

Evaluation of the Safe and Civil Schools Leadership Program: A Technical Report

Prepared by

Keith C. Herman, Francis Huang, & Wendy M. Reinke
Missouri Prevention Science Institute
University of Missouri



June 2026

Take Home Message:

SCSL may be a promising leadership model, but this trial shows how difficult it is to translate leadership training into measurable changes in school climate and student experience—especially amid pandemic-era implementation disruptions.

Abstract

This technical report presents findings from a group-randomized trial evaluating the Safe and Civil Schools Leadership program (SCSL), a professional development and coaching model designed to help school leaders improve school safety, student behavior, disciplinary systems, and school climate. Prior research suggests that school leadership influences student outcomes indirectly through organizational conditions, with school safety and disciplinary climate emerging as especially important mediators. SCSL was developed to strengthen leaders' use of distributed leadership, data-based decision-making, schoolwide behavior planning, and positive behavior support practices. The study included 90 schools recruited across six cohorts, with schools randomly assigned to SCSL or a control/business-as-usual condition. Survey data were collected from students, teachers/staff, and principals across multiple time points. The analytic sample included 59,863 student responses from 83 schools and 10,588 teacher/staff responses. Outcomes included student and staff reports of disciplinary structure, student support, safety, bullying, school problems, belonging, suspension, and fighting, as well as principal-reported school climate and data-use measures. Using ANCOVA models with school-level cluster robust standard errors and Benjamini-Hochberg corrections for multiple comparisons, results indicated no statistically significant intervention effects on student, teacher/staff, or principal outcomes. Although SCSL was implemented with high fidelity and was positively rated by participating leaders, the program did not produce measurable improvements in the primary outcomes. Findings suggest that leadership-focused professional development and coaching alone may be insufficient to change student and staff experiences of school climate and discipline, and that future interventions may require more intensive implementation supports or stronger mechanisms connecting leadership practices to classroom and schoolwide behavior systems.

1. Introduction

Considerable research has shown that school leadership is central to developing high performing schools (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Louis, Leithwood, Wahlstrom, & Anderson, 2010). Over four decades of research suggests that school leadership has significant, albeit indirect effects on student outcomes through multiple school organizational factors (Bosker & Witziers, 1995; Hallinger & Heck, 1996a, 1996b, 1998; Robinson, Lloyd, & Rowe, 2008; Waters, Marzano, & McNulty, 2003; Witziers, Bosker, & Kruger, 2003). However, much of the evidence has relied on cross-sectional data and case studies (Hallinger & Heck, 2010a). Although more recent longitudinal studies have improved the empirical basis for linking leadership to student outcomes (Dumay, Boonen, & Van Damme, 2013; Hallinger & Heck, 2010a, 2010b; Heck & Hallinger, 2009), the field still has limited experimental evidence establishing whether specific leadership practices or leadership professional development programs produce measurable improvements in school climate, student behavior, or student outcomes (Heck & Hallinger, 2010; Neumerski, 2013).

Prior research suggests that school learning climate, and particularly school safety and disciplinary climate, may be among the most important mediators linking leadership to student outcomes. A previous IES-funded study used the school organizational framework developed by Bryk, Sebring, Allensworth, Luppescu, and Easton (2010) to examine multiple mediators through which leadership may influence student learning, including professional capacity, parent and community ties, and school learning climate. Findings from that work indicated that school learning climate, as measured by school safety, peer behavior, and student expectations, was the only mediator through which school leadership was significantly associated with student learning outcomes in urban settings (Sebastian & Allensworth, 2012, 2013; Sebastian, Allensworth, & Huang, 2014; Sebastian, Allensworth, & Stevens, 2014). Cross-national evidence from PISA similarly found that school disciplinary climate was consistently linked to academic performance across many countries (OECD, 2010a, 2010b, 2013a, 2013b).

Research on principal practice also supports this focus. Grissom and Loeb (2011) found that principals' organizational management skills were the only set of leadership skills associated with school performance, with school safety serving as a key component of this domain. Horng, Klasik, and Loeb (2010) similarly found that time spent on organizational management activities was positively associated with school outcomes. Although principals are often expected to serve as instructional leaders, counselors, coaches, and managers, one of their most fundamental responsibilities is ensuring that students experience schools as safe, orderly, and supportive environments (Connelly, 2013).

Despite the importance of this leadership role, many school leaders report feeling underprepared to address student misbehavior, school discipline, and school safety. Leadership preparation programs have been criticized for insufficiently connecting theory to practice, lacking authentic experiences, and failing to prepare leaders for the complexities of school improvement (Camburn, Goldring, Sebastian, May, & Huff, 2015; Darling-Hammond, LaPointe, Meyerson, T., & Cohen, 2007; Levine, 2005). These limitations are

especially important because many principals first serve as assistant principals, where student behavior, discipline, and safety are often central responsibilities (Glanz, 1994; Kealey, 2002; Oleszewski, Shoho, & Barnett, 2012). Although principal professional development programs are common, rigorous evidence of their efficacy remains limited, and school safety and student behavior are not typically central foci of leadership professional development (Camburn et al., 2015; Darling-Hammond et al., 2007; Clendenin, 2009; Lisle, 2002; Sheras, Cornell, & Bostain, 1996; Timmons, 2010).

The Safe and Civil Schools Leadership program (SCSL) was developed to address this gap. SCSL is a research-based leadership training and coaching model designed to help school leaders promote positive student behavior, improve school safety, and strengthen schoolwide discipline systems (Sprick, Howard, Wise, Marcum, & Haykin, 1998). The program emphasizes distributed leadership, data-based decision-making, schoolwide behavior planning, and the use of leadership teams to guide implementation, consistent with broader distributed leadership perspectives in school leadership research (Spillane, 2006; Spillane, Camburn, & Pareja, 2007; Spillane, Halverson, & Diamond, 2004). Prior non-experimental evaluations and reports suggested that SCSL and related Safe and Civil Schools models may be associated with improvements in perceptions of school safety, student relationships, discipline referrals, and school climate (Jacobsen & Polin, 2006; Rickert, 2005; Smolkowski, Strycker, & Ward, 2016; Ward & Gersten, 2013). However, despite its promise and widespread adoption, SCSL had not previously been tested through a rigorous randomized controlled trial and did not meet the highest standards of evidence established by review groups such as the What Works Clearinghouse (2008).

To fill this gap, the current study evaluated the impact of the SCSL program on school leadership practices, school safety, school climate, student behavior, and related student and staff outcomes across schools from multiple districts representing rural, urban, and suburban settings. The primary innovations of the study were the implementation of a widely used training program focused on school leaders' skills in promoting school safety and behavior management and the first randomized controlled trial of its effects on perceptions of school safety and climate, student behaviors, and academic-related school outcomes.

SCSL Program Description

SCSL is a research-based series of training and coaching materials designed to help school leaders develop effective skills for promoting positive student behavior and school safety (Sprick, Howard, Wise, Marcum, & Haykin, 1998). Certified trainers provide two full days of training and ongoing coaching to school leaders. SCSL is grounded in social learning theory and includes research-based principles of systematic data collection and decision-making. The SCSL model targets school leaders' use of effective school-wide discipline practices by promoting positive relationships with all students. The acronym STOIC summarizes the guiding principles for creating effective and orderly interactions in all school settings: Structure/organize all settings for success; Teach students how to behave responsibly; Observe student behavior; Interact positively with students; and Correct irresponsible behavior fluently.

SCSL is based on a distributed leadership model that emphasizes data-based decision making practices. SCSL promotes collaborative decision making by instituting a Behavior Leadership Team composed of the school principal, an influential teacher leader, and other personnel involved in supporting student behavior success, such as school psychologists. Decisions made by the Behavior Leadership Team follow principles of effective data-based decision-making guidelines that uses data from multiple sources—teachers, students, leaders, and parents—to guide the review of existing policies/procedures, adoption and implementation of new policies, and the maintenance of effective current policies.

Table 1. SCSL Topics and Content

Topics	Content
Leadership Skills	Clarifying a Vision, building trust, using rituals, behavior modification
Schoolwide Behavior Plan	Guiding principles, mission and staff statement, student success guidelines
Behavior Leader Teams	How to create teams to support the schoolwide behavior plan
Data-Based Decisions	How to collect and use meaningful data in continuous cycle
Instructional Approach	Clear expectations, increasing consistency, planning for emergencies
Common Areas	Improving behavior in halls, cafeteria, playground
Classrooms	Unified classroom management model, assisting teachers with behavior
Referrals and Solutions	How to manage discipline referrals
Meaningful Work	Assigning jobs to improve student motivation and responsibility

In addition to the SCSL Workshops, participants have access to online materials designed to supplement and enhance skill development, and a certified SCSL coach provides onsite coaching to intervention schools. The online training is available for both years of the program during which the SCSL coach prompts principals to complete simple assignments on a monthly basis. The SCSL coach also meets with each leadership team twice each year to review the content of the workshop, assess areas of difficulty with implementation, and problem solve any implementation barriers.

2. Method

2.1 Participants

A total of 90 schools were recruited over six cohorts to be part of the project (see Table 1). Approximately half of the schools were randomly assigned to the intervention condition and the other half was randomized into the control/business-as-usual condition. There were 90 schools that provided staff/teacher responses and 88 schools that provided student responses. Two schools that provided staff responses had students who were in grades K-2 and were not eligible for student portion of the survey. Covid primarily affected student participation in years 2 through 4.

In total, there were 65,129 anonymous student survey responses across four time periods. Of the student respondents, 3,851 (5.9%) indicated that they were not responding truthfully or were classified as inattentive responders (i.e., responding too fast). The use of validity screening items has been shown to improve the overall quality of survey results (Cornell et al., 2012; Furlong et al., 2017). This reduced the sample to 61,278 students.

In SY 2019-2020, five schools ($n_c = 2$, $n_t = 3$) from one school district did not collect student responses in Spring 2020 (due to Covid) and did not resume data collection (other schools recruited in the time period resumed the following school year) and we excluded an additional 1,415 students ($n_c = 724$, $n_t = 691$). For the analysis using student data, all other schools ($n = 83$) were used in the analysis that had two or more data collection points ($n = 59,863$; see Table 2).

Table 1. Number of schools per cohort and time period providing responses for student or staff surveys.

Cohort	Time							
	1 (Fall)		2 (Spring)		3 (Fall)		4 (Spring)	
	Students	Staff	Students	Staff	Students	Staff	Students	Staff
1: 2017-2019	21	21	21	21	20	20	20	20
2: 2018-2020	13	15	13	15	13	15	4 ^a	13
3: 2019-2021	23	23	4 ^a	22	17	17	18	18
4: 2020-2022 ^b	2	2	2	2	2	2	2	2
5: 2021-2023	22	22	22	22	18	18	16	16
6: 2022-2024	7	7	7	7	3	3	3	3
Total	88	90	69	89	73	75	63	72

Notes. ^aSpring 2020 (start of Covid). ^bSchools operating remotely. Cohorts 2, 3, and 4 were impacted by Covid (e.g., schools operating virtually).

Table 2. Number of student responses per time period.

Cohort	Time			
	1 (Fall)	2 (Spring)	3 (Fall)	4 (Spring)
1: 2017-2019	6,053	4,573	5,320	5,665
2: 2018-2020	4,321	4,171	4,485	255
3: 2019-2021	3,584	190 ^a	2,043	2,507
4: 2020-2022 ^b	363	319	377	363
5: 2021-2023	3,785	3,528	2,659	2,168
6: 2022-2024	1,122	1,091	512	409
Total	19,229	13,874	15,399	11,371

Notes. ^aSpring 2020 (start of Covid). ^bSchools operating remotely. Cohorts 2, 3, and 4 were impacted by Covid (e.g., schools operating virtually).

For the teacher/staff survey, 12,894 survey responses were initially collected. Of those responses, 2,077 were responses that were too quick and were dropped (e.g., surveys were opened to view some questions and then closed). An additional 31 respondents indicated that they were not paying attention to the survey and 198 had no responses and were excluded. The final analytic sample for the staff survey was 10,588 with a median response time of 16 minutes. The number of responses per time period can be seen in Table 3.

Table 3. Number of teacher/staff responses per time period.

Cohort	Time			
	1 (Fall)	2 (Spring)	3 (Fall)	4 (Spring)
1: 2017-2019	629	657	682	697
2: 2018-2020	578	553	532	426
3: 2019-2021	907	853	507	548
4: 2020-2022 ^b	55	49	51	61
5: 2021-2023	709	605	555	446
6: 2022-2024	180	191	56	61
Total	3,058	2,908	2,383	2,239

All descriptive statistics for students and teachers/staff can be found in Tables 4 and 5, respectively. Differences of participants in the treatment and control conditions were not meaningfully different, given the small effect size difference (all d s < 0.20).

Table 4. Descriptive statistics for student participants (across multiple time points).

Student level: n	Control		Treatment		<i>d</i>
	31,140		28,723		
Gender	M (SD)	n (%)	M (SD)	n (%)	0.05
Male		14,636 (47.0)		13,271 (46.2)	
Female		15,200 (48.8)		14,475 (50.4)	
Unknown		1,304 (4.2)		977 (3.4)	
Race\Ethnicity					0.16
White		16,363 (52.5)		13,540 (47.1)	
Black		5,589 (17.9)		6,936 (24.1)	
Hispanic		3,963 (12.7)		3,436 (12.0)	
Other		1,658 (5.3)		1,398 (4.9)	
Two+		3,567 (11.5)		3,413 (11.9)	
Grade					0.17
3-5		14,048 (45.1)		13,968 (48.6)	
6-8		11,481 (36.9)		8,442 (29.4)	
9-12		5,611 (18.0)		6,313 (22.0)	
FRPM					0.07
No		13,808 (44.3)		12,041 (41.9)	
Yes		12,901 (41.4)		11,882 (41.4)	
Don't know		4,431 (14.2)		4,800 (16.7)	
School-level: n	41		42		
% FRPM	56.37 (27.80)		57.36 (28.68)		0.04
Size	465.82 (224.13)		467.98 (327.81)		0.01
% Minority	48.55 (33.37)		49.93 (34.61)		0.04
School type					0.16
Alternative		1 (2.4)		1 (2.4)	
Combined		2 (4.9)		1 (2.4)	
Elementary		23 (56.1)		25 (59.5)	
Middle		10 (24.4)		9 (21.4)	
High		5 (12.2)		6 (14.3)	

Notes. FRPM = free or reduced price meals eligible. Differences based on p values for categorical variables using chi-square tests and t-tests using cluster-robust standard errors for continuous variables were not statistically significant.

Table 5. Descriptive statistics for teacher/staff participants (across multiple time points).

	Control 5,215	Treatment 5,373	<i>d</i>
Gender	n (%)	n (%)	0.02
Male	846 (16.2)	826 (15.4)	
Female	4,018 (77.0)	4,186 (77.9)	
Missing	351 (6.7)	361 (6.7)	
Years teaching			0.07
1-2 years	460 (8.8)	513 (9.5)	
3-5 years	828 (15.9)	739 (13.8)	
6-10 years	893 (17.1)	909 (16.9)	
More than 10 years	2,684 (51.5)	2,857 (53.2)	
Missing	350 (6.7)	355 (6.6)	
School level			0.20
Elementary School Teachers	2,433 (46.7)	2,577 (48.0)	
Middle School Teachers	1,309 (25.1)	1,030 (19.2)	
High School Teachers	722 (13.8)	1,063 (19.8)	
Teaches Multiple School Levels	386 (7.4)	333 (6.2)	
Other School Faculty/Staff	365 (7.0)	370 (6.9)	
Race/Ethnicity			0.05
Black	349 (6.7)	375 (7.0)	
Hispanic	99 (1.9)	118 (2.2)	
Other	34 (0.7)	20 (0.4)	
Two or more	110 (2.1)	132 (2.5)	
White	4,254 (81.6)	4,349 (80.9)	
Missing	369 (7.1)	379 (7.1)	

Notes. All differences were not statistically significant (all $ps > .10$).

2.2.a. Measures: Student

For scales, mean scores were taken for all the scale items. Response options for items ranged from 1 (Strongly disagree) to 4 (Strongly agree). The scales for Disciplinary structure and Student support were from the Authoritative School Climate Survey (Konold et al., 2014). Reverse coding was performed as appropriate. Modifications were made to the scales for students in grades 3 and 4 who were provided a shortened version of the survey (items with the highest factor loadings were selected). Other outcomes were single item measures (e.g., bullied, suspension, fighting) that were recoded as 0 = no or 1 = yes or once or more.

Disciplinary structure: Two items for grades 3 and 4 (e.g., “The school rules are fair”, “The punishment for breaking the rules is the same for all students”) and 7 items for grades 5+.

Student support: Two items for grades 3 and 4 (Most teachers and other adults at this school: “Care about all students”, “Want all students to do well”) and 8 items for grades 5+.

Safety: Three items for grades 3 and 4 (“I feel safe at this school”, “I feel safe going to and from this school”, “I sometimes stay at home because I do not feel safe at this school”) and 8 items for grades 5+.

School problems: Only asked of middle and high school students. Four items that asked if the following types of behaviors occur at school (“Students arrive later at school”, “Absenteeism”, “Classroom disturbance”, “Profanity”).

Bullied: Single item: “I have been bullied at school this year” (0 = Never, 1 = once or twice, about once per week, more than once per week”). M = 30%.

Bullied others: Single item: “I have bullied others at school this year” (0 = Never, 1 = once or twice, about once per week, more than once per week”). M = 11%.

School belonging: Single item: “I feel like I belong at this school.” (1 = Strongly disagree, 4 = strongly agree).

Suspension: Have been suspended at school (0 = no, 1 = once or more). M = 5%.

Fighting: Single item: “During the past 12 months, how many times were you in a physical fight on school property?” (0 = 0 times, 1 = 1 or more). M = 15%.

2.2.b. Measures: Teacher/Staff

Scales for the teacher survey ranged from 1 = Strongly disagree to 6 = Strongly agree. The scales of Justness and Fairness (both part of disciplinary structure), Student support, and the Prevalence of teasing and bullying were taken from the Authoritative School Climate Survey-Teacher/Staff version (Huang et al., 2015).

Justness: Related to disciplinary structure, composed of three items such as “Students get suspended for minor things” and “When students are accused of doing something wrong, they get a chance to explain.”

Fairness: Also related to disciplinary structure, composed of four items such as: “If a student does something wrong, he or she will definitely be punished” and “The punishment for breaking the rules is the same for all students.”

Student support: Composed of 10 items such as: “Students know whom to go to for help if they have been treated badly by another student” and “Teachers/staff take action to solve the problem when students report bullying.”

Safety: Composed of six items which include items such as: “I feel safe at this school” and “Students at this school threaten to hurt other students.”

Respect for students: Composed of four items which asks if most teachers and adult at this school: “Care about all students” or “Want all students to do well.”

Prevalence of Teasing and Bullying: Scale composed of five items which include items such as: “Bullying is a problem at this school” and “Students in this school are teased or put down because of their race or ethnicity.”

School problems: Scale composed of four items which asks if the following types of problems occur at this school often: “Absenteeism”, “Profanity”, “Classroom disturbance,” and “Student arrive late at this school.”

2.3 Analytic Strategy

Given the seasonality of school climate (i.e., the tendency for climate to fluctuate within a school year with less favorable measures in the Spring) (Huang et al., 2023), we used an ANCOVA approach to evaluate the outcomes separately in the Spring (Time 2), Fall (Time 3), and Spring (Time 4), while controlling for school-aggregated baseline measures from Fall (Time 1). Both individual and school-level covariates were included in all models. The primary coefficient of interest is the binary school-level treatment indicator. Cluster robust standard errors (at the school level) were used to account for the clustered nature of the data (Huang & Li, 2022). All outcomes were standardized so treatment effect coefficients can be interpreted as a standardized mean difference or Cohen's *d*. As several outcomes were investigated and keeping in line with What Works Clearinghouse (2022) guidelines, we accounted for multiple comparisons using a Benjamini-Hochberg (1995) correction.

3. Results

Student and Teacher/Staff Outcomes

Across the nine different student outcomes during the time periods (time 2 to time 4 from spring-fall-spring), none of the student-reported outcomes were statistically significant (using an alpha = .05). For the seven teacher/staff outcomes, almost all the treatment effects were not statistically significant, with the exception of Student Support at Time 3 (fall in the following year) where intervention schools had slightly lower reported measures compared to control schools ($d = -0.10$, $p = .045$). However, after adjusting for multiple comparisons, the reported difference was also not statistically significant.

Table 6.
Regression results for student-reported outcomes.

	Disciplinary Structure			Student Support			Safety		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-4.281*** (0.715)	-2.256*** (0.613)	-2.261*** (0.667)	-3.404*** (0.662)	-2.779*** (0.693)	-3.010*** (0.793)	-1.506* (0.681)	-0.373 (0.489)	-2.206* (1.008)
Female	0.034 (0.025)	0.052* (0.027)	0.062* (0.027)	0.000 (0.020)	0.025 (0.021)	0.004 (0.024)	-0.043* (0.017)	-0.041* (0.019)	-0.049* (0.024)
Black ¹	-0.202*** (0.027)	-0.183*** (0.033)	-0.222*** (0.030)	-0.177*** (0.029)	-0.124*** (0.035)	-0.229*** (0.031)	-0.161*** (0.038)	-0.183*** (0.032)	-0.181*** (0.038)
Hispanic ¹	-0.184*** (0.028)	-0.177*** (0.026)	-0.208*** (0.029)	-0.190*** (0.028)	-0.173*** (0.028)	-0.204*** (0.036)	-0.220*** (0.029)	-0.198*** (0.034)	-0.235*** (0.034)
Other ¹	-0.088* (0.039)	-0.196*** (0.043)	-0.208*** (0.043)	-0.144*** (0.043)	-0.133*** (0.038)	-0.163** (0.054)	-0.099* (0.043)	-0.190*** (0.034)	-0.167*** (0.041)
Two or more ¹	-0.256*** (0.031)	-0.185*** (0.028)	-0.235*** (0.027)	-0.232*** (0.033)	-0.197*** (0.032)	-0.242*** (0.037)	-0.189*** (0.031)	-0.205*** (0.034)	-0.178*** (0.032)
ES ²	0.121 (0.083)	0.273*** (0.070)	0.292*** (0.064)	0.265** (0.084)	0.338*** (0.093)	0.294*** (0.067)	0.269*** (0.060)	0.333*** (0.062)	0.266** (0.081)
HS ²	0.080 (0.063)	-0.166*** (0.048)	-0.164 (0.123)	0.070 (0.049)	-0.167*** (0.050)	-0.131+ (0.074)	-0.117 (0.095)	-0.327*** (0.071)	-0.182+ (0.104)
FRPM: Y	0.061** (0.023)	0.041+ (0.022)	0.075** (0.024)	0.110*** (0.024)	0.082*** (0.023)	0.095*** (0.028)	0.056* (0.023)	0.043+ (0.025)	0.066** (0.023)
FRPM: U	0.128** (0.047)	0.095** (0.036)	0.157*** (0.038)	0.145*** (0.034)	0.139*** (0.032)	0.202*** (0.037)	0.126** (0.047)	0.051 (0.041)	0.039 (0.045)
<i>School-level</i>									
Treatment	0.038 (0.040)	0.035 (0.037)	0.059 (0.052)	0.027 (0.031)	0.027 (0.033)	0.019 (0.044)	0.071 (0.059)	0.085+ (0.048)	0.010 (0.073)
Size	0.000 (0.001)	0.002** (0.001)	0.022 (0.018)	0.001+ (0.001)	0.002** (0.001)	0.016 (0.011)	0.004* (0.002)	0.004* (0.002)	0.007 (0.018)
% FRPM	0.000 (0.001)	-0.001 (0.001)	0.000 (0.002)	-0.001 (0.001)	-0.002+ (0.001)	-0.001 (0.001)	-0.003+ (0.001)	-0.004** (0.002)	0.000 (0.003)
% minority	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	-0.002 (0.001)	-0.003** (0.001)	0.000 (0.002)
n	13,872	15,396	11,367	13,872	15,396	11,367	13,872	15,396	11,367
r ²	0.099	0.101	0.090	0.122	0.152	0.128	0.125	0.123	0.126

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals (YES or Unknown). Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ES = elementary school. HS = high school. ¹White is the reference group. ²Middle school is the reference group.

Table 6 (continued).
Regression results for student-reported outcomes.

	School Problems ^a			Bullied			Bullied Others		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-3.550*** (0.656)	-1.732* (0.757)	-2.363+ (1.338)	0.042*** (0.012)	0.032** (0.011)	0.057* (0.027)	0.021** (0.007)	0.001 (0.003)	0.015* (0.006)
Female	0.128*** (0.026)	0.140*** (0.026)	0.158*** (0.030)	-0.004 (0.005)	-0.020*** (0.004)	-0.013* (0.005)	-0.018*** (0.003)	-0.010*** (0.002)	-0.011*** (0.003)
Black ¹	-0.012 (0.039)	-0.015 (0.040)	0.104 (0.064)	-0.024* (0.010)	-0.019* (0.010)	-0.020+ (0.011)	0.008 (0.007)	0.004 (0.005)	0.012+ (0.007)
Hispanic ¹	0.058 (0.050)	0.066 (0.043)	0.102* (0.043)	0.039** (0.012)	0.036*** (0.008)	0.018 (0.012)	0.027*** (0.008)	0.026*** (0.006)	0.022** (0.008)
Other ¹	0.031 (0.061)	0.110* (0.046)	0.005 (0.085)	0.012 (0.012)	0.036*** (0.011)	0.025+ (0.014)	0.023* (0.009)	0.024*** (0.007)	0.027** (0.009)
Two or more ¹	0.106* (0.049)	0.161*** (0.036)	0.137** (0.047)	0.028** (0.010)	0.023** (0.008)	0.027* (0.012)	0.007 (0.005)	0.005 (0.004)	0.011* (0.006)
ES ²	a	a	a	0.003 (0.010)	0.008 (0.009)	0.010 (0.014)	-0.003 (0.005)	0.003 (0.003)	0.002 (0.006)
HS ²	0.026 (0.092)	0.187** (0.060)	0.102 (0.113)	-0.035*** (0.009)	-0.010 (0.008)	-0.020 (0.019)	0.004 (0.008)	0.004 (0.004)	0.017*** (0.005)
FRPM: Y	-0.044 (0.032)	-0.067** (0.023)	0.006 (0.037)	0.008 (0.007)	0.015* (0.006)	0.019** (0.007)	0.002 (0.003)	0.006* (0.003)	0.001 (0.004)
FRPM: U	a	a	a	0.001 (0.011)	0.007 (0.007)	0.002 (0.012)	-0.002 (0.005)	0.001 (0.004)	0.005 (0.006)
<i>School-level</i>									
Treatment	0.001 (0.075)	-0.003 (0.061)	0.135 (0.113)	-0.003 (0.007)	0.005 (0.006)	0.001 (0.011)	-0.002 (0.004)	0.004 (0.003)	-0.002 (0.004)
Size	-0.013 (0.012)	0.000 (0.010)	0.003 (0.016)	0.000 (0.000)	0.000*** (0.000)	0.001 (0.003)	0.000 (0.000)	0.000+ (0.000)	-0.001 (0.001)
% FRPM	0.001 (0.002)	0.005* (0.002)	0.007+ (0.004)	0.000 (0.000)	0.000+ (0.000)	0.000 (0.000)	0.000 (0.000)	0.000+ (0.000)	0.000** (0.000)
% minority	0.001 (0.002)	0.002 (0.001)	-0.004 (0.003)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
n	7,346	8,078	5,513	13,872	15,396	11,367	13,872	15,396	11,367
r ²	0.060	0.053	0.050	0.020	0.017	0.027	0.011	0.013	0.009

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals (YES or Unknown). Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ^aNot asked of students in grades 3 and 4. ES = elementary school. HS = high school. ¹White is the reference group. ²Middle school is the reference group.

Table 6 (continued).
Regression results for student-reported outcomes.

	School Belonging			Suspension			Fighting ^a		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	3.063*** (0.072)	3.167*** (0.047)	2.925*** (0.101)	0.028* (0.013)	0.007 (0.009)	0.101*** (0.022)	0.170*** (0.028)	0.129*** (0.022)	0.164*** (0.023)
Female	-0.069** (0.023)	-0.065** (0.021)	-0.059* (0.026)	-0.043*** (0.004)	-0.035*** (0.004)	-0.052*** (0.005)	-0.127*** (0.011)	-0.104*** (0.012)	-0.099*** (0.012)
Black ¹	-0.165*** (0.035)	-0.154*** (0.031)	-0.156*** (0.036)	0.066*** (0.009)	0.022** (0.007)	0.041*** (0.012)	0.088*** (0.017)	0.089*** (0.018)	0.105*** (0.017)
Hispanic ¹	-0.123*** (0.033)	-0.136*** (0.022)	-0.131*** (0.031)	0.061*** (0.010)	0.038*** (0.006)	0.044*** (0.009)	0.084*** (0.019)	0.066*** (0.018)	0.084*** (0.017)
Other ¹	-0.088* (0.041)	-0.102* (0.042)	-0.155*** (0.045)	0.038** (0.014)	0.030*** (0.008)	0.021+ (0.012)	0.093*** (0.020)	0.057*** (0.017)	0.078** (0.027)
Two or more ¹	-0.242*** (0.037)	-0.178*** (0.028)	-0.214*** (0.037)	0.042*** (0.007)	0.022*** (0.006)	0.029** (0.010)	0.079*** (0.012)	0.085*** (0.014)	0.113*** (0.025)
ES ²	0.411*** (0.051)	0.431*** (0.042)	0.488*** (0.045)	-0.038*** (0.011)	0.006 (0.007)	-0.026** (0.010)	0.094*** (0.020)	0.068*** (0.017)	0.055*** (0.014)
HS ²	-0.146+ (0.075)	-0.248*** (0.049)	-0.202* (0.090)	-0.002 (0.016)	0.015 (0.010)	0.054* (0.022)	-0.068** (0.022)	-0.054** (0.017)	-0.058* (0.024)
FRPM: Y	0.019 (0.026)	0.036 (0.025)	0.062** (0.022)	0.020** (0.006)	0.005 (0.005)	0.015* (0.007)	0.005 (0.012)	0.003 (0.010)	0.022** (0.008)
FRPM: U	0.152*** (0.036)	0.086** (0.027)	0.173*** (0.038)	0.001 (0.008)	0.006 (0.006)	0.000 (0.009)	a	a	a
<i>School-level</i>									
Treatment	0.016 (0.047)	0.040 (0.033)	0.027 (0.042)	0.022+ (0.011)	0.000 (0.006)	0.001 (0.009)	0.013 (0.017)	0.001 (0.015)	0.002 (0.014)
Size	0.002 (0.002)	0.002*** (0.000)	0.011 (0.012)	0.000 (0.001)	0.000 (0.000)	-0.009** (0.003)	-0.002 (0.004)	-0.002 (0.003)	-0.001 (0.003)
% FRPM	-0.003** (0.001)	-0.003*** (0.001)	-0.002+ (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)
% minority	0.000 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.001*** (0.000)	0.001*** (0.000)	0.000 (0.000)	0.000 (0.001)	0.000 (0.001)	-0.001* (0.000)
n	13,872	15,396	11,367	13,872	15,396	11,367	9,528	10,335	7,356
r ²	0.102	0.111	0.113	0.067	0.036	0.037	0.083	0.077	0.064

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals (YES or Unknown). Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ^aNot asked of students in grades 3 and 4. ES = elementary school. HS = high school. ¹White is the reference group. ²Middle school is the reference group.

Table 7.
Regression results for teacher outcomes.

	Fairness			Justness			Student Support		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-3.377*** (0.362)	-2.753*** (0.434)	-3.089*** (0.497)	-6.954*** (1.145)	-4.183*** (0.602)	-3.601** (1.352)	-5.697*** (0.570)	-4.965*** (0.900)	-4.758*** (1.148)
<i>Teacher level</i>									
Female	-0.221*** (0.053)	-0.200*** (0.058)	-0.253*** (0.069)	0.010 (0.054)	0.067 (0.067)	0.045 (0.066)	-0.169** (0.057)	-0.104 (0.070)	-0.199** (0.066)
Black ¹	0.227** (0.081)	0.103 (0.076)	0.280** (0.107)	-0.033 (0.084)	-0.004 (0.166)	0.037 (0.120)	0.019 (0.082)	0.033 (0.105)	0.077 (0.083)
Hispanic ¹	0.168 (0.111)	0.161 (0.121)	0.264+ (0.137)	-0.424** (0.136)	-0.146 (0.150)	0.118 (0.101)	0.003 (0.133)	0.217+ (0.120)	-0.090 (0.137)
Missing ¹	-0.265 (0.611)	-0.255 (0.271)	0.317 (0.275)	-0.053 (0.387)	-0.386 (0.483)	0.357 (0.220)	-0.512 (0.734)	-0.146 (0.280)	0.250 (0.198)
Other ¹	0.039 (0.108)	-0.013 (0.212)	-0.110 (0.219)	-0.106 (0.257)	0.132 (0.143)	0.298 (0.191)	-0.262 (0.183)	-0.165 (0.302)	-0.079 (0.215)
Two or more ¹	-0.096 (0.096)	-0.114 (0.108)	-0.107 (0.120)	-0.203+ (0.104)	-0.087 (0.135)	-0.131 (0.181)	-0.276** (0.094)	-0.059 (0.108)	-0.250 (0.165)
3-5 years ²	-0.097 (0.065)	-0.344*** (0.086)	-0.349*** (0.074)	0.133+ (0.073)	0.019 (0.092)	0.137 (0.095)	-0.011 (0.071)	-0.190* (0.082)	-0.237** (0.084)
6-10 years ²	-0.105 (0.071)	-0.376*** (0.089)	-0.206** (0.074)	0.145* (0.065)	0.094 (0.089)	0.244* (0.096)	-0.017 (0.079)	-0.234** (0.078)	-0.146+ (0.083)
> 10 years ²	-0.010 (0.064)	-0.273*** (0.081)	-0.129+ (0.077)	0.365*** (0.065)	0.229** (0.084)	0.398*** (0.095)	0.096 (0.069)	-0.072 (0.068)	-0.052 (0.080)
MS Teachers ³	-0.324*** (0.070)	-0.358** (0.116)	-0.346** (0.129)	-0.095 (0.058)	-0.048 (0.063)	-0.249*** (0.063)	-0.226*** (0.063)	-0.232** (0.079)	-0.301*** (0.089)
HS Teachers ³	-0.399*** (0.091)	-0.618*** (0.166)	-0.720*** (0.104)	-0.025 (0.088)	0.068 (0.088)	-0.339* (0.143)	-0.083 (0.085)	-0.327** (0.121)	-0.342*** (0.102)
Multiple Levels ³	-0.074 (0.088)	-0.239* (0.095)	-0.161+ (0.097)	-0.059 (0.085)	-0.069 (0.086)	-0.136 (0.088)	-0.145* (0.072)	-0.256** (0.091)	-0.158 (0.105)
Other Faculty/Staff ³	0.048 (0.075)	0.049 (0.101)	-0.057 (0.083)	-0.051 (0.080)	-0.076 (0.106)	-0.132 (0.085)	-0.131+ (0.079)	-0.050 (0.088)	-0.234+ (0.122)
<i>School level</i>									
Treatment	0.064 (0.058)	-0.089 (0.072)	-0.112 (0.086)	-0.003 (0.045)	0.008 (0.046)	-0.045 (0.054)	-0.081+ (0.043)	-0.100* (0.048)	-0.106 (0.066)
Size	0.006 (0.010)	0.039* (0.017)	0.039* (0.019)	-0.016 (0.010)	-0.022* (0.010)	-0.024+ (0.014)	-0.023** (0.007)	-0.006 (0.013)	-0.024 (0.015)
% FRPM	0.002+ (0.001)	0.004* (0.002)	0.006* (0.003)	0.000 (0.001)	0.001 (0.001)	-0.002 (0.002)	0.000 (0.001)	0.000 (0.001)	0.001 (0.002)
% Minority	-0.002+ (0.001)	-0.003 (0.002)	-0.001 (0.002)	0.000 (0.001)	-0.002 (0.001)	0.000 (0.002)	0.000 (0.001)	0.001 (0.001)	0.000 (0.002)
n	2,719	2,222	2,081	2,719	2,222	2,081	2,719	2,222	2,081
r ²	0.257	0.205	0.197	0.174	0.097	0.119	0.242	0.237	0.217

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, *p < .05, **p < .01, ***p < .001. FRPM = free or reduced price meals. Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ¹Reference group is White. ²Reference group is 1-3 years of experience. ³Reference group is elementary school teacher.

Table 7 (continued).
Regression results for teacher outcomes.

	Safety			Respect			PTB		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-4.252*** (0.394)	-3.148*** (0.871)	-2.874*** (0.769)	-4.474*** (0.711)	-2.877** (0.998)	-3.676** (1.331)	-2.539*** (0.155)	-2.240*** (0.210)	-2.136*** (0.230)
<i>Teacher level</i>									
Female	-0.233*** (0.047)	-0.112 (0.069)	-0.151** (0.058)	-0.178*** (0.054)	-0.116+ (0.065)	-0.184** (0.071)	0.247*** (0.065)	0.132+ (0.073)	0.098 (0.064)
Black ¹	0.328*** (0.063)	0.349*** (0.098)	0.467*** (0.081)	-0.210* (0.091)	-0.250* (0.122)	-0.187 (0.119)	-0.315** (0.101)	-0.250** (0.093)	-0.248** (0.089)
Hispanic ¹	0.076 (0.118)	0.337* (0.148)	-0.017 (0.167)	-0.112 (0.127)	0.122 (0.116)	0.105 (0.119)	-0.077 (0.136)	-0.172 (0.155)	0.010 (0.133)
Missing ¹	0.008 (0.351)	-0.613+ (0.341)	-0.316 (0.293)	-0.405 (0.698)	-0.266 (0.330)	0.308 (0.218)	-0.275 (0.555)	0.305 (0.484)	-0.130 (0.282)
Other ¹	0.020 (0.214)	0.249 (0.161)	-0.053 (0.190)	-0.066 (0.188)	0.000 (0.322)	-0.032 (0.202)	-0.221 (0.167)	-0.179 (0.197)	0.025 (0.322)
Two or more ¹	-0.020 (0.084)	-0.044 (0.153)	-0.001 (0.158)	-0.380*** (0.103)	-0.205+ (0.124)	-0.526** (0.170)	0.079 (0.098)	-0.063 (0.146)	0.018 (0.201)
3-5 years ²	-0.083 (0.074)	-0.347*** (0.068)	-0.279*** (0.075)	-0.017 (0.080)	-0.218* (0.086)	-0.145+ (0.077)	0.062 (0.084)	0.260*** (0.075)	0.263*** (0.077)
6-10 years ²	-0.078 (0.069)	-0.324*** (0.080)	-0.179* (0.079)	-0.051 (0.091)	-0.256** (0.080)	-0.114+ (0.069)	0.013 (0.071)	0.156+ (0.087)	0.069 (0.078)
> 10 years ²	0.005 (0.062)	-0.299*** (0.072)	-0.147+ (0.076)	0.092 (0.075)	-0.052 (0.074)	0.007 (0.074)	-0.077 (0.077)	0.154* (0.075)	0.031 (0.075)
MS Teachers ³	-0.213*** (0.059)	-0.214* (0.093)	-0.239* (0.116)	-0.142+ (0.076)	-0.154* (0.068)	-0.203* (0.089)	0.272*** (0.065)	0.458*** (0.128)	0.416*** (0.114)
HS Teachers ³	-0.123 (0.078)	-0.307+ (0.167)	-0.345** (0.123)	-0.065 (0.084)	-0.188+ (0.099)	-0.317** (0.102)	0.122 (0.086)	0.415*** (0.110)	0.392** (0.138)
Multiple Levels ³	-0.046 (0.073)	-0.136 (0.091)	-0.034 (0.103)	-0.079 (0.063)	-0.127 (0.085)	-0.002 (0.110)	0.162* (0.079)	0.311** (0.110)	0.254** (0.097)
Other Faculty/Staff ³	0.002 (0.061)	-0.024 (0.090)	0.088 (0.106)	-0.191* (0.091)	-0.026 (0.096)	-0.199+ (0.114)	0.061 (0.061)	0.044 (0.091)	0.068 (0.094)
<i>School level</i>									
Treatment	-0.002 (0.041)	-0.064 (0.063)	-0.041 (0.079)	-0.060 (0.046)	-0.042 (0.047)	-0.040 (0.063)	0.021 (0.041)	0.118+ (0.060)	0.105 (0.069)
Size	-0.002 (0.013)	-0.007 (0.022)	-0.007 (0.021)	-0.025** (0.008)	-0.021* (0.010)	-0.025+ (0.014)	-0.004 (0.008)	-0.001 (0.014)	0.005 (0.015)
% FRPM	0.002+ (0.001)	0.002 (0.002)	0.002 (0.003)	0.000 (0.001)	-0.001 (0.002)	0.000 (0.002)	-0.001 (0.001)	-0.001 (0.001)	-0.002 (0.002)
% Minority	-0.003** (0.001)	-0.002 (0.002)	-0.002 (0.002)	0.001 (0.001)	0.002+ (0.001)	0.001 (0.002)	0.003*** (0.001)	0.002 (0.001)	0.001 (0.002)
n	2,719	2,222	2,081	2,719	2,222	2,081	2,719	2,222	2,081
r ²	0.400	0.309	0.256	0.141	0.115	0.128	0.277	0.234	0.240

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, *p < .05, **p < .01, ***p < .001. FRPM = free or reduced price meals. Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ¹Reference group is White. ²Reference group is 1-3 years of experience. ³Reference group is elementary school teacher. PTB = prevalence of teasing and bullying.

Table 7 (continued).
Regression results for teacher outcomes.

	School problems		
	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-4.051*** (0.259)	-4.168*** (0.505)	-3.421*** (0.476)
<i>Teacher level</i>			
Female	0.190** (0.059)	0.152* (0.072)	0.058 (0.061)
Black ¹	-0.348*** (0.063)	-0.311*** (0.080)	-0.328*** (0.091)
Hispanic ¹	-0.120 (0.130)	-0.182 (0.127)	0.032 (0.131)
Missing ¹	-0.220 (0.479)	0.356 (0.405)	-0.014 (0.294)
Other ¹	-0.046 (0.217)	-0.308* (0.138)	-0.202 (0.280)
Two or more ¹	-0.058 (0.088)	-0.094 (0.129)	-0.276+ (0.145)
3-5 years ²	0.120 (0.078)	0.272*** (0.078)	0.117 (0.079)
6-10 years ²	0.049 (0.078)	0.340*** (0.091)	0.078 (0.092)
> 10 years ²	-0.046 (0.067)	0.169* (0.079)	-0.015 (0.082)
MS Teachers ³	0.200** (0.064)	0.157 (0.119)	0.241* (0.121)
HS Teachers ³	0.235* (0.100)	0.324** (0.117)	0.459*** (0.138)
Multiple Levels ³	0.108 (0.087)	0.084 (0.102)	0.070 (0.130)
Other Faculty/Staff ³	0.076 (0.064)	-0.074 (0.077)	-0.081 (0.092)
<i>School level</i>			
Treatment	0.044 (0.050)	0.095 (0.066)	0.081 (0.077)
Size	-0.003 (0.010)	-0.024+ (0.014)	-0.015 (0.014)
% FRPM	-0.001 (0.001)	-0.005+ (0.003)	-0.003 (0.002)
% Minority	0.002 (0.001)	0.002 (0.002)	0.001 (0.002)
n	2,719	2,222	2,081
r ²	0.310	0.277	0.235

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals. Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Cluster robust standard errors within parenthesis. ¹Reference group is White. ²Reference group is 1-3 years of experience. ³Reference group is elementary school teacher.

Principal Outcomes

In addition to student and teacher/staff survey responses, principals and assistant principals also responded to surveys. Note that some schools had multiple assistant principals while others did not and their responses were not required; as a result, we do not further analyze the assistant principal data. Of the 90 schools, one principal only provided responses at Time 1 and was not included in the analyses. Of the remaining schools (n = 89), 13 schools experienced principal turnover and we factor that in the analysis as well by

including a dummy coded turnover variable if the principal was new at the school. Robust standard errors were used.

Results for six outcomes are shown in Table 9. The outcomes mirror the scales asked of students and teachers/staff. An additional scale was included which was the Student Aggression Toward Teachers/Staff scale which included five items such as: “A student stole or damaged teachers’ personal property” and “A student threatened to harm a teacher.”

Table 8. Number principals and assistant principals (Ast. Prin.) surveys responses.

Cohort	Time							
	1 (Fall)		2 (Spring)		3 (Fall)		4 (Spring)	
	Principal	Ast. Prin.	Principal	Ast. Prin.	Principal	Ast. Prin.	Principal	Ast. Prin.
1: 2017-2019	21	24	20	19	20	22	20	26
2: 2018-2020	15	12	15	12	15	10	15	5
3: 2019-2021	23	26	23	27	17	20	18	23
4: 2020-2022 ^b	2	1	2	1	2	1	2	1
5: 2021-2023	22	20	22	18	18	17	16	12
6: 2022-2024	7	4	7	4	4	0	3	0
Total	90	87	89	81	76	70	74	67

Notes. ^aSpring 2020 (start of Covid). ^bSchools operating remotely. Cohorts 2, 3, and 4 were impacted by Covid (e.g., schools operating virtually).

Note that the current study has the student and teacher measures as the primary outcomes. Unlike previous results, measures are based on single informants only (i.e., the principal). Principals, as a result of the intervention, may be more aware of their environment and as a result be more critical. Regression results of six measures are shown in Table 9. Based on the results, intervention schools reported higher Prevalence of Teasing and Bullying (in Time 3) ($d = .52$, $p = .010$) and Student Aggression Toward Teachers/Staff (in Time 2) ($d = 0.39$, $p = .040$). After accounting for multiple comparisons, none of the results were statistically significant.

Table 9.
Regression results for principal outcomes.

	Disciplinary Structure			Student Support			PTB		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-6.791*** (1.367)	-6.378*** (1.285)	-6.392*** (1.434)	-5.714** (1.868)	-3.588+ (2.074)	-3.978 (2.743)	-1.752*** (0.477)	-1.525+ (0.879)	-1.658** (0.512)
Treatment	-0.160 (0.186)	-0.357 (0.224)	-0.064 (0.213)	-0.278 (0.195)	-0.430+ (0.222)	-0.110 (0.253)	0.231 (0.174)	0.519* (0.195)	0.303 (0.195)
Female	0.398 (0.272)	-0.047 (0.278)	-0.010 (0.258)	-0.252 (0.239)	-0.176 (0.259)	-0.305 (0.251)	-0.199 (0.202)	-0.195 (0.212)	-0.180 (0.210)
Black ¹	-0.183 (0.431)	0.487 (0.442)	0.211 (0.531)	0.286 (0.413)	-0.034 (0.476)	0.027 (0.509)	0.354 (0.334)	0.324 (0.447)	0.139 (0.420)
Hispanic ¹	0.659+ (0.356)	0.478 (0.596)	1.097 (0.689)	0.483 (0.387)	0.220 (0.431)	0.707* (0.330)	-1.016** (0.300)	-0.546 (0.488)	-1.171** (0.409)
Two or more ¹	0.230 (0.493)	0.814+ (0.472)	-0.081 (0.736)	0.349 (0.494)	-0.426 (0.686)	0.074 (0.700)	0.566 (0.401)	0.369 (0.490)	-1.069 (0.815)
3-5 years ²	-0.237 (0.232)	0.361 (0.375)	0.248 (0.430)	0.028 (0.298)	-0.031 (0.774)	0.044 (0.601)	-0.101 (0.227)	0.547 (0.651)	0.392 (0.378)
6-10 years ²	-0.137 (0.294)	0.683+ (0.401)	0.648 (0.402)	-0.220 (0.319)	0.389 (0.707)	0.185 (0.581)	-0.112 (0.209)	-0.009 (0.633)	-0.191 (0.328)
More than 10 years ²	0.269 (0.196)	0.387 (0.382)	0.423 (0.449)	-0.300 (0.262)	0.266 (0.829)	0.009 (0.540)	-0.118 (0.223)	0.133 (0.685)	0.421 (0.313)
Turnover	0.808+ (0.456)	0.385 (0.443)	-0.838* (0.364)	0.455 (0.425)	0.058 (0.910)	-0.675 (0.479)	-0.001 (0.289)	-0.020 (0.755)	-0.074 (0.365)
Size	0.004 (0.040)	0.046 (0.055)	-0.068 (0.054)	0.001 (0.052)	-0.069 (0.046)	-0.016 (0.064)	-0.014 (0.038)	-0.014 (0.039)	0.001 (0.039)
% FRPM	0.003 (0.006)	0.011+ (0.006)	-0.007 (0.007)	0.004 (0.005)	-0.002 (0.006)	-0.002 (0.006)	0.009+ (0.005)	0.013** (0.005)	0.004 (0.005)
% Minority	0.001 (0.005)	-0.010 (0.006)	0.009 (0.005)	0.000 (0.005)	0.008 (0.006)	0.005 (0.006)	-0.004 (0.004)	-0.014* (0.005)	-0.009+ (0.005)
Alternative ⁴	-0.413 (0.440)	0.071 (0.441)	-1.423** (0.411)	0.091 (0.426)	-0.264 (0.653)	-0.904+ (0.498)	-0.743* (0.292)	-0.586 (0.529)	0.273 (0.317)
Combined ⁴	-0.003 (0.402)	-1.486** (0.440)	-1.776* (0.786)	-0.649 (0.478)	-0.965** (0.350)	-0.747** (0.273)	-0.027 (0.227)	0.686 (0.685)	-0.380* (0.166)
High ⁴	0.015 (0.263)	0.136 (0.437)	0.582 (0.396)	-0.283 (0.459)	0.090 (0.396)	-0.416 (0.616)	-0.012 (0.287)	0.192 (0.309)	0.335 (0.406)
Middle ⁴	0.078 (0.260)	-0.065 (0.283)	0.000 (0.305)	-0.208 (0.238)	0.180 (0.289)	-0.016 (0.303)	0.243 (0.211)	0.166 (0.240)	0.358 (0.262)
n	89	76	74	89	76	74	89	76	74
r ²	0.487	0.479	0.490	0.401	0.416	0.403	0.647	0.570	0.640

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals. Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Robust standard errors within parenthesis. ¹Reference group is White. ²Reference group is 1-3 years of experience. ³Reference group is elementary school teacher. ⁴Reference group is Elementary school.

Table 9 (continued).
Regression results for principal outcomes.

	School Problems			Aggression ⁵			Safety		
	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)	Sp (T2)	Fa (T3)	Sp (T4)
(Intercept)	-2.649*** (0.556)	-0.323 (0.947)	-2.239** (0.690)	-1.071* (0.404)	-1.441* (0.678)	-0.991 (0.608)	-5.193*** (0.800)	-1.914+ (1.107)	-2.869* (1.147)
Treatment	0.189 (0.199)	0.040 (0.216)	0.184 (0.224)	0.387* (0.184)	0.150 (0.227)	0.267 (0.242)	-0.065 (0.157)	-0.335+ (0.196)	-0.121 (0.198)
Female	-0.194 (0.219)	-0.185 (0.278)	0.157 (0.240)	-0.386+ (0.227)	-0.021 (0.255)	0.159 (0.293)	-0.089 (0.204)	-0.458* (0.217)	-0.081 (0.224)
Black ¹	-0.155 (0.309)	-0.158 (0.398)	-0.501 (0.374)	-0.001 (0.327)	0.106 (0.530)	-0.903+ (0.531)	-0.068 (0.343)	0.464 (0.426)	0.227 (0.501)
Hispanic ¹	-0.444 (0.451)	0.124 (0.480)	-0.314 (0.588)	-0.697+ (0.390)	0.461 (0.293)	-0.711 (0.615)	0.796** (0.284)	0.512* (0.232)	0.409 (0.324)
Two or more ¹	0.556 (0.402)	0.431 (0.649)	-0.620 (0.642)	-0.988** (0.333)	0.743 (0.596)	-1.073 (1.101)	3.509*** (0.534)	1.064 (0.994)	2.015* (0.974)
3-5 years ²	0.363 (0.310)	-0.865 (0.541)	0.387 (0.648)	-0.241 (0.286)	0.042 (0.502)	0.230 (0.622)	-0.022 (0.273)	0.299 (0.587)	0.049 (0.370)
6-10 years ²	0.351 (0.267)	-1.319* (0.513)	-0.388 (0.604)	-0.415 (0.316)	0.066 (0.467)	-0.090 (0.526)	-0.114 (0.244)	0.737 (0.517)	0.172 (0.347)
More than 10 years ²	0.405+ (0.233)	-1.428* (0.541)	0.372 (0.624)	-0.286 (0.254)	-0.082 (0.468)	0.066 (0.513)	0.159 (0.242)	0.579 (0.610)	0.222 (0.333)
Turnover	-0.012 (0.430)	-1.084+ (0.586)	0.209 (0.418)	-0.005 (0.289)	0.187 (0.492)	-0.175 (0.487)	0.652* (0.312)	0.231 (0.572)	0.083 (0.378)
Size	0.006 (0.054)	0.050 (0.062)	0.030 (0.055)	0.006 (0.035)	0.091 (0.057)	0.064 (0.051)	0.069+ (0.036)	-0.035 (0.042)	0.050 (0.045)
% FRPM	-0.002 (0.007)	0.008 (0.006)	0.000 (0.004)	0.006 (0.004)	0.011* (0.005)	-0.001 (0.006)	0.006+ (0.004)	-0.001 (0.005)	0.005 (0.005)
% Minority	0.000 (0.005)	-0.001 (0.005)	-0.001 (0.006)	0.003 (0.004)	-0.008 (0.006)	0.000 (0.007)	-0.001 (0.004)	0.003 (0.005)	-0.004 (0.006)
Alternative ⁴	0.277 (0.351)	1.651** (0.494)	1.264*** (0.332)	0.298 (0.400)	1.221 (0.854)	0.931+ (0.509)	-0.285 (0.364)	-1.905*** (0.527)	-1.821** (0.670)
Combined ⁴	-0.002 (0.461)	0.008 (0.447)	0.466 (0.375)	-0.127 (0.466)	0.437 (0.356)	1.081** (0.317)	-0.044 (0.272)	-0.911* (0.388)	-0.602+ (0.322)
High ⁴	0.287 (0.407)	-0.172 (0.495)	0.602 (0.542)	-0.394 (0.256)	-0.701* (0.325)	-0.225 (0.375)	0.367 (0.323)	0.572 (0.391)	0.043 (0.340)
Middle ⁴	0.223 (0.268)	0.332 (0.327)	0.264 (0.266)	-0.236 (0.227)	-0.716* (0.285)	-0.344 (0.312)	-0.472* (0.228)	-0.444+ (0.249)	-0.466+ (0.253)
n	89	76	74	89	76	74	89	76	74
R2	0.479	0.459	0.498	0.591	0.438	0.457	0.637	0.561	0.618

Notes. T2 = time 2. T3 = time 3. T4 = time 4. +p < .10, * p < .05, ** p < .01, *** p < .001. FRPM = free or reduced price meals. Size is in hundreds of students. Models control for cohort effects and school-level baseline measures. Robust standard errors within parenthesis. ¹Reference group is White. ²Reference group is 1-3 years of experience. ³Reference group is elementary school teacher. ⁴Reference group is Elementary school. ⁵Student aggression toward teachers/staff.

In addition to the school related measures, principals were asked data use questions (see Appendix A). Given the limited sample size, we used exploratory factor analysis to assess which of the items formed into coherent factors. Of the 10 items, 8 items were used (two items had low loadings or double loadings) to form three factors:

- **Data importance:** two items (“I believe using data and evaluation tools leads to better student outcomes”, “I believe using data and evaluation tools is an important ongoing activity for principals”)
- **Data identification:** two items (“I can use data to identify student strengths and needs”, “I can use data to identify building strengths and needs”)
- **Data use:** four items (e.g., “I can present and interpret data with staff in my building”, “I can develop a survey and use its results to improve my school”)

Growth models assessing the three data use scales were fit. Results are shown in Table 10 and findings indicated no differential change in scores over time. A visual inspection of changes over time is also shown in Figure 1.

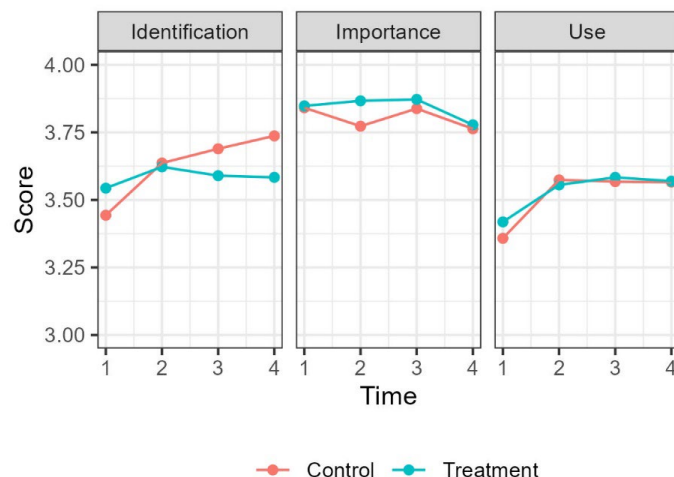
Table 10.

Growth models for data usage scales.

	Importance	Identification	Usage
(Intercept)	3.359*** (0.098)	3.391*** (0.096)	3.845*** (0.074)
Treatment	0.048 (0.134)	0.170 (0.141)	0.043 (0.097)
Time	0.063+ (0.032)	0.094** (0.031)	-0.017 (0.029)
Treatment × Time	-0.013 (0.044)	-0.085+ (0.045)	-0.002 (0.036)
n	329	329	329
R ² Adj.	0.008	0.012	-0.004

Notes. Cluster robust standard errors in parenthesis.

Figure 1. Data usage scales over time by intervention status



4. Discussion

This group-randomized trial evaluated whether the Safe and Civil Schools Leadership (SCSL) program improved school climate, discipline incidents, student behavior, and principal outcomes. The study was motivated by a substantial body of prior research suggesting that school leadership matters for student outcomes largely through indirect pathways, especially through the organizational conditions leaders help create (Leithwood et al., 2004; Louis et al., 2010; Hallinger & Heck, 1996a, 1996b, 1998; Robinson et al., 2008; Waters et al., 2003). Within this literature, school safety, student behavior, and disciplinary climate have emerged as especially important features of the school environment. Prior work identified school learning climate as a key mediator linking leadership to student learning (Bryk et al., 2010; Sebastian & Allensworth, 2012, 2013; Sebastian, Allensworth, & Huang, 2014; Sebastian, Allensworth, & Stevens, 2014), and cross-national evidence from PISA found that disciplinary climate was consistently linked to academic performance across many countries (OECD, 2010a, 2010b, 2013a, 2013b). SCSL was therefore a promising intervention because it targeted a theoretically and practically important domain of leadership practice: improving schoolwide systems for student behavior, safety, and discipline (Sprick et al., 1998).

The rationale for the trial was further strengthened by research on principals' daily work. Studies of principal practice suggest that organizational management skills, including attention to safe and orderly school environments, may be more consistently associated with school performance than many other leadership activities (Grissom & Loeb, 2011; Horng et al., 2010). At the same time, leadership preparation and professional development programs have been criticized for insufficiently preparing school leaders to manage the practical demands of student behavior, discipline, and school safety (Camburn et al., 2015; Darling-Hammond et al., 2007; Levine, 2005; Timmons, 2010). SCSL addressed a clear need in the field by providing a structured, leadership-focused model organized around distributed leadership, behavior leadership teams, and data-based decision-making (Spillane, 2006; Spillane et al., 2004, 2007).

In this group-randomized trial, the data suggested that the SCSL training program did not produce the hypothesized benefits on student, teacher, or principal outcomes. The trial involved high quality implementation of the program, so the lack of impact is not likely due to low fidelity of implementation. For instance, all intervention school leadership teams attended multiple days of training in the program by a certified SCSL trainer. They also attended 10 webinars led by the same trainer over the subsequent two years of the project. Additionally, all schools received ongoing coaching by a SCSL coach. Both the trainings and coaching were highly rated, and feedback from school leaders consistently suggested the program was beneficial to them. Despite the promise of the program and the positive feedback about it, we did not detect meaningful differences between the treatment schools and the waitlisted schools on student, teacher, or principal surveys.

The absence of statistically significant effects after adjustment for multiple comparisons suggests that participation in SCSL, as implemented in this trial, did not lead to measurable improvements in the primary student and teacher/staff outcomes. Across nine student-

reported outcomes and seven teacher/staff outcomes, there was no evidence of consistent treatment effects. The only teacher/staff outcome that reached nominal significance before adjustment was Student Support at Time 3, but the effect was small, in the opposite direction from the hypothesized benefit, and no longer statistically significant after accounting for multiple comparisons. Principal-reported outcomes showed nominally significant differences for Prevalence of Teasing and Bullying and Student Aggression Toward Teachers/Staff at selected time points, but these effects also did not remain statistically significant after multiple-comparison correction. Moreover, these principal-reported findings should be interpreted cautiously because they were based on single informants and may reflect changes in principals' awareness or interpretation of school conditions rather than changes in student behavior or school climate.

The findings are important because SCSL was a well-developed, widely disseminated, and positively received leadership professional development program. Prior non-experimental reports suggested that SCSL or related Safe and Civil Schools approaches may be associated with reductions in discipline referrals and improvements in perceptions of safety and school climate (Jacobsen & Polin, 2006; Rickert, 2005; Smolkowski et al., 2016; Ward & Gersten, 2013). The present trial contributes to the leadership and school climate literature by providing rigorous experimental evidence for a program that had previously been supported primarily by non-experimental evaluations and practitioner reports. Although prior theory and observational evidence pointed to school safety, discipline, and student behavior as promising mechanisms through which leadership could affect student outcomes (Bryk et al., 2010; Grissom & Loeb, 2011; Horng et al., 2010; OECD, 2010b, 2013a; Sebastian & Allensworth, 2012, 2013), the present findings indicate that a leadership-focused professional development and coaching model alone may not be sufficient to produce detectable changes in student, staff, or principal survey outcomes over the study period.

Several interpretations are possible. First, schoolwide discipline and climate practices may require more intensive, direct, or sustained implementation support than was provided through leadership training, webinars, and coaching. Second, leadership professional development may influence leader knowledge, attitudes, or perceived usefulness without necessarily changing the daily experiences of students and staff. Third, the pathway from principal training to leadership team action, from leadership team action to staff implementation, and from staff implementation to student experience may be long and difficult to detect, especially in a multi-cohort field trial. This interpretation is consistent with broader leadership research suggesting that principals' effects on student outcomes are indirect and mediated through complex organizational systems rather than through direct effects alone (Hallinger & Heck, 1998; Leithwood et al., 2004; Louis et al., 2010). Fourth, school climate and discipline outcomes may be influenced by many contextual factors outside the control of a leadership training program, including staffing, student needs, district policies, school resources, and broader disruptions.

It is also worth noting that the trial was impacted by COVID-19. The pandemic affected student participation in some years and created substantial disruptions for schools. However, we did not detect trends that suggested substantial differences between cohorts before and after the pandemic. Thus, while COVID-19 likely complicated implementation

and data collection, the available evidence did not indicate that the overall pattern of findings was primarily driven by pandemic-era cohort differences.

Overall, this study provides a rigorous test of a theoretically promising and widely used leadership intervention. The results do not support the conclusion that SCSL produced measurable improvements in student, teacher/staff, or principal outcomes in this trial. At the same time, the findings underscore the importance of continuing to evaluate leadership professional development programs using strong research designs, particularly because the existing evidence base for principal professional development remains limited (Camburn et al., 2015; Darling-Hammond et al., 2007). The broader literature continues to suggest that school safety, discipline, and learning climate are important for student success (OECD, 2010b, 2013a; Sebastian & Allensworth, 2012, 2013), but this trial indicates that translating that knowledge into effective principal professional development may require more intensive supports, stronger implementation mechanisms, or intervention models that more directly connect leadership practices to classroom and schoolwide behavior systems.

References

- Bosker, R. J., & Witziers, B. (1995). A meta-analytical approach regarding school effectiveness: The size of school effects and the effect size of educational leadership. Technical Report: ERIC.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.
- Camburn, E., Goldring, E., Sebastian, J., May, H., & Huff, J. (2015). An examination of the benefits, limitations, and challenges of conducting randomized experiments with principals. *Educational Administration Quarterly*.
- Clendenin, R. T. (2009). *A qualitative study of the perceived professional development needs of principals*. (Doctor of Education), University of Virginia.
- Connelly, G. (2013). Student safety - A principal's first priority. *Communicator*, 36(5).
- Darling-Hammond, L., LaPointe, M., Meyerson, D., T., O. M., & Cohen, C. (2007). *Preparing school leaders for a changing world: Lessons from exemplary leadership development programs*. Stanford, CA: Stanford Educational Leadership Institute.
- Dumay, X., Boonen, T., & Van Damme, J. (2013). Principal leadership long-term indirect effects on learning growth in mathematics. *Elementary School Journal*, 114(2), 225–251.
- Glanz, J. (1994). Redefining the roles and responsibilities of assistant principals. *Clearing House*, 67, 283–287.
- Grissom, J. A., & Loeb, S. (2011). Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills. *American Educational Research Journal*, 48(5), 1091–1123.
<https://doi.org/10.3102/0002831211402663>
- Hallinger, P., & Heck, R. H. (1996a). The principal's role in school effectiveness: A review of methodological issues, 1980-95. In K. Leithwood, J. Chapman, P. Hallinger & A. Hart (Eds.), *The international handbook of educational leadership and administration* (pp. 723–784). Dordrecht, Netherlands: Kluwer.
- Hallinger, P., & Heck, R. H. (1996b). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, 3(21), 5–44.
- Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement*, 9(2), 157–191.
- Hallinger, P., & Heck, R. H. (2010a). Collaborative leadership and school improvement: Understanding the impact on school capacity and student learning. *School Leadership & Management*, 30(2), 95–110.

- Hallinger, P., & Heck, R. H. (2010b). Leadership for learning: Does collaborative leadership make a difference in school improvement? *Educational Management Administration & Leadership*, 38(6), 654–678.
- Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal*, 46(3), 659–689.
- Heck, R. H., & Hallinger, P. (2010). Collaborative leadership effects on school improvement: Integrating unidirectional- and reciprocal-effects models. *Elementary School Journal*, 111(2), 226–252.
- Hornig, E. L., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education*, 116(4), 491–523.
- Jacobsen, M. D., & Polin, M. R. (2006). A district's role in building safe and civil schools. *Principal Leadership*, 7(4), 36–40.
- Kealey, R. J. (2002). *Some Things Assistant Principals Do*. Washington, DC: National Catholic Educational Association.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. Wallace Foundation.
- Levine, A. (2005). *Educating school leaders*. New York: The Education School Project.
- Lisle, E. (2002). *Safe schools: Staff development training opportunities*. (Doctor of Education), University of Florida.
- Louis, K. S., Leithwood, K., Wahlstrom, K., & Anderson, S. (2010). *Learning from leadership project: Investigating the links to improved student learning*. <http://www.wallacefoundation.org/KnowledgeCenter/KnowledgeTopics/CurrentAreasofFocus/EducationLeadership/Documents/Learning-from-Leadership-Investigating-Links-Final-Report.pdf>
- Neumerski, C. M. (2013). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly*, 49(2), 310–347. <https://doi.org/10.1177/0013161X12456700>
- OECD. (2010a). *PISA 2009 results: Executive summary*. Paris, France: Organization for Economic Co-operation and Development.
- OECD. (2010b). *PISA 2009 Results: What makes a school successful? - Resources, Policies and Practices*. Paris, France: Organization for Economic Co-operation and Development.
- OECD. (2013a). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)*. Paris, France: PISA, OECD Publishing.
- OECD. (2013b). *PISA 2012 Results: What students know and can do – student performance in mathematics, reading and science (Volume I)*. Paris, France: PISA, OECD Publishing.

- Oleszewski, A., Shoho, A., & Barnett, B. (2012). The development of assistant principals: A literature review. *Journal of Educational Administration, 50*(3), 264–286.
- Rickert, C. S. (2005). A blueprint for safe and civil schools. *Principal Leadership, 6*(1), 44–46, 48–49.
- Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Educational Administration Quarterly, 44*(5), 635–674.
- Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly, 48*(4), 626–663.
- Sebastian, J., & Allensworth, E. (2013). How do secondary principals influence teaching and learning? *Principal's Research Review, 8*(4).
- Sebastian, J., Allensworth, E., & Huang, H. (2014). The role of teacher leadership in how principals influence classroom instruction and student learning. Paper presented at the University Council of Educational Administration, Washington, DC.
- Sebastian, J., Allensworth, E., & Stevens, D. (2014). The influence of school leadership on classroom instruction: Examining configurations of organizational supports. *Teachers College Record*.
- Sheras, P. L., Cornell, D. G., & Bostain, D. S. (1996). The Virginia Youth Violence Project: Transmitting psychological knowledge on youth violence to schools and communities. *Professional Psychology: Research and Practice, 27*(4), 401–406.
<https://doi.org/10.1037/0735-7028.27.4.401>
- Smolkowski, K., Strycker, L., & Ward, B. (2016). Scale-up of Safe & Civil Schools' model for school-wide positive behavioral interventions and supports. *Psychology in the Schools, 53*(4), 339–358.
- Spillane, J. P. (2006). *Distributed leadership*. San Francisco, CA: Jossey Bass.
- Spillane, J. P., Camburn, E., & Pareja, A. S. (2007). Taking a distributed perspective to the school principal's workday. *Leadership and Policy in Schools, 6*(1), 103–125.
- Spillane, J. P., Halverson, R., & Diamond, J. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies, 36*(1), 3–34.
- Sprick, R., Howard, L., Wise, B. J., Marcum, K., & Haykin, M. (1998). *Administrator's desk reference of behavior management: Leadership guide*. Longmont, CO: Sporis West.
- Timmons, S. L. (2010). *Principal perceptions of training needs in school safety in Virginia*. (Doctor of Education), Virginia Polytechnic Institute and State University.
- Ward, B., & Gersten, R. (2013). A randomized evaluation of the Safe and Civil Schools model for positive behavioral interventions and supports at elementary schools in a large urban school district. *School Psychology Review, 42*(3), 317–333.

Waters, T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 Years of research tells us about the effect of leadership on student achievement. A working Paper*. Aurora C. O.: Mid-Continent Regional Educational Lab.

Witziers, B., Bosker, R. J., & Kruger, M. L. (2003). Educational Leadership and Student Achievement: The Elusive Search for an Association. *Educational Administration Quarterly*, 39(3), 398–425.