Abstract Title Page

Title: Do Intervention Impacts on Academic Achievement Vary by School Climate? Evidence from a Randomized Trial in Urban Elementary Schools

Authors and Affiliations:

Meghan P. McCormick
New York University

Elise Cappella
New York University

Erin E. O'Connor
New York University

Sandee G. McClowry
New York University

Abstract Body

Background: Given established links between social-emotional skills and academic achievement, there is growing support for implementing universal social/behavioral interventions in early schooling (Jones & Bouffard, 2012). Advocates have been particularly interested in implementing such programming in low-income urban schools where students are likely to start school with lower levels of social-emotional and academic skills than their more affluent peers (Jones & Bouffard, 2012; Raver, 2002). Yet, there is inconsistent evidence that such programs improve students' academic achievement over and above typical educational practice (SRCDC, 2010). One possible constraint to understanding mixed evidence about intervention efficacy is the limited information on how program effects differ across school settings. It could be that universal social/behavioral programs are highly effective in some types of schools and less so in others, thus confounding overall understanding of intervention efficacy. Moreover, although some work has considered how demographic characteristics – like school poverty – differentiate social/behavioral program impacts on student outcomes, fewer studies have examined the moderating role of the school-level social processes (e.g., social norms, relationships) within which interventions are typically embedded.

School climate reflects the norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures of the school (NSCC, 2007; Thapa et al., 2013), and is a useful construct for understanding social processes at the school-level. A prevention research perspective suggests that schools with the poorest climates have the most to gain from school-based interventions that explicitly target social interactions (e.g., Cicchetti & Aber, 1998; Van Lier et al., 2004). Contrasting work argues that social/behavioral programs will be most effective for improving student outcomes in settings where extant norms *already* support positive academic and social-behavioral development (Aber et al., 1998; Hughes et al., 2005).

Purpose of Study: The current study considers these alternating views by examining whether key theory-driven domains of school climate – leadership, accountability, and safety/respect (see Nathanson et al., 2013) – moderate impacts of one social/behavioral intervention – *INSIGHTS* – on low-income urban kindergarten and first grade students' math and reading achievement, sustained attention, and disruptive behaviors. A randomized trial using intent-to-treat analyses identified empirical support for benefits of *INSIGHTS* on these four student outcomes (see O'Connor et al., 2014). It is unclear, however, whether students in different schools benefited similarly. Learning more about variation in *INSIGHTS*' impacts across schools can inform targeting of social/behavioral programs, and allocation of funds towards the types of schools and students that may benefit most from such intervention.

Setting: This study took place in 22 low-income urban public elementary schools. All classrooms were regular education, with an average of 16.57 students (SD = 3.54). Schools had an average attendance rate of 86.26% (SD = .19) and an average size of 465 students (SD = 158.46). Schools also had high percentages of students who were racial/ethnic minorities (Black, M = .77, SD = .13; Hispanic, M = .40, SD = .27) and eligible for free or reduced lunch (M = .80, SD = .16).

Participants: Ninety-one percent of participating children were age five or six when they enrolled in the study (M = 5.38 SD = 0.61). Half (52%) of the children were male. Eighty-seven percent of children qualified for free or reduced lunch. Seventy-five percent of children were black, non-Hispanic, 16% were Hispanic, non-black, and the remaining children were biracial. 22% of students' parents graduated from a 2- or 4-year college.

Intervention: *INSIGHTS* is a comprehensive social/behavioral intervention with teacher, parent, and classroom programs. In brief, *INSIGHTS* provides teachers and parents with a temperament framework for supporting the individual differences of children. Key to temperament theory is the concept of goodness of fit, or notion that it is important for a child's temperament to be in consonance with the demands,

expectations, and opportunities of the child's proximal environment. Although temperament itself should not be targeted by intervention, the environment can be modified to appropriately respond to a child's temperament and improve goodness of fit. Using this framework, *INSIGHTS* helps parents and teachers recognize a child's temperament and respond with warmth and discipline strategies that support adaptive social-emotional and behavioral outcomes (McClowry et al., 2005; McClowry et al., 2010). *INSIGHTS*' ultimate goal is to the support students' academic development.

The *INSIGHTS* intervention implemented in this study included: (a) teacher sessions, (b) parent sessions, and (c) universal classroom sessions (see McClowry et al., 2010). Teachers and parents attended 10 weekly 2-hr facilitated sessions based on a structured curriculum that included didactic content and professionally produced vignettes as well as handouts and group activities. During the same 10 weeks, the classroom program was delivered in 45-min lessons to all students in the classrooms of participating teachers. Teachers were engaged in the children's sessions, especially when students practiced resolving dilemmas.

*Intervention fidelity***.** Facilitators followed scripts, used material checklists, documented sessions, and received ongoing training and supervision. Deviations or clinical concerns were discussed weekly in meetings with the program developer. Fidelity coding revealed that 93% of the curriculum was adequately covered, on average.

INSIGHTS dosage. The average number of teacher sessions attended was 9.44 (*SD* = 0.91). The average number of classroom sessions attended by the participating children was 8.30 (*SD* = 2.25). There was significant variation in parent participation across schools; participation ranged from 23% of parents attending more than 80% of sessions to 66% attending more than 80% of sessions.

*Attention-control condition***.** Schools not assigned to *INSIGHTS* participated in a 10-week, supplemental reading program after school for children whose parents consented. Dosage and fidelity were both deemed adequate in the attention-control condition.

Research Design: Eleven schools were randomized to *INSIGHTS*; the remaining eleven schools were assigned to the attention-control condition. Half of the children were in the *INSIGHTS* program (n = 225); the remaining child participants (n = 210) were in the attention-control. Similarly, approximately half of teachers (n = 57) participated in the *INSIGHTS* program; the remaining teachers (n = 65) were enrolled in the attention-control. Examination of pretest variables suggests group equivalence except for children's reading scores, which favored the control group.

Data Collection: Data used in this study were multi-informant and longitudinal. Prior to the implementation of the study, data on school climate and school demographic characteristics were assessed. Time 1 intervention (T1) data were then collected at baseline in the winter of the kindergarten year prior to students and classrooms receiving 10 weeks of kindergarten intervention. Time 2 (T2) data were collected following intervention in the late spring of the kindergarten year. Time 3 (T3) data were collected in the fall of first grade prior to 10 weeks of first grade intervention. Time 4 (T4) data were collected after the first grade intervention in the late winter of the first grade year, followed by Time 5 (T5) data in late spring. Importantly, the main variable of interest -- Treatment -- was measured as a dummy variable in all analyses, wherein 1 = *INSIGHTS*; 0 = attention-control.

Outcome variables. Outcome variables mirror the ones assessed in the study examining intent-to-treat effects of *INSIGHTS*. Reading and math achievement were assessed using raw scores from the Letter-Word Identification and Applied Problems subtests of the Woodcock-Johnson III Tests of Achievement, Form B (WJ-III; Woodcock McGrew, & Mather, 2001). Child sustained attention was measured with the Attention Sustained subtest from the Leiter International Performance Scale—Revised (Roid & Miller, 1997). Child behavior problems were measured with the 36-item Sutter–Eyberg Student Behavior Inventory, the teacher-report version of the Eyberg Child Behavior Inventory (Eyberg & Pincus, 1999; α = .97).

*Moderators***.** The school climate moderators – leadership, accountability, and safety/respect – were measured using aggregated reports from all teachers in the school on the New York City School Survey (see

http://schools.nyc.gov/Accountability/tools/ survey/default.htm). Leadership was measured using aggregated teacher reports of the quality of instructional leadership provided by principals and other administrative staff at their school. Accountability was measured by aggregating teacher perceptions on the extent to which the school had high standards for student work (Childress et al., 2011). Safety and respect measured the extent to which teachers felt that their school provided students and themselves with physical and emotional safety. Scores were measured on a 1 – 4 scale. A recent report provides evidence for reliability and validity of these measures (see Nathanson et al., 2013).

Covariates. Covariates included school demographic characteristics, child demographic characteristics, and child temperament (see O'Connor et al., 2014).

Data Analysis: There was 0% to 20% missing data across study variables. Twenty separate datasets were thus imputed by chained equations, using STATA MICE in STATA version 12. STATA ran each set of analyses 20 times and aggregated the findings across the imputed datasets.

*Validity of school climate constructs***.** The school climate dimensions used in this study – leadership, accountability, and safety/respect – have shown initial evidence of reliability and validity (see Nathanson et al, 2013; Rockoff & Speroni, 2008). In addition, the presented findings will demonstrate further conceptual and empirical evidence for the validity and reliability of constructs. Notably, correlations between school climate domains were moderate.

Growth curve modeling. Three-level individual growth modeling was used to examine change over four waves of data for each outcome (Singer & Willett, 2003). The metric of time used was time point. Time was centered at the last assessment time point so that the intercept would represent the average level of the outcome at the final intervention follow-up point (T5). The current study builds on the intent-to-treat analyses conducted by O'Connor et al. (2014) and thus follows a similar estimation strategy. As displayed in Appendix C, each student's outcome score at the intercept was modeled as a grand mean outcome score at the first post-treatment time point (T2) as well as a residual term that demonstrates deviations in outcome scores at T2 about the grand mean. Additionally, each student's rate of change across time on each outcome score was modeled as a grand mean rate of change in the outcome (β_1), as well as a residual term that demonstrates deviations in the slope (u_{1i}).

Moderated impacts. To examine moderated between-child impacts of *INSIGHTS* on student outcomes, cross-level interactions between treatment and the moderators of interest (leadership, accountability, safety/respect) were then added in separate models. Cross-level interactions between time, treatment, and the moderators of interest (leadership, accountability, safety/respect) were also included. Effect sizes were calculated for statistically significant impacts (see Feingold, 2009). Wald tests were used to determine whether moderated impacts were significantly different from one another.

Results: *Leadership moderation impacts*. Results from the leadership model (see Table 1) revealed significant within-child treatment x leadership effects on math (γ = -.26, *SE* = .11, *p* = .03; E.S. = .16) and reading achievement (γ = -.49, *SE* = .19, *p* = .03; E.S. = .21). Models examining within-child moderated effects on sustained attention and disruptive behaviors were not statistically significant. As illustrated in Figures 1a and 1b, findings suggest that impacts of *INSIGHTS* on math and reading achievement were larger in schools with lower baseline levels of leadership. There were no significant between-child moderated effects on any of the outcomes.

*Accountability moderation impacts***.** Results from the accountability model (see Table 2) revealed significant within-child treatment x accountability effects on math (γ = -.42, *SE* = .14, *p* = .02; E.S. = .25) and reading achievement (γ = -.70, *SE* = .26, *p* < .03; E.S. = .29). As illustrated in Figures 2a and 2b, findings suggest that impacts of *INSIGHTS* on math and reading achievement were larger in schools with lower baseline levels of leadership. In addition, between-child analyses revealed a significant treatment x accountability effect for disruptive behaviors (γ = .32, *SE* = .14, *p* = .04; E.S. = .27), suggesting that overall, schools with lower levels of accountability had bigger overall impacts on disruptive behaviors, relative to treatment schools with higher levels of baseline accountability. Models examining within-child moderated

effects on sustained attention and disruptive behaviors were not non-significant Moderated effects on math and reading achievement and sustained attention were also non-significant.

*Safety and respect moderation impacts***.** Results from the safety/respect model (see Table 3) revealed significant within-child treatment x safety/respect effects on math (γ = -.49, *SE* = .16, p = .02; E.S. = .29) and reading achievement (γ = -.86, *SE* = .29, $p < .01$; E.S. = .36) as well as sustained attention (γ = -1.33, *SE* = .36, $p < .01$; E.S. = .31). As illustrated in Figures 3a, 3b, and Figure 3c, findings suggest that impacts of *INSIGHTS* on math and reading achievement and sustained attention were larger in schools with lower baseline levels of safety/respect. In addition, between-child analyses revealed a significant treatment x safety/respect effect for reading achievement (γ = -3.24, *SE* = 1.46, p = .03; E.S. = .45), suggesting that overall, schools with lower levels of safety/respect had bigger overall impacts on disruptive behaviors, relative to treatment schools with higher levels of baseline safety/respect. Moderated within- and between-child effects on disruptive behaviors were not statistically significant. Additionally, there were no significant moderated between-child effects on math achievement and sustained attention.

*Sensitivity analyses***.** Wald tests revealed that the within-child moderated effects on math achievement were not statistically different in the models examining accountability and safety/respect. All other within-child moderated effects were significantly different from one another. In addition, the between-child moderated effect on disruptive behaviors for accountability was significantly different from the between-child moderated effect on reading achievement for safety/respect.

Conclusions: The current study is one of the first to consider the role of school climate in understanding moderated impacts of social/behavioral interventions on student achievement, attention, and behaviors. The major lesson from this work is that context matters. Across student outcomes, program impacts on achievement were generally larger, and sometimes driven by, schools that had less leadership, accountability and safety/respect prior to implementation of the intervention. Although there are nuanced reasons to explain heterogeneity of effects, future evaluators of social/behavioral programs should build on this work to determine whether such moderated impacts are replicated across diverse implementation settings. Similar to Bierman et al. (2010), it may be important to consider varied cities and types of school settings, while explicitly collecting data on school climate and other characteristics to later understand impact variation. In the past, researchers have tested many promising social/behavioral interventions in schools with relatively supportive climates. However, future research may benefit from targeting implementation at schools with more negative climates, given the current study's findings indicating enhanced effects for high-need schools.

Perhaps the biggest lesson from this study however is for policymakers, who are currently engaged in distributing funding to expand and implement social/behavioral interventions in a variety of settings across the country. Importantly, policymakers are paying increased attention to the role of school climate in promtoing student achievement (Weissbourd et al., 2013). Indeed, the U.S. Department of Education, the Institute for Educational Sciences, and President Obama's Bully Prevention Partnership endorse school climate renewal as a strategy for increasing student learning and achievement, and enhancing school connectedness. Given the supportive political context, it may be possible to combine assessment of school climate with efforts to implement social/behavioral interventions in high-need contexts. Indeed, findings suggest the importance of considering the overall climate and characteristics of the school before allocating resources to programs that aim to improve student achievement. Implementing such a strategy may actually be quite feasible in some larger urban areas where there are administrative surveys and outside quality reviews that provide information about school climate (Coburn et al, 2013). In doing so, it is critical that policymakers use such information to identify the schools that are most in need of school-based programming, rather than penalizing schools who are struggling to create supportive climates.

Appendix A. References

Aber, J. L., Jones, S. M., Brown, J. L., Chaudry, N., & Samples, F. (1998). Resolving conflict creatively: Evaluating the developmental effects of a school-based violence prevention program in neighborhood and classroom context. *Development and Psychopathology*, *10*(02), 187-213.

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, *78*(2), 156.

Childress, S., Higgins, M. C., Ishimaru, A., & Takahashi, S. (2011). Managing for results at the New York City Department of Education. Retrieved from: http://dash.harvard.edu/handle/1/9544460.

Cicchetti, D., & Aber, J. L. (1998). Contextualism and developmental psychopathology. *Development and Psychopathology*, *10*(02), 137-141.

Coburn, C., Penuel, W., & Geil, K. (2013). *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts*. Working Paper. New York: The William T. Grant Foundation.

Eyberg, S.M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory-Revised: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, *14*(1), 43.

Hughes, J. N., Cavell, T. A., Meehan, B. T., Zhang, D., & Collie, C. (2005). Adverse school context moderates the outcomes of selective interventions for aggressive children. *Journal of Consulting and Clinical Psychology*, *73*(4), 731.

Jones, S. M., & Bouffard, S. M. (2012). Social and Emotional Learning in Schools: From Programs to Strategies. Social Policy Report. Volume 26, Number 4. *Society for Research in Child Development*.

McClowry, S. G., Snow, D. L., & Tamis-LeMonda, C. S. (2005). An evaluation of the effects of INSIGHTS on the behavior of inner city primary school children. *The Journal of Primary Prevention*, *26*(6), 567-584.

McClowry, S. G., Snow, D. L., Tamis-LeMonda, C. S., & Rodriguez, E. T. (2010). Testing the efficacy of *INSIGHTS* on student disruptive behavior, classroom management, and student competence in inner city primary grades. *School Mental Health, 2,* 3–35.

Nathanson, L., McCormick, M., & Kemple, J. (2013). Strengthening assessments of school climate: Lessons from the NYC school survey. New York: The Research Alliance for New York City Schools. Retrieved from http://files.eric.ed.g ov/fulltext/ED543180.pdf.

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

O'Connor, E. E., Cappella, E., McCormick, M. P., & McClowry, S. G. (2014). An examination of the efficacy of INSIGHTS in enhancing the academic and behavioral development of children in early grades. *Journal of Educational Psychology*. Advance online publication. http://dx.doi.org/10.1037/a0036615.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report, 16*(3), 3-19.

Rockoff, J. & Speroni, C. (2008, November). Reliability, consistency, and validity of the NYC Environmental Surveys: A preliminary analysis. Working Paper. Columbia University Graduate School of Business.

Roid, G. H., & Miller, L. J. (1997). *Leiter international performance scale-revised: Examiners manual*. Wood Dale, IL: Stoelting.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Social and Character Development Research Consortium. (2010). Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children. *Washington*.

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

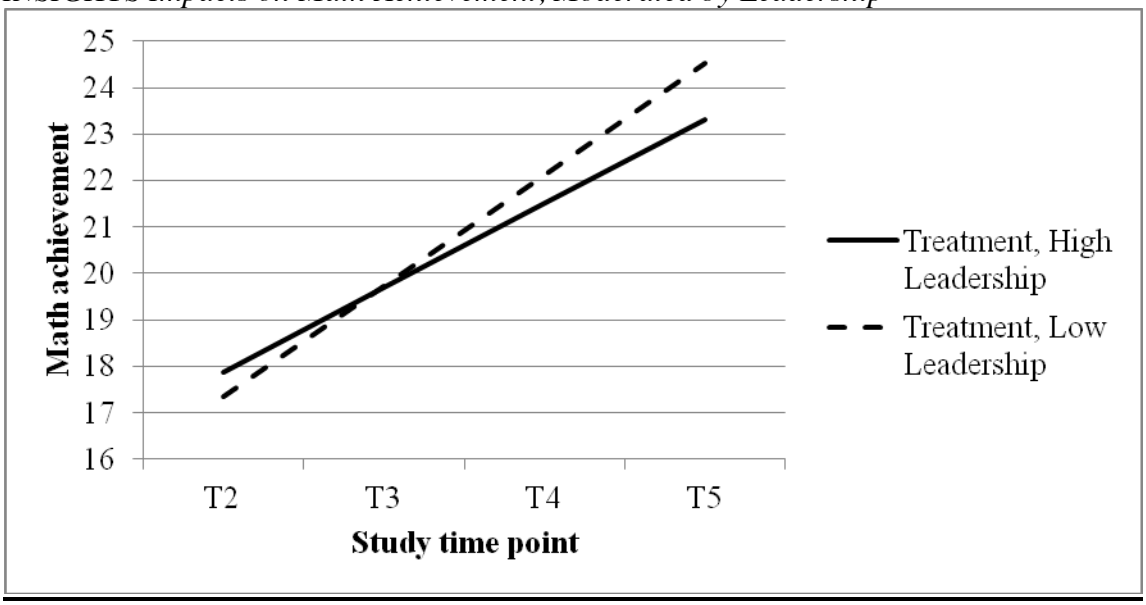
Van Lier, P. A., Muthén, B. O., van der Sar, R. M., & Crijnen, A. A. (2004). Preventing disruptive behavior in elementary schoolchildren: impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology*, *72*(3), 467.

Weissbourd, R., Bouffard, S. M., & Jones, S. M. (2013). School Climate and Moral and Social Development. *School Climate Practices for Implementation and Sustainability*, 30.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). Woodcock Johnson Psychoeducational Battery-Third Edition. Chicago: Riverside Publishing.

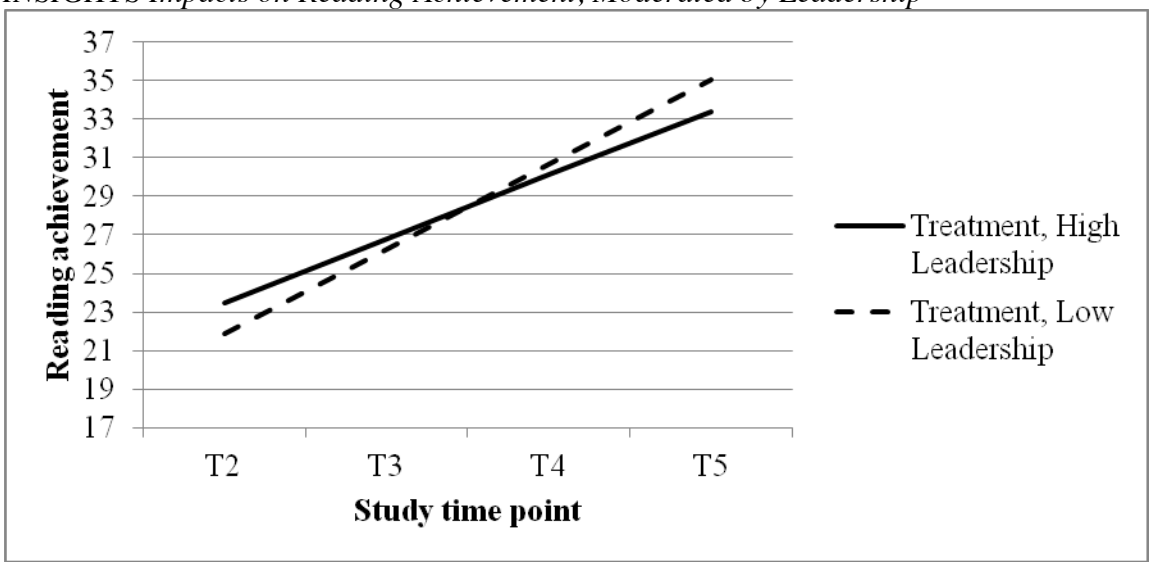
Appendix B. Tables and Figures

Figure 1a
INSIGHTS Impacts on Math Achievement, Moderated by Leadership



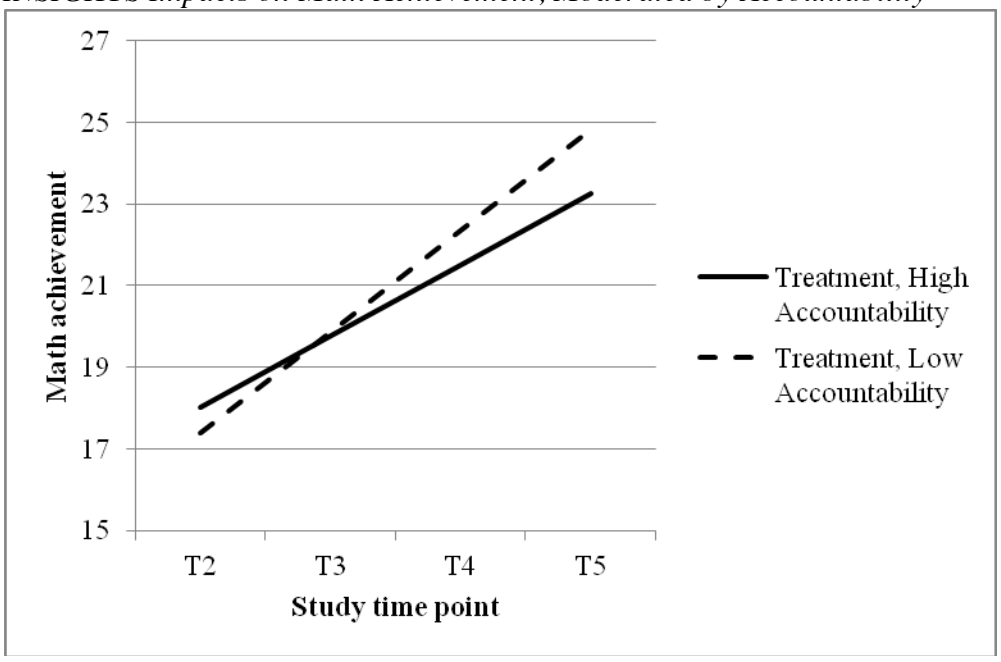
NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 1b
INSIGHTS Impacts on Reading Achievement, Moderated by Leadership



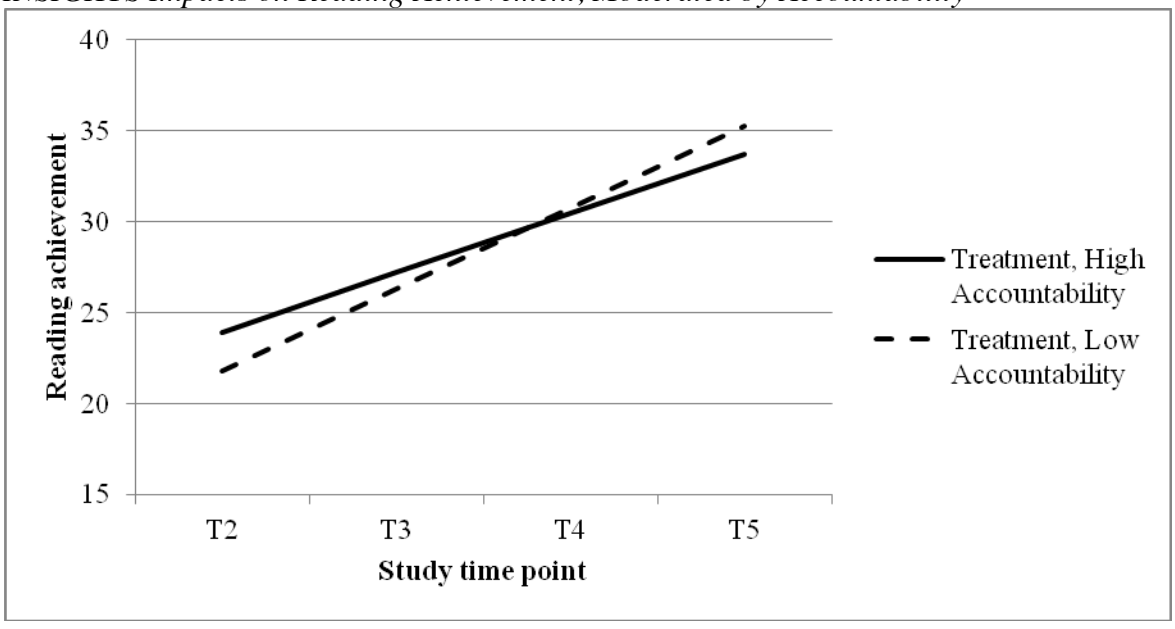
NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 2a
INSIGHTS Impacts on Math Achievement, Moderated by Accountability



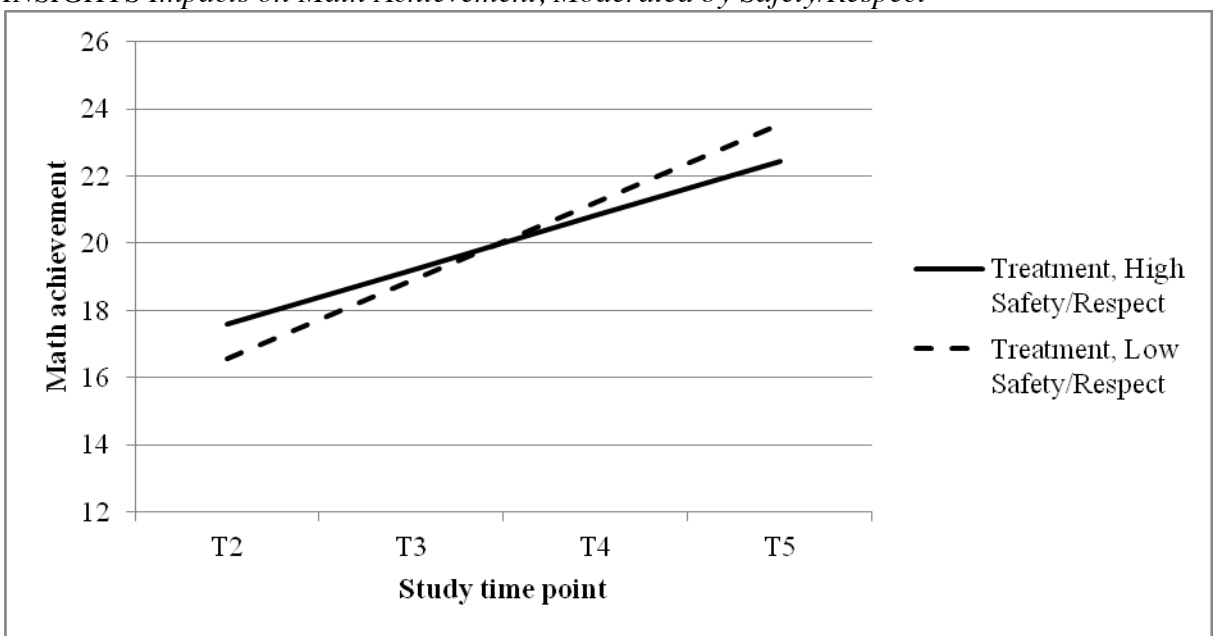
NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 2b
INSIGHTS Impacts on Reading Achievement, Moderated by Accountability



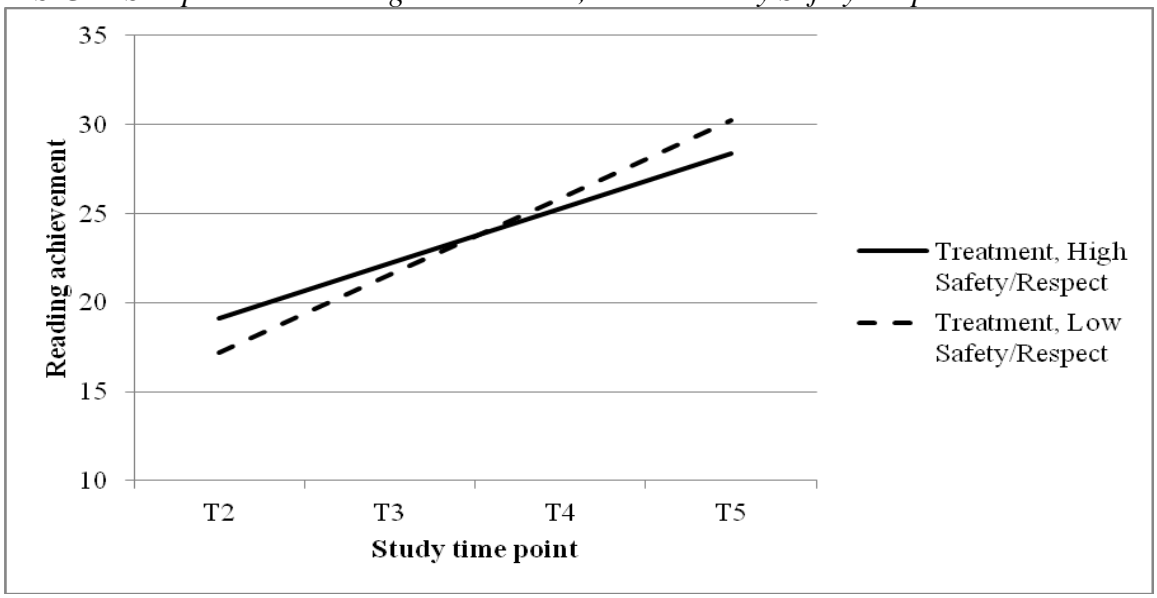
NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 3a
INSIGHTS Impacts on Math Achievement, Moderated by Safety/Respect



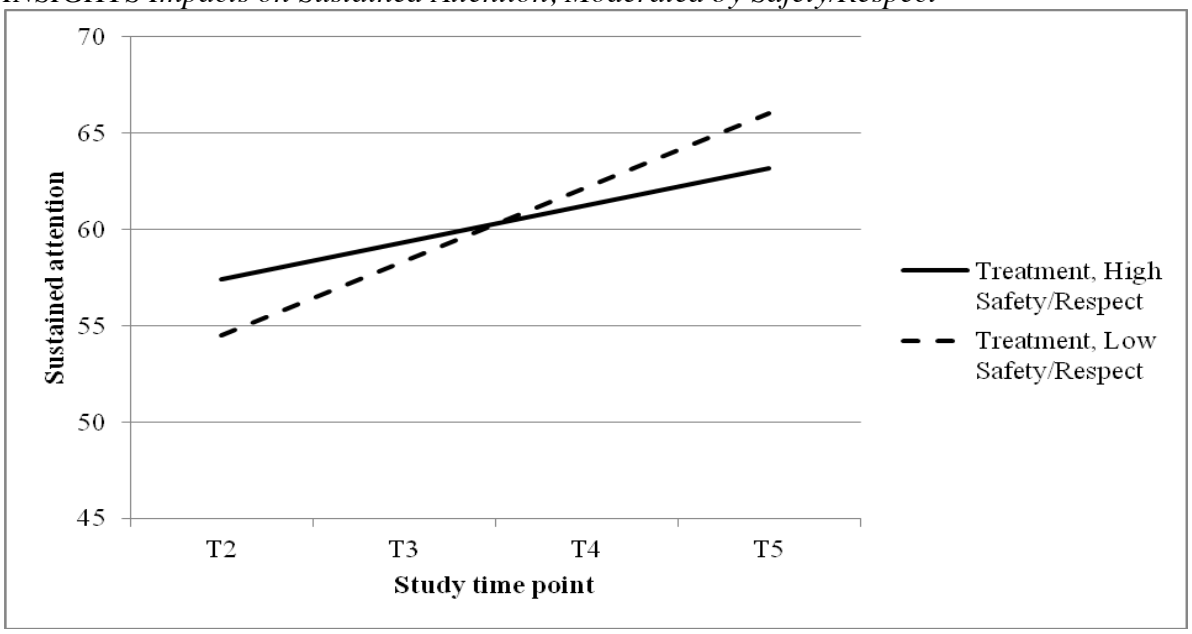
NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 3b
INSIGHTS Impacts on Reading Achievement, Moderated by Safety/Respect



NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariates.

Figure 3c
INSIGHTS Impacts on Sustained Attention, Moderated by Safety/Respect



NOTE: "High" reflects the level of the moderator at 1 SD greater than the mean; "Low" is the level of the moderator at 1 SD less than the mean. Models control for all covariat

Table 1
Model Summary for Individual Growth Models Examining Sustained Attention, Behavior Problems, Math Achievement, and Reading Achievement, Moderated by School Leadership

Fixed Effects	Math achievement			Reading achievement			Sustained attention			Behavior problems		
	γ		SE	γ		SE	γ		SE	γ		SE
Between-child estimates												
Intercept	23.57	**	0.64	32.60	**	1.37	58.49	**	1.57	2.48	**	0.18
Child female	-0.34		0.31	-0.45		0.64	1.27	†	0.68	-0.21	*	0.09
Child Black	-0.53		0.47	0.99		0.98	0.33		1.04	0.11		0.13
Child Hispanic	-0.82	†	0.48	0.18		1.01	0.20		1.08	-0.21		0.13
Elig. free/reduced lunch	-1.09	*	0.49	-2.22	*	1.03	-0.26		1.01	0.05		0.13
Math achievement, T1	0.26	**	0.03	0.23	**	0.07	0.09		0.08	-0.01		0.01
Reading achievement, T1	0.13	**	0.02	0.49	**	0.04	0.15	**	0.04	0.01		0.01
Behavior problems, T1	-0.20		0.13	-0.59	*	0.27	-0.76	**	0.29	0.49	**	0.03
Sustained attention, T1	0.03	*	0.01	0.04	†	0.02	0.17	**	0.03	0.01	†	0.01
Task persistent	-0.13		0.24	0.34		0.50	0.25		0.54	-0.03		0.07
Negative reactivity	-0.59	*	0.25	-0.42		0.51	0.26		0.55	0.13	†	0.07
Withdrawn	0.32		0.21	-0.62		0.43	-0.94	*	0.46	0.01		0.06
Activity	0.32		0.20	0.93	*	0.42	0.14		0.44	-0.04		0.06
School % black	-0.49		1.91	-1.74		4.31	-7.95		5.92	-0.04		0.56
School % Hispanic	-1.04		0.72	-2.42		1.67	3.98		2.46	-0.03		0.21
School average attendance	0.44		1.43	2.49		3.15	5.21		4.19	-0.15		0.41
School size	-0.01		0.01	-0.01		0.01	-0.01		0.01	0.01		0.01
School leadership	0.21		0.40	0.91		0.91	0.23		1.25	-0.11		0.12
Treatment	-0.02		0.33	1.65	*	0.75	1.65	†	1.00	-0.19	*	0.08
Treatment x leadership	-0.42		0.45	-1.65		1.04	0.25		1.43	0.17		0.14
Within-child estimates												
Slope	2.10	**	0.09	3.86	**	0.16	3.13	**	0.19	0.01		0.02
Treatment x leadership	-0.26	*	0.11	-0.49	*	0.19	-0.06		0.24	-0.01		0.03

Random Effects	*Variance*		*SE*	*Variance*		*SE*	*Variance*		*SE*	*Variance*		*SE*
Child intercept	4.75	**	0.61	31.19	**	3.34	22.11	**	2.91	0.57	**	0.06
Child slope	0.61	**	0.23	1.11	**	0.39	4.61	**	1.19	0.07	**	0.12
Corr., slope & intercept	0.03		0.26	5.88	**	1.20	-2.21		1.40	0.12	**	0.02
School intercept	0.01		0.01	0.39		0.74	2.53		1.37	0.01		0.01
Residual variation	10.69	**	0.56	40.24	**	1.72	49.53	**	2.58	0.64		0.03
Fit Indices												
Pseudo R^2	0.45			0.41			0.31			0.16		
Percent reduction in AIC	0.25			0.26			0.22			0.17		

** $p < 0.01$, * $p < 0.05$, † $p < 0.1$; $N = 435$

Models include control variables for cohort 1 (1 = Yes; 0 = No) and cohort 2 (1 = Yes; 0 = No).

Table 2

Model Summary for Individual Growth Models Examining Sustained Attention, Behavior Problems, Math Achievement, and Reading Achievement, Moderated by Accountability

	Math achievement		Reading achievement		Sustained attention		Behavior problems	
Fixed Effects	γ	SE	γ	SE	γ	SE	γ	SE
Between-child estimates								
Intercept	23.56 **	0.63	32.58 **	1.37	58.40 **	1.56	2.41 **	0.18
Child female	-0.32	0.30	-0.39	0.64	1.26 †	0.68	-0.20 *	0.08
Child Black	-0.51	0.47	1.04	0.98	0.34	1.04	0.10	0.13
Child Hispanic	0.81 †	0.48	0.22	1.01	0.19	1.07	-0.22 †	0.13
Elig. free/reduced lunch	-1.12 *	0.49	-2.32 *	1.02	-0.23	1.10	0.07	0.13
Math achievement, T1	0.26 **	0.03	0.23 **	0.07	0.09	0.08	-0.01	0.01
Reading achievement, T1	0.13 **	0.02	0.49 **	0.04	0.15 **	0.15	0.01	0.01
Behavior problems, T1	-0.20	0.13	-0.65 *	0.27	-0.76 **	0.29	0.49 **	0.04
Sustained attention, T1	0.03 *	0.01	0.04 †	0.03	0.17 **	0.03	0.01 †	0.01
Task persistent	-0.13	0.24	0.27	0.51	0.26	0.54	-0.04	0.07
Negative reactivity	-0.58 *	0.25	-0.43	0.52	0.26	0.55	0.13 †	0.07
Withdrawn	-0.31	0.21	-0.61	0.41	-0.95 *	0.45	0.01	0.06
Activity	0.31	0.19	0.93 *	0.42	0.14	0.44	-0.04	0.05
School % black	0.01	1.58	0.04	3.72	-6.81	5.14	-0.25	0.44
School % Hispanic	-0.97	0.82	-1.12	1.98	3.23	2.78	0.24	0.23
School average attendance	0.87	1.20	0.71	2.07	-5.57	3.47	-0.17	0.33
School size	0.01	0.01	-0.01	0.01	-0.01	0.01	0.01	0.01
School accountability	0.01	0.33	-0.20	0.77	0.32	1.03	-0.23 **	0.09
Treatment	0.01	0.34	1.70 *	0.79	1.65	1.07	-0.18 *	0.09
Treatment x accountability	-0.37	0.48	-0.43	1.15	-0.03	1.51	0.32 *	0.14
Within-child estimates								
Slope	2.11 **	0.08	3.88 **	0.16	3.14 **	0.20	0.01	0.01
Treatment x accountability	-0.42 **	0.14	-0.70 **	0.26	-0.21	0.33	-0.01	0.01
Random Effects	*Variance*	*SE*	*Variance*	*SE*	*Variance*	*SE*	*Variance*	*SE*

Child intercept	4.74	**	0.61	31.36	**	3.56	22.11	**	2.91	0.57	**	0.06
Child slope	0.59	**	0.23	1.13	**	0.33	4.59	**	1.19	0.07	**	0.02
Corr., slope & intercept	0.02		0.27	5.96	**	1.18	-2.19		1.40	0.13	**	0.02
School intercept	0.01		0.01	0.55		0.74	2.58	†	1.38	0.01		0.01
Residual variation	10.70	**	0.56	40.16	**	1.64	49.54	**	2.57	0.64	**	0.03
Fit Indices												
Pseudo R^2	0.45			0.41			0.31			0.16		
Percent reduction in AIC	0.28			0.27			0.23			0.13		

** $p < 0.01$, * $p < 0.05$, † $p < 0.1$; $N = 435$

Models include control variables for cohort 1 (1 = Yes; 0 = No) and cohort 2 (1 = Yes; 0 = No).

Table 3

Model Summary for Individual Growth Models Examining Sustained Attention, Behavior Problems, Math Achievement, and Reading Achievement, Moderated by Safety and Respect

	Math achievement			Reading achievement			Sustained attention			Behavior problems		
Fixed Effects	γ		SE	γ		SE	γ		SE	γ		SE
Between-child estimates												
Intercept	22.79	**	0.51	30.10	**	1.17	57.86	**	1.27	2.46	**	0.15
Child female	-0.35		0.30	-0.34		0.64	1.32	*	0.67	-0.20	*	0.08
Child Black	-0.70		0.45	0.95		0.95	0.23		0.99	0.15		0.12
Child Hispanic	-1.04	*	0.46	-0.04		0.97	0.02		1.02	-0.16		0.13
Elig. free/reduced lunch	-0.04		0.03	-0.01		0.07	0.06		0.08	0.01		0.01
Math achievement, T1	0.28	**	0.03	0.25	**	0.07	0.11		0.07	-0.01		0.01
Reading achievement, T1	0.13	**	0.02	0.49	**	0.04	0.14	**	0.05	0.01		0.01
Behavior problems, T1	-0.15		0.13	-0.43		0.27	-0.66	*	0.28	0.50	**	0.03
Sustained attention, T1	0.03	*	0.01	-0.04	†	0.04	0.18	**	0.03	0.01		0.01
Task persistent	-0.13		0.24	0.39		0.51	0.33		0.54	-0.02		0.07
Negative reactivity	-0.57	*	0.25	0.53		0.52	0.17		0.54	0.12	†	0.07
Withdrawn	-0.34	†	0.20	-0.65		0.42	-0.74	†	0.44	0.01		0.05
Activity	0.29		0.20	0.92	*	0.41	0.17		0.44	-0.03		0.05
% black	-1.45		1.90	-5.98		4.77	-14.08	*	5.62	-0.11		0.59
% Hispanic	-0.98		0.73	-2.89		1.84	2.50		2.17	-0.14		0.22
Average attendance	0.49		1.39	5.77	†	3.34	-0.36		3.85	-0.10		0.42
School size	0.01		0.01	-0.01	†	0.01	-0.01		0.01	-0.01		0.01
Safety and respect	0.51		0.47	2.59	*	1.15	2.89	*	1.35	-0.06		0.14
Treatment	-0.14		0.36	-1.27		0.90	2.77	**	1.06	-0.17		0.11
Treatment x safety and respect	-0.78		0.59	-3.24	*	1.46	-1.51		1.65	0.03		0.18
Within-child estimates												
Slope	1.98	**	0.09	3.71	**	0.16	2.88	**	0.20	0.01		0.02
Treatment x safety and respect	-0.49	**	0.16	-0.86	**	0.29	-1.33	**	0.36	0.01		0.04
Random Effects	*Variance*		*SE*	*Variance*		*SE*	*Variance*		*SE*	*Variance*		*SE*

Intercept	4.83	**	0.60	32.55	**	3.37	22.14	**	2.86	0.57	**	0.06
Slope	0.57	**	0.22	1.06	**	0.36	3.85	**	1.12	0.07	**	0.02
Corr., slope & intercept	0.02		0.26	5.86	**	1.17	-2.54	*	1.32	0.13	**	0.02
School intercept	0.01		0.01	0.74		0.80	1.61		1.07	0.01		0.01
Level-1 residual	10.62	**	0.54	39.92	**	1.67	49.99	**	2.55	0.64	**	0.03
Fit Indices												
Pseudo R^2	0.45			0.42			0.30			0.16		
Percent reduction in AIC	0.23			0.22			0.19			0.13		

** $p < 0.01$, * $p < 0.05$, † $p < 0.1$; $N = 435$

Models include control variables for cohort 1 (1 = Yes; 0 = No) and cohort 2 (1 = Yes; 0 = No).

Appendix C

The final model for moderated impact analyses (modeled separately for each outcome and each school context variable (school poverty, leadership, accountability, safety/respect) is:

Level 1: $Y_{tij} = \pi_{0ij} + \pi_{1ij(}\text{Assessment point - 4})_{tij} + \varepsilon_{tij}$

Level 2: $\pi_{0ij} = \beta_{00j} + \beta\text{Treatment}j + \delta_j + r_{0ij}$

$\pi_{1ij} = \beta_{10j} + r_{1ij}$

Level 3: $\beta_{00j} = \gamma_{000} + \gamma\text{SchoolContextVariable} + X + \eta_j + u_{00j}$

$\beta_{10j} = \gamma_{100} + \gamma\text{SchoolContextVariable} + u_{10j}$

In the Level 1 model, Y_{tij} is the academic outcome at time t of student i in school j. π_{0ij} is the academic outcome score at T5 for student i in site j. Assessment point - 4_{tij} is a measure of time for student i in school j. π_{1ij} is the growth trajectory for student i in school j. e_{tij} is a random error term that represents the residual (or unexplained) variation in the outcome. In the Level 2 model, the students' academic outcome score at T5 is modeled as a function of the mean achievement score in within school j (β_{00j}) and a random effect (r_{0ij}) that allows the intercept to vary randomly around the student (Level 2) mean. δ_j is a vector of student-level covariates. The growth rate in academic scores for student i in school j is modeled as a function of the mean growth in academic outcomes in school j (β_{10j}) and a random effect that allows students' trajectories to vary randomly around the student mean trajectory (r_{1ij}). At Level 3, β_{00j} is a function of the overall mean academic score score at T5 (γ_{000}) and a random effect that allows the intercept to vary around the school level mean. X is a vector of school-level covariates. Finally, β_{10j} is a function of the overall mean growth rate (γ_{100}) in academic outcomes in site j and a random effect that allows the achievement trajectory to vary randomly around the overall mean trajectory. The coefficients of interest for this study are the cross level interactions for the slope (growth) and intercept (overall effects) between dimensions of school climate and Treatment.