr>
<tr>
<td>CGM_T</td>
<td>5.736</td>
<td>1.096</td>
<td>5.234</td>
</tr>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>16.606</td>
<td>0.221</td>
<td>75.033</td>
</tr>
<tr>
<td>S1</td>
<td>0.330</td>
<td>0.077</td>
<td>4.261</td>
</tr>
<tr>
<td>S2</td>
<td>-0.092</td>
<td>0.052</td>
<td>-1.754</td>
</tr>
<tr>
<td>S3</td>
<td>-0.637</td>
<td>0.224</td>
<td>-2.845</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>0.584</td>
<td>0.129</td>
<td>4.513</td>
</tr>
</tbody>
</table>
3e. Climate measure: Expectations

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood

H0 Value -73924.914
H0 Scaling Correction Factor 1.958
for MLR

Information Criteria

Akaike (AIC) 147915.827
Bayesian (BIC) 148184.491
Sample-Size Adjusted BIC 148079.617
(n* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.716</td>
<td>0.089</td>
<td>8.062</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.759</td>
<td>0.111</td>
<td>-24.951</td>
</tr>
<tr>
<td>SPED</td>
<td>-2.256</td>
<td>0.086</td>
<td>-26.153</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.744</td>
<td>0.120</td>
<td>-6.203</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.268</td>
<td>0.052</td>
<td>24.537</td>
</tr>
<tr>
<td>X1X2</td>
<td>-0.091</td>
<td>0.070</td>
<td>-1.301</td>
</tr>
<tr>
<td>X1X3</td>
<td>-0.196</td>
<td>0.106</td>
<td>-1.853</td>
</tr>
<tr>
<td>CENTE ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.066</td>
<td>0.006</td>
<td>10.981</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.000</td>
<td>0.010</td>
<td>-0.023</td>
</tr>
<tr>
<td>LEP</td>
<td>0.037</td>
<td>0.020</td>
<td>1.832</td>
</tr>
<tr>
<td>SPED</td>
<td>-0.058</td>
<td>0.012</td>
<td>-4.959</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.021</td>
<td>0.008</td>
<td>2.593</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.022</td>
<td>0.009</td>
<td>2.411</td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTE</td>
<td>-0.023</td>
<td>0.008</td>
<td>-2.811</td>
</tr>
<tr>
<td>Residual Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_E</td>
<td>7.740</td>
<td>0.276</td>
<td>28.080</td>
</tr>
<tr>
<td>CENTE</td>
<td>0.309</td>
<td>0.005</td>
<td>59.096</td>
</tr>
</tbody>
</table>
### Between Level

<table>
<thead>
<tr>
<th></th>
<th>PLAN_E ON</th>
<th>S1 ON</th>
<th>S2 ON</th>
<th>S3 ON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CGM_E</td>
<td>CGM_E</td>
<td>CGM_E</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.141</td>
<td>0.152</td>
<td>-0.926</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.393</td>
<td>0.275</td>
<td>1.433</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.062</td>
<td>0.404</td>
<td>-2.625</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.428</td>
<td>0.943</td>
<td>-0.455</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.428</td>
<td>0.914</td>
<td>0.791</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.240</td>
<td>0.568</td>
<td>7.468</td>
</tr>
</tbody>
</table>

### PLAN_E ON

<table>
<thead>
<tr>
<th></th>
<th>RECGMPOP</th>
<th>CGMMOB</th>
<th>CGMFRL</th>
<th>CGMRDI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.049</td>
<td>-3.989</td>
<td>-6.624</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>0.792</td>
<td>0.810</td>
<td>0.914</td>
</tr>
<tr>
<td></td>
<td>4.017</td>
<td>-5.035</td>
<td>-8.175</td>
<td>0.791</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.429</td>
</tr>
</tbody>
</table>

### Intercepts

<table>
<thead>
<tr>
<th></th>
<th>PLAN_E</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16.618</td>
<td>0.317</td>
<td>0.114</td>
<td>-0.668</td>
</tr>
<tr>
<td></td>
<td>0.206</td>
<td>0.082</td>
<td>0.052</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td>80.679</td>
<td>3.893</td>
<td>2.176</td>
<td>-3.006</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.030</td>
<td>0.003</td>
</tr>
</tbody>
</table>

### Residual Variances

<table>
<thead>
<tr>
<th></th>
<th>PLAN_E</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.474</td>
<td>0.006</td>
<td>0.015</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>0.122</td>
<td>0.026</td>
<td>0.017</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>3.867</td>
<td>0.213</td>
<td>0.901</td>
<td>2.466</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.831</td>
<td>0.368</td>
<td>0.014</td>
</tr>
</tbody>
</table>
PLAN_R: Models Predicting Reading Test Scores

4a. Climate measure: Safety

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood

H0 Value -78927.062
H0 Scaling Correction Factor 1.943 for MLR

Information Criteria

Akaike (AIC) 157920.125
Bayesian (BIC) 158189.072
Sample-Size Adjusted BIC 158084.199
(n* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_R ON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRADE</td>
<td>0.569</td>
<td>0.080</td>
<td>7.122</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.838</td>
<td>0.117</td>
<td>-24.316</td>
</tr>
<tr>
<td>SPED</td>
<td>-1.983</td>
<td>0.094</td>
<td>-21.201</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.670</td>
<td>0.110</td>
<td>-6.076</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.306</td>
<td>0.053</td>
<td>24.684</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.293</td>
<td>0.066</td>
<td>4.404</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.089</td>
<td>0.114</td>
<td>0.783</td>
</tr>
</tbody>
</table>

| CENTS ON | | | |
| CENTGPA | 0.017 | 0.006 | 2.608 | 0.009 |
| LUNCH | -0.025 | 0.008 | -3.030 | 0.002 |
| LEP | -0.234 | 0.020 | -11.490 | 0.000 |
| SPED | -0.060 | 0.012 | -4.913 | 0.000 |
| GENDER | -0.063 | 0.015 | -4.257 | 0.000 |
| RACETHRT | 0.016 | 0.011 | 1.456 | 0.145 |

| Intercepts | | | |
| CENTS | 0.060 | 0.014 | 4.287 | 0.000 |

| Residual Variances | | | |
| PLAN_R | 10.607 | 0.293 | 36.176 | 0.000 |
| CENTS | 0.334 | 0.009 | 38.081 | 0.000 |
### Between Level

<table>
<thead>
<tr>
<th></th>
<th>PLAN_R</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>ON CGM_S</td>
<td>-0.138</td>
<td>0.167</td>
<td>-0.825</td>
</tr>
<tr>
<td>S2</td>
<td>ON CGM_S</td>
<td>0.369</td>
<td>0.173</td>
<td>2.132</td>
</tr>
<tr>
<td>S3</td>
<td>ON CGM_S</td>
<td>0.240</td>
<td>0.620</td>
<td>0.387</td>
</tr>
<tr>
<td></td>
<td>CGMRDI</td>
<td>0.071</td>
<td>1.202</td>
<td>0.059</td>
</tr>
<tr>
<td>PLAN_R ON RECGMPOP</td>
<td>0.038</td>
<td>0.013</td>
<td>2.939</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>CGMMOB</td>
<td>-1.573</td>
<td>0.825</td>
<td>-1.906</td>
</tr>
<tr>
<td></td>
<td>CGMFRL</td>
<td>-7.322</td>
<td>1.061</td>
<td>-6.901</td>
</tr>
<tr>
<td></td>
<td>CGMRDI</td>
<td>0.429</td>
<td>1.124</td>
<td>0.382</td>
</tr>
<tr>
<td></td>
<td>CGM_S</td>
<td>2.639</td>
<td>0.868</td>
<td>3.040</td>
</tr>
</tbody>
</table>

### Intercepts

<table>
<thead>
<tr>
<th></th>
<th>PLAN_R</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_R</td>
<td>16.730</td>
<td>0.203</td>
<td>0.085</td>
<td>2.399</td>
</tr>
<tr>
<td>S1</td>
<td>0.163</td>
<td>0.056</td>
<td>2.928</td>
<td>0.003</td>
</tr>
<tr>
<td>S2</td>
<td>0.509</td>
<td>0.209</td>
<td>2.441</td>
<td>0.015</td>
</tr>
</tbody>
</table>

### Residual Variances

<table>
<thead>
<tr>
<th></th>
<th>PLAN_R</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_R</td>
<td>0.634</td>
<td>0.014</td>
<td>0.023</td>
<td>0.612</td>
</tr>
<tr>
<td>S1</td>
<td>0.009</td>
<td>0.020</td>
<td>0.477</td>
<td>0.633</td>
</tr>
<tr>
<td>S2</td>
<td>0.307</td>
<td>0.129</td>
<td>2.375</td>
<td>0.018</td>
</tr>
</tbody>
</table>

### 4b. Climate measure: Antisocial Peers

#### MODEL FIT INFORMATION

- Number of Free Parameters: 33
- Loglikelihood:
  - H0 Value: -73168.019
  - H0 Scaling Correction Factor: 1.882 for MLR
- Information Criteria:
  - Akaike (AIC): 146402.038
  - Bayesian (BIC): 146670.985
  - Sample-Size Adjusted BIC: 146566.112
  - n* = (n + 2) / 24

#### MODEL RESULTS
### APPENDIX C (continued)

Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
</table>

**Within Level**

<table>
<thead>
<tr>
<th>PLAN_R ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GRADE</td>
<td>0.581</td>
<td>0.078</td>
<td>7.409</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.870</td>
<td>0.116</td>
<td>-24.796</td>
</tr>
<tr>
<td>SPED</td>
<td>-1.978</td>
<td>0.093</td>
<td>-21.226</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.677</td>
<td>0.110</td>
<td>-6.186</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.312</td>
<td>0.054</td>
<td>24.419</td>
</tr>
<tr>
<td>X1X2</td>
<td>0.126</td>
<td>0.089</td>
<td>1.415</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.093</td>
<td>0.119</td>
<td>0.789</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CENTA ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTGPA</td>
<td>0.002</td>
<td>0.006</td>
<td>0.410</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.002</td>
<td>0.010</td>
<td>0.228</td>
</tr>
<tr>
<td>LEP</td>
<td>0.118</td>
<td>0.015</td>
<td>7.800</td>
</tr>
<tr>
<td>SPED</td>
<td>0.084</td>
<td>0.012</td>
<td>7.216</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.093</td>
<td>0.013</td>
<td>-7.164</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.006</td>
<td>0.010</td>
<td>0.605</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intercepts</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTA</td>
<td>0.028</td>
<td>0.012</td>
<td>2.294</td>
</tr>
</tbody>
</table>

**Residual Variances**

| PLAN_R | 10.622 | 0.292 | 36.387 | 0.000 |
| CENTA  | 0.212  | 0.004 | 57.215 | 0.000 |

**Between Level**

<table>
<thead>
<tr>
<th>S1 ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>0.056</td>
<td>0.215</td>
<td>0.262</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2 ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>0.827</td>
<td>0.219</td>
<td>3.773</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3 ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM_A</td>
<td>0.182</td>
<td>0.791</td>
<td>0.230</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>-0.156</td>
<td>1.134</td>
<td>-0.137</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PLAN_R ON</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RECGMPOP</td>
<td>0.030</td>
<td>0.013</td>
<td>2.238</td>
</tr>
<tr>
<td>CGMMOB</td>
<td>-3.063</td>
<td>0.929</td>
<td>-3.298</td>
</tr>
<tr>
<td>CGMFRSL</td>
<td>-7.798</td>
<td>1.082</td>
<td>-7.208</td>
</tr>
<tr>
<td>CGMRDI</td>
<td>0.120</td>
<td>1.078</td>
<td>0.111</td>
</tr>
<tr>
<td>CGM_A</td>
<td>2.333</td>
<td>1.101</td>
<td>2.119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intercepts</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_R</td>
<td>16.708</td>
<td>0.220</td>
<td>76.041</td>
</tr>
<tr>
<td>S1</td>
<td>0.150</td>
<td>0.084</td>
<td>1.779</td>
</tr>
<tr>
<td>S2</td>
<td>-0.321</td>
<td>0.069</td>
<td>-4.627</td>
</tr>
<tr>
<td>S3</td>
<td>-0.486</td>
<td>0.204</td>
<td>-2.382</td>
</tr>
</tbody>
</table>

**Residual Variances**
### APPENDIX C (continued)

PLAN_R  | 0.693 | 0.143 | 4.864 | 0.000  
S1      | 0.012 | 0.023 | 0.521 | 0.603  
S2      | 0.017 | 0.028 | 0.599 | 0.549  
S3      | 0.363 | 0.140 | 2.600 | 0.009  

4c. Climate measure: Prosocial Peers

MODEL FIT INFORMATION

Number of Free Parameters 33

Loglikelihood

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H0 Value</td>
<td>-75525.351</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0 Scaling Correction Factor</td>
<td>1.805</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

for MLR

Information Criteria

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Akaike (AIC)</td>
<td>151116.702</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian (BIC)</td>
<td>151385.649</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample-Size Adjusted BIC</td>
<td>151280.776</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n* = (n + 2) / 24)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MODEL RESULTS

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two-Tailed</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>S.E. Est./S.E.</td>
<td>P-Value</td>
</tr>
</tbody>
</table>

Within Level

PLAN_R ON

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GRADE</td>
<td>0.589</td>
<td>0.078</td>
<td>7.510</td>
</tr>
<tr>
<td>LEP</td>
<td>-2.787</td>
<td>0.115</td>
<td>-24.338</td>
</tr>
<tr>
<td>SPED</td>
<td>-1.919</td>
<td>0.091</td>
<td>-20.970</td>
</tr>
<tr>
<td>LUNCH</td>
<td>-0.647</td>
<td>0.106</td>
<td>-6.089</td>
</tr>
<tr>
<td>CENTGPA</td>
<td>1.308</td>
<td>0.053</td>
<td>24.630</td>
</tr>
<tr>
<td>X1X2</td>
<td>-0.019</td>
<td>0.083</td>
<td>-0.223</td>
</tr>
<tr>
<td>X1X3</td>
<td>0.079</td>
<td>0.104</td>
<td>0.758</td>
</tr>
</tbody>
</table>

CENTP ON

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTGPA</td>
<td>-0.013</td>
<td>0.006</td>
<td>-2.219</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.033</td>
<td>0.008</td>
<td>4.386</td>
</tr>
<tr>
<td>LEP</td>
<td>0.168</td>
<td>0.016</td>
<td>10.633</td>
</tr>
<tr>
<td>SPED</td>
<td>0.127</td>
<td>0.014</td>
<td>8.863</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.030</td>
<td>0.009</td>
<td>-3.198</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.008</td>
<td>0.009</td>
<td>0.958</td>
</tr>
</tbody>
</table>

Intercepts

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CENTP</td>
<td>-0.042</td>
<td>0.009</td>
<td>-4.777</td>
</tr>
</tbody>
</table>

Residual Variances

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAN_R</td>
<td>10.530</td>
<td>0.291</td>
<td>36.212</td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

CENTP              0.257      0.006     44.825      0.000

Between Level
S1       ON
  CGM_P   -0.046      0.218   -0.209      0.834
S2       ON
  CGM_P             0.858      0.303   2.834      0.005
S3       ON
  CGM_P   -0.232      0.846   -0.274      0.784
  CGMRDI         0.080      1.091   0.074      0.941

PLAN_R ON
  RECGMPOP       0.037      0.014   2.684      0.007
  CGMPOP         -3.707     0.970   -3.820      0.000
  CGMFRL         -7.749     1.047   -7.401      0.000
  CGMRDI         0.249      1.028   0.243      0.808
  CGM_P           3.069     1.244   2.467      0.014

Intercepts
PLAN_R       16.656     0.215  77.305      0.000
S1           0.173      0.073   2.363      0.018
S2 -0.569      0.065  -8.719      0.000
S3 -0.490      0.207  -2.367      0.018

Residual Variances
PLAN_R       0.655      0.148   4.412      0.000
S1           0.010      0.023   0.429      0.668
S2           0.072      0.035   2.040      0.041
S3           0.395      0.141   2.797      0.005

4d. Climate measure: Teacher-Student Interactions

MODEL FIT INFORMATION

Number of Free Parameters                           33

Loglikelihood
    H0 Value                        -74509.811
    H0 Scaling Correction Factor    1.665
    for MLR

Information Criteria
    Akaike (AIC)                    149085.621
    Bayesian (BIC)                  149354.567
    Sample-Size Adjusted BIC        149249.694
    (n* = (n + 2) / 24)

MODEL RESULTS
## Two-Tailed

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
</table>

### Within Level

**PLAN_R ON**
- GRADE: 0.577 0.080 7.247 0.000
- LEP: -2.894 0.115 -25.200 0.000
- SPED: -2.003 0.093 -21.432 0.000
- LUNCH: -0.681 0.111 -6.136 0.000
- CENTGPA: 1.323 0.055 24.074 0.000
- X1X2: 0.057 0.092 0.619 0.536
- X1X3: 0.099 0.098 1.011 0.312

**CENTT ON**
- CENTGPA: 0.102 0.004 23.926 0.000
- LUNCH: 0.021 0.008 2.645 0.008
- LEP: 0.023 0.017 1.315 0.188
- SPED: 0.015 0.009 1.735 0.083
- GENDER: 0.044 0.007 6.389 0.000
- RACETHRT: 0.020 0.008 2.655 0.008

### Intercepts
- CENTT: -0.062 0.008 -8.170 0.000

### Residual Variances
- PLAN_R: 10.646 0.292 36.425 0.000
- CENTT: 0.236 0.005 43.572 0.000

### Between Level

**S1 ON**
- CGM_T: -0.491 0.465 -1.055 0.291

**S2 ON**
- CGM_T: 0.980 0.447 2.192 0.028

**S3 ON**
- CGM_T: -1.100 0.740 -1.487 0.137
- CGMRDI: 0.197 0.969 0.203 0.839

**PLAN_R ON**
- RECGMPOP: 0.047 0.013 3.736 0.000
- CGMMOB: -3.108 0.897 -3.466 0.001
- CGMFRL: -9.087 0.930 -9.772 0.000
- CGMRDI: 0.284 0.920 0.308 0.758
- CGM_T: 5.536 1.096 5.053 0.000

### Intercepts
- PLAN_R: 16.753 0.201 83.324 0.000
- S1: 0.181 0.072 2.521 0.012
- S2: -0.087 0.067 -1.291 0.197
- S3: -0.529 0.197 -2.682 0.007

### Residual Variances
### 4e. Climate measure: Expectations

**MODEL FIT INFORMATION**

Number of Free Parameters: 33

Loglikelihood

- H0 Value: -77324.574
- H0 Scaling Correction Factor: 1.605 for MLR

Information Criteria

- Akaike (AIC): 154715.147
- Bayesian (BIC): 154983.811
- Sample-Size Adjusted BIC: 154878.938 (n* = (n + 2) / 24)

**MODEL RESULTS**

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN_R ON GRADE</td>
<td>0.579</td>
<td>0.080</td>
<td>7.275</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>-2.897</td>
<td>0.115</td>
<td>-25.204</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>-1.995</td>
<td>0.093</td>
<td>-21.429</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>-0.690</td>
<td>0.109</td>
<td>-6.347</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>1.318</td>
<td>0.053</td>
<td>24.920</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>-0.164</td>
<td>0.098</td>
<td>-1.677</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.102</td>
<td>0.941</td>
<td>0.347</td>
</tr>
<tr>
<td><strong>CENTE ON</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTGPA</td>
<td>0.066</td>
<td>0.006</td>
<td>10.981</td>
<td>0.000</td>
</tr>
<tr>
<td>LUNCH</td>
<td>0.000</td>
<td>0.010</td>
<td>-0.023</td>
<td>0.982</td>
</tr>
<tr>
<td>LEP</td>
<td>0.037</td>
<td>0.020</td>
<td>1.832</td>
<td>0.067</td>
</tr>
<tr>
<td>SPED</td>
<td>-0.058</td>
<td>0.012</td>
<td>-4.959</td>
<td>0.000</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.021</td>
<td>0.008</td>
<td>2.593</td>
<td>0.010</td>
</tr>
<tr>
<td>RACETHRT</td>
<td>0.022</td>
<td>0.009</td>
<td>2.411</td>
<td>0.016</td>
</tr>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTE</td>
<td>-0.023</td>
<td>0.008</td>
<td>-2.811</td>
<td>0.005</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>PLAN_R</td>
<td>10.632</td>
<td>0.290</td>
<td>36.674</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>CENTE</td>
<td>0.309</td>
<td>0.005</td>
<td>59.096</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Between Level**

<table>
<thead>
<tr>
<th>Level</th>
<th>PLAN_R</th>
<th>CENTE</th>
<th>Between Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>ON</td>
<td>CGM_E</td>
<td>-0.276</td>
</tr>
<tr>
<td>S2</td>
<td>ON</td>
<td>CGM_E</td>
<td>0.498</td>
</tr>
<tr>
<td>S3</td>
<td>ON</td>
<td>CGM_E</td>
<td>-0.690</td>
</tr>
<tr>
<td></td>
<td>CGMRDI</td>
<td></td>
<td>0.406</td>
</tr>
</tbody>
</table>

**PLAN_R ON**

<table>
<thead>
<tr>
<th>Variable</th>
<th>PLAN_R</th>
<th>RECGMPOP</th>
<th>0.040</th>
<th>0.011</th>
<th>3.690</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.182</td>
<td>0.076</td>
<td>2.383</td>
<td>0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>0.172</td>
<td>0.066</td>
<td>2.631</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>-0.538</td>
<td>0.186</td>
<td>-2.898</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Intercepts**

<table>
<thead>
<tr>
<th>Variable</th>
<th>PLAN_R</th>
<th>16.749</th>
<th>0.176</th>
<th>95.066</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.182</td>
<td>0.076</td>
<td>2.383</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>0.172</td>
<td>0.066</td>
<td>2.631</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>-0.538</td>
<td>0.186</td>
<td>-2.898</td>
<td>0.004</td>
<td></td>
</tr>
</tbody>
</table>

**Residual Variances**

<table>
<thead>
<tr>
<th>Variable</th>
<th>PLAN_R</th>
<th>0.466</th>
<th>0.164</th>
<th>2.843</th>
<th>0.004</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.011</td>
<td>0.025</td>
<td>0.454</td>
<td>0.650</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>0.018</td>
<td>0.026</td>
<td>0.715</td>
<td>0.475</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>0.424</td>
<td>0.162</td>
<td>2.610</td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>
Rebecca D. Taylor
Department of Psychology
University of Illinois at Chicago
1007 West Harrison Street (M/C 285)
Chicago, IL 60607-7137

EDUCATION

2011 Ph.D., Psychology (Community & Prevention Research), University of Illinois at Chicago
Minor: Statistics, Methods, and Measurement
Dissertation: The Multilevel Effects of School Climate and Gender on Academic Achievement in Urban High School Students
Committee: Roger P. Weissberg (Chair), Herbert Walberg, Mary Murphy, Stephanie Riger, & Kimberly Kendziora

2008 M.A., Psychology (Community & Prevention Research), University of Illinois at Chicago
Thesis: The Follow-up Effects of School-based Social and Emotional Learning Interventions
Committee: Roger P. Weissberg (Chair), Joseph A. Durlak, & Susan R. Goldman

2002 B.A., Cum Laude Psychology, Pomona College, Claremont, CA
Primary Advisor: Deborah Burke

MAJOR INTERESTS
Social and emotional competencies in young people; prevention of adolescent risk-taking behaviors through skill building; positive youth development; school and classroom climate assessment; multiple methodologies and multilevel analytical approaches to studying individuals nested in settings; ecological evaluation of schools and community organizations; impact of stereotype threat in learning and academic performance; development of LGBTQ young people; intersection of research and public policy.

ACADEMIC HONORS AND FELLOWSHIPS

Aug 2010-Aug 2011
University Fellowship, UIC Graduate College. Received a renewal of the $20,000 University Fellowship stipend and a tuition waiver (worth approximately $11,000) to support research in dissertation year.

August 2005-August 2009
Jacob K Javits Fellowship, US Department of Education. Selected by panel of psychology experts appointed by the Javits Fellowship Board as a promising early scholar. Over four years, received $120,000 in stipends and four tuition waivers (worth approximately $44,000) to support doctoral work at UIC.
Aug 2004-August 2005
University Fellowship, UIC Graduate College. Selected as a University Fellow and awarded a $20,000 stipend and a tuition waiver (worth approximately $11,000) to recruit me to UIC and to support my incoming year of graduate study.

May 2002
Phi Beta Kappa, Pomona College. Inducted into Phi Beta Kappa honor society.

PUBLICATIONS


In Preparation:


TECHNICAL AND POLICY REPORTS


PRESENTATIONS


Tran, N., Taylor, R. D., & Dymnicki, A. B. (2005, September). *School-based action research: The challenge of collaborating to create change.* Symposium conducted at the annual meeting of the Midwest ECO, Saugatuck, MI.


RESEARCH EXPERIENCE

Sept 2004 – Present  
*Research Assistant, Collaborative for Academic, Social and Emotional Learning (CASEL), Chicago, IL*

*Meta-analytic Research*
*Collected and coded studies for use in a meta-analysis of over 400 studies on the impact of universal, targeted, and after-school social and emotional learning (SEL) programs.*

*Contributed to the conceptualization of research questions and the design of the analytic plan for the meta-analytic reviews during regular research group meetings.*

*Responsible for organizing and analyzing meta-analytic data, and reporting post-intervention and follow-up results.*

*SEL and Climate Assessment Research*
*Collaborated with experienced teachers and other school personnel to generate items for use in a formative assessment to evaluate and promote students social and emotional competence. Focused on students at the middle school-level.*
*Assisted in the organization of the above items and others into a formative assessment tool for use in kindergarten through 12th grade, under the guidance of Robert Marzano
*Attended and documented ongoing meetings of an expert panel on the assessment of social and emotional competencies, with the goal of reviewing the available tools and state of the field, as well as suggesting future directions. Assisted in the review of a measure being developed by this panel.
* Assisted in the analysis of a school climate assessment developed by CASEL staff.

Oct 2006 – June 2007  
Community Practicum/Internship, Broadway Youth Center/Howard Brown Health Center, Chicago, IL
* Conducted ecological evaluation of agency and developed organizational profile
* Researched mentoring and mental health issues for gay, lesbian, bisexual, transgender, queer and questioning (GLBTQQ) youth
*Wrote foundational documents for use in multiple grant proposals for mentor program
*Assisted in the training of mentors for GLBTQQ youth

May – Nov 2005  
Research Assistant, Illinois State Board of Education (ISBE), Chicago, IL
* Surveyed the social and emotional learning policies for every school district in Illinois
*Wrote a policy paper summarizing the survey and its implications

TEACHING EXPERIENCE

Course Instructor
Statistical Methods in Behavioral Sciences; Dept. of Psychology, UIC (Spring 2008)

Teaching Assistantships
Introduction to Research Methods; Dept. of Psychology, UIC (Fall 2009)
Developmental Psychology; Dept. of Psychology, UIC (Fall 2004)

Invited Guest Lectures
Introduction to Research Methods in Psychology
   Topic: Qualitative Research Methods (Fall 2009; for 2 instructors)


Teaching portfolio materials available for request: Student Evaluations, Student Comments, Syllabus, Sample Lecture Slides and Notes, Sample Class Activity, Sample Exam