Integrated Technical Report: Randomized Evaluations of Safe & Civil School Leadership and Safe & Civil School Leadership Plus Start on Time

Prepared by

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

June 2026

Take Home Message:

Principal-focused leadership training was promising and well received across both studies, but neither SCSL alone nor SCSL plus START reliably translated into measurable improvements in the school climate and safety experiences reported by teachers and students.

Abstract

This integrated technical report summarizes findings from two randomized evaluations of Safe & Civil School Leadership (SCSL), a leadership training and coaching model designed to help school leaders improve school safety, student behavior, disciplinary systems, and school climate (Sprick, Howard, Wise, Marcum, & Haykin, 1998). The first evaluation tested SCSL as a stand-alone leadership professional development program across 90 schools recruited over six cohorts. The second evaluation tested SCSL plus Start on Time (START), a companion tardiness-reduction and hallway-management intervention, across 43 schools in Oklahoma. Both studies were motivated by evidence that school leaders influence student outcomes indirectly through the organizational conditions they help create, particularly school safety, disciplinary climate, and student behavior (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Sebastian & Allensworth, 2012, 2013; Sebastian, Allensworth, & Huang, 2014; Sebastian, Allensworth, & Stevens, 2014). Across both evaluations, schools were randomly assigned to intervention or business-as-usual control conditions, and outcomes were assessed using student, teacher/staff, and administrator surveys. The SCSL-only evaluation included 59,863 student responses from 83 schools and 10,588 teacher/staff responses. The SCSL plus START evaluation included 265 administrator responses, 4,141 teacher responses, and 32,861 student responses from 43 schools. Results across both studies provided limited evidence of intervention impact. The SCSL-only trial did not produce statistically significant improvements in student, teacher/staff, or principal outcomes after accounting for multiple comparisons. In the SCSL plus START trial, administrator reports showed some positive effects on disciplinary structure, but these effects were not reflected in teacher or student reports; teacher outcomes showed no significant treatment effects, and student outcomes showed no positive effects. Together, the findings suggest that although SCSL and SCSL plus START were promising, well-received, and supported through training and coaching, principal-focused leadership interventions did not reliably translate into measurable improvements in the school climate and safety experiences reported by teachers and students.

1. Introduction

Schools are expected to provide safe, orderly, and supportive environments for learning, yet many schools continue to struggle with discipline incidents, bullying, student victimization, and perceptions of school safety (Every Student Succeeds Act, 2015; National Center for Education Statistics, 2015). Prior research suggests that school leadership plays a central role in shaping these conditions. School leaders influence student outcomes indirectly through the organizational systems they help create, including professional capacity, family and community ties, school learning climate, school safety, and student behavior (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Louis, Leithwood, Wahlstrom, & Anderson, 2010). Within this literature, school safety and disciplinary climate have emerged as especially important school-level conditions linking leadership to student outcomes (Sebastian & Allensworth, 2012, 2013; Sebastian et al., 2014; Sebastian, Allensworth, & Stevens, 2014). Cross-national evidence from PISA also suggests that disciplinary climate is consistently associated with academic performance across countries (OECD, 2010a, 2010b, 2013a, 2013b).

Despite the importance of school safety and discipline as leadership responsibilities, many school leaders report feeling underprepared to address student misbehavior, school discipline, and schoolwide behavior systems (Ricciardi & Petrosko, 2000; Timmons, 2010). Leadership preparation and professional development programs often focus broadly on instructional leadership or general school improvement and may give less explicit attention to school safety, disciplinary structure, hallway behavior, and student conduct (Camburn, Goldring, Sebastian, May, & Huff, 2015; Darling-Hammond, LaPointe, Meyerson, T., & Cohen, 2007). This gap is important because principals and assistant principals often spend substantial time managing student behavior and promoting safe school environments (Glanz, 1994; Kealey, 2002; Oleszewski, Shoho, & Barnett, 2012).

Safe & Civil School Leadership (SCSL) was developed to address this need. 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 et al., 1998). The program emphasizes distributed leadership, data-based decision-making, schoolwide behavior planning, and leadership team implementation, consistent with distributed leadership perspectives in school leadership research (Spillane, 2006; Spillane, Camburn, & Pareja, 2007; Spillane, Halverson, & Diamond, 2004). Its guiding STOIC principles include: Structure and organize all settings for success; Teach students how to behave responsibly; Observe student behavior; Interact positively with students; and Correct irresponsible behavior fluently.

The second evaluation examined SCSL in combination with Start on Time (START). START is a companion program in the broader Safe & Civil Schools series that focuses specifically on reducing tardiness and improving schoolwide hallway management. START targets a visible and valued student behavior--arriving to class on time--through a step-by-step process involving all school staff. By helping schools create a unified approach to hallway transitions and tardiness reduction, START is intended to build staff buy-in and support broader SCSL efforts to create safe, predictable, and nurturing school climates. Prior non-experimental reports suggested that SCSL and related Safe & Civil Schools models may be associated

with improvements in school safety, student relationships, discipline referrals, and school climate (Jacobsen & Polin, 2006; Rickert, 2005; Smolkowski, Strycker, & Ward, 2016; Ward & Gersten, 2013).

The two studies summarized in this report therefore address a common question: Can principal-focused leadership training and coaching improve school climate, safety, student behavior, and related outcomes? The first study evaluated SCSL as a stand-alone intervention. The second evaluated SCSL plus START, adding a targeted tardiness-reduction and hallway-management component.

2. Program Descriptions

Safe & Civil School Leadership

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 et al., 1998). Certified trainers provide training and ongoing coaching to school leaders. SCSL is grounded in social learning theory and includes research-based principles including data-based decision-making. The SCSL model targets school leaders' use of effective schoolwide discipline practices by promoting positive relationships.

SCSL is based on a cyclical process of reviewing data from multiple sources, revising the Schoolwide Behavior Plan based on that review, creating improved policies and procedures, implementing new policies, and maintaining effective current policies. These steps are taken collaboratively by the school Behavior Leadership Team, which includes the principal and various school staff involved in supporting student behavior success.

SCSL topics include leadership skills, schoolwide behavior planning, behavior leadership teams, data-based decision-making, instructional approaches to behavior, common area behavior, classroom management support, discipline referrals and solutions, and meaningful work for students.

Start on Time

START is a companion program to SCSL that focuses on schoolwide hallway management and tardiness reduction. START is designed to help principals and school staff implement a unified approach to improving on-time arrival to class. The rationale for START is that tardiness is a concrete, observable, and schoolwide behavior that can be addressed through relatively simple staff actions. Successful implementation of START may help principals build staff confidence and commitment to broader school climate and discipline strategies.

In the SCSL plus START evaluation, START was implemented alongside SCSL to test whether adding a targeted schoolwide behavior intervention strengthened effects on school climate, safety, disciplinary structure, and student behavior.

3. Study 1: Evaluation of SCSL Alone

Design and Participants

A total of 90 schools were recruited over six cohorts to be part of the SCSL-only project. Approximately half of the schools were randomly assigned to the intervention condition and the other half were 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 in grades K-2 and were not eligible for the 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 indicated that they were not responding truthfully or were classified as inattentive responders. This reduced the sample to 61,278 students. 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). In SY 2019-2020, five schools from one district did not collect student responses in Spring 2020 due to COVID and did not resume data collection; an additional 1,415 students were excluded. For analyses using student data, all other schools with two or more data collection points were included, resulting in 59,863 student responses from 83 schools.

For the teacher/staff survey, 12,894 responses were initially collected. Responses that were too quick, inattentive, or empty were excluded, resulting in a final analytic sample of 10,588 teacher/staff responses. Differences between treatment and control conditions were not meaningfully different, with small effect size differences.

Measures

Student measures included disciplinary structure, student support, safety, school problems, bullying victimization, bullying others, school belonging, suspension, and fighting. Scales were adapted from the Authoritative School Climate Survey, with shortened forms used for students in grades 3 and 4 (Konold et al., 2014).

Teacher/staff measures included justness, fairness, student support, safety, respect for students, prevalence of teasing and bullying, and school problems. These scales were drawn from the Authoritative School Climate Survey-Teacher/Staff version (Huang et al., 2015).

Principal and assistant principal measures included scales paralleling the student and teacher/staff outcomes, along with additional measures of data importance, data identification, and data use.

Analytic Strategy

Given the seasonality of school climate, the study used an ANCOVA approach to evaluate outcomes separately in Spring, Fall, and Spring follow-up periods while controlling for

school-aggregated baseline measures from Fall (Huang et al., 2023). Both individual and school-level covariates were included in all models. The primary coefficient of interest was 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). Outcomes were standardized so treatment effect coefficients could be interpreted as standardized mean differences. Multiple comparisons were addressed using a Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Results

Across nine student outcomes during the follow-up periods, none of the student-reported outcomes were statistically significant using an alpha of .05. For the seven teacher/staff outcomes, almost all treatment effects were not statistically significant. The exception was Student Support at Time 3, where intervention schools had slightly lower reported measures compared to control schools; however, after adjustment for multiple comparisons, this difference was not statistically significant.

Principal outcomes also did not show consistent evidence of intervention impact. Intervention schools reported higher prevalence of teasing and bullying at one time point and higher student aggression toward teachers/staff at another time point, but these findings did not remain statistically significant after accounting for multiple comparisons. Growth models assessing data importance, data identification, and data use indicated no differential change in scores over time.

Summary of Study 1

The SCSL-only trial did not produce the hypothesized benefits on student, teacher/staff, or principal outcomes. The program was implemented with substantial training and coaching support, and feedback from school leaders suggested that the program was useful and well received. However, the intervention did not produce meaningful differences between treatment and waitlisted schools on student, teacher/staff, or principal surveys. Although COVID-19 affected some data collection and school operations, the overall pattern of findings did not indicate clear or consistent intervention effects.

---

4. Study 2: Evaluation of SCSL Plus START

Design and Participants

The second study evaluated SCSL plus START using a group randomized design. Forty-three schools in Oklahoma were randomly assigned to receive SCSL plus START or to a business-as-usual control condition. Data on primary outcomes and proposed mechanisms were collected at baseline and at 6-month, 1-year, and 2-year follow-up periods.

The study included 43 public schools serving students in grades 5 through 12. Twenty-three schools were assigned to the treatment group and 20 schools were assigned to the control group. The sample included 25 middle schools, 11 high schools, and 7 schools serving

multiple levels. Most schools were located in rural settings, with additional schools in suburban and urban contexts.

After deleting invalid responses, the analytic sample included 265 principal and assistant principal responses, 4,141 teacher responses, and 32,861 student responses from 43 schools. The student sample was diverse and included substantial representation of students eligible for free or reduced-price lunch. Student response rates were affected by COVID-19, especially during Spring 2020 when many schools shut down or moved to virtual learning platforms.

Intervention Supports

The SCSL plus START intervention included principal training workshops, online training materials, and ongoing coaching. A certified trainer trained principals from SCSL schools in two full-day workshops. Participants also had access to online materials designed to supplement and enhance skill development. A certified SCSL coach provided onsite coaching to intervention schools, prompted principals to complete monthly assignments, and met with leadership teams at least twice each year to review workshop content, assess implementation challenges, and problem solve barriers.

Implementation supports included detailed training manuals, session-by-session protocols, online resources monitored for use and activity, onsite coaching, a buddy system to help principals support one another, and independent site observations with performance feedback. Workshop integrity was supported through supervision by Dr. Sprick, trainer checklists, and feedback on trainer performance.

Measures

Measures included student, teacher, and administrator reports of school climate and safety. Key constructs included prevalence of teasing and bullying, victimization, disciplinary structure, academic expectations, student support, perceptions of school safety, aggressive attitudes, use of data, and school problems. Many measures were drawn from the Authoritative School Climate Survey and related validated school climate scales (Huang et al., 2015; Huang, Cornell, & Konold, 2014; Konold et al., 2014). Additional items from the Trends Climate Survey were used because the survey was developed specifically as an online data tool for the SCSL intervention (Sprick et al., 1998; Sprick et al., 2015).

Analytic Strategy

Because of the nested data structure, random intercept multilevel models were used (Raudenbush & Bryk, 2002). Analyses were conducted using R with restricted maximum likelihood estimation. Three regression models were modeled separately for three follow-up time points while controlling for the baseline measure at the first spring, aggregated at the school level. Models included treatment condition, individual demographic variables, school characteristics, and cohort indicators. Because outcomes were standardized, treatment effect coefficients can be interpreted as effect sizes.

Results

Administrator Outcomes

Administrator reports showed statistically significant positive intervention effects on school disciplinary structure at the second and third time periods. However, no statistically significant effects were found for perceptions of school safety, academic expectations, or school problems. A negative treatment effect was found at the fourth time period on perceived student support.

Secondary administrator outcomes showed no statistically significant intervention effects for concerns about safety and discipline, student aggression toward teachers, prevalence of teasing and bullying, or willingness to seek help. A decline in respect for students was found in the intervention group at the final time point.

Teacher Outcomes

Teacher-reported outcomes showed no significant treatment effects for the five main outcomes: school disciplinary structure, perceptions of school safety, academic expectations, school problems, and student support. Two school discipline subscales, justness and fairness, also showed no statistically significant treatment effects. Secondary school climate outcomes showed no statistically significant treatment effects across all time periods after controlling for covariates.

Student Outcomes

Student-reported outcomes also showed limited evidence of intervention impact. There were no statistically significant treatment effects on perceptions of school safety, academic expectations, school problems, or aggressive attitudes. Negative treatment effects were found for school disciplinary structure and student support at the second spring follow-up period.

Summary of Study 2

The SCSL plus START study provided limited evidence of intervention impact. Administrator reports suggested some improvement in disciplinary structure during two follow-up periods, but these effects were not mirrored by teacher or student reports. Teacher outcomes showed no significant treatment effects, and student outcomes showed no positive effects. The student findings included negative effects on disciplinary structure and student support at one time point. These findings should be interpreted cautiously given the influence of COVID-19 on implementation, school operations, and survey response patterns.

---

5. Integrated Discussion

Taken together, the two randomized evaluations provide a rigorous test of principal-focused leadership interventions intended to improve school safety, disciplinary structure, student

behavior, and school climate. Both studies were grounded in a strong theoretical and empirical rationale. Prior research suggests that school leaders influence student outcomes through organizational conditions, and school safety and disciplinary climate have been identified as key pathways through which leadership may matter (Bryk et al., 2010; Sebastian & Allensworth, 2012, 2013; Sebastian et al., 2014; Sebastian, Allensworth, & Stevens, 2014). SCSL was designed to strengthen those pathways through leadership training, data-based decision-making, schoolwide behavior planning, and coaching (Sprick et al., 1998). The SCSL plus START study extended this model by adding a targeted tardiness-reduction and hallway-management intervention intended to create an early, concrete behavior change and build staff buy-in.

Across both studies, however, the results did not provide consistent evidence that these interventions improved the school climate and safety experiences reported by teachers and students. In the SCSL-only evaluation, no meaningful intervention effects were found across student, teacher/staff, or principal outcomes after accounting for multiple comparisons. In the SCSL plus START evaluation, administrator-reported disciplinary structure improved at some time points, but teacher and student reports did not show corresponding improvement. Teacher outcomes were consistently null, and student outcomes were either null or negative at isolated time points.

These findings suggest a critical distinction between leadership training that is well received and leadership training that changes the daily experiences of students and staff. In both studies, the intervention models were coherent, practical, and supported through training and coaching. Yet positive reception by school leaders did not reliably translate into measurable improvements in school climate, safety, disciplinary structure, or student support. This pattern underscores the complexity of schoolwide change. Principal knowledge and leadership team planning may be necessary but not sufficient to alter staff behavior, student routines, hallway interactions, disciplinary consistency, or relational climate.

One explanation is that the pathway from principal training to student outcomes is long and indirect, consistent with prior leadership research emphasizing mediated and organizational pathways between leadership and student outcomes (Leithwood et al., 2004; Louis et al., 2010; Sebastian & Allensworth, 2012). The intervention must first change principal knowledge, priorities, and leadership behavior. Those changes must then influence leadership team functioning, staff expectations, schoolwide procedures, adult consistency, and daily student interactions. Only after those steps might students and teachers perceive improvements in school climate and safety. Even a well-designed intervention may produce limited effects if implementation weakens at any point in this chain.

A second explanation is that principal-focused interventions may require more direct staff-level implementation supports. SCSL and SCSL plus START relied on principals and leadership teams to guide schoolwide change. However, teacher and student outcomes depend heavily on what adults across the school do consistently in classrooms, hallways, cafeterias, and other common areas. If staff do not implement practices consistently, students may not experience meaningful changes in structure, support, or safety. Future models may need to include more intensive teacher and staff coaching, clearer

implementation benchmarks, stronger feedback systems, and more direct support for translating leadership plans into daily routines.

A third explanation concerns informant differences. Administrators may notice changes in policies, procedures, planning routines, or leadership team processes, whereas teachers and students may report whether those changes are visible in everyday interactions. In the SCSL plus START study, the positive administrator-reported effects on disciplinary structure, coupled with null teacher and student findings, may indicate that leaders perceived improvements in systems that were not yet fully experienced by staff or students. Multi-informant measurement is therefore essential for evaluating school climate interventions.

A fourth issue is the role of COVID-19. Both evaluations were affected by pandemic-era disruptions, though the SCSL plus START study was especially shaped by data collection from 2017 to 2022. The pandemic disrupted school routines, hallway transitions, staff consistency, student attendance, and the meaning of school climate itself. Because START was specifically focused on tardiness and hallway management, disruptions to in-person schooling may have reduced the opportunity for full implementation. COVID-19 also affected survey participation, especially in Spring 2020. These disruptions complicate interpretation and may have reduced the likelihood of detecting intervention effects.

Despite these limitations, the combined findings make an important contribution. They show the value of testing widely used and theoretically promising school leadership programs through rigorous randomized designs. The results caution against assuming that positive participant feedback, strong program materials, or high conceptual alignment with the literature will necessarily produce measurable changes in student and staff outcomes. The findings also suggest that improving school safety and climate may require more intensive, sustained, and systemwide intervention models that go beyond principal training.

6. Implications for Practice

The results suggest that leadership training alone may not be enough to produce visible improvements in the daily experiences of teachers and students. Schools seeking to improve safety, disciplinary structure, and student behavior may need interventions that combine principal leadership development with direct staff training, classroom and common-area coaching, implementation monitoring, and continuous feedback from students and teachers.

For programs such as SCSL and SCSL plus START, future implementation efforts may benefit from stronger mechanisms for ensuring that leadership team decisions are enacted consistently across school settings. START may remain conceptually useful because tardiness is visible, measurable, and connected to hallway management, but it may require strong staff participation and stable in-person routines to produce schoolwide effects.

Implications for Research

Future research should examine implementation variability across schools to determine whether stronger implementation is associated with more favorable outcomes. Analyses should also test whether changes in leadership skills, data use, disciplinary structure, student support, and staff consistency mediate intervention effects. Given the COVID-19 disruptions, sensitivity analyses by cohort, timing, and level of pandemic exposure may help clarify whether results differed before, during, and after pandemic-related disruptions.

Future intervention studies may also need to measure intermediate implementation processes more directly, including staff adoption, hallway supervision practices, classroom management consistency, leadership team functioning, and student exposure to changed routines. These measures may help explain whether null findings reflect failure of the intervention theory, insufficient implementation, insufficient dosage, or contextual disruption.

Conclusion

Across two randomized evaluations, SCSL alone and SCSL plus START did not produce consistent improvements in student, teacher/staff, or administrator-reported school climate and safety outcomes. The second study extended the first by adding START, a tardiness-reduction and hallway-management intervention, but the combined model still did not yield consistent positive effects as reported by teachers and students. The findings underscore a central lesson for school improvement: training principals is important, but changing school climate requires sustained, systemwide implementation that reaches the everyday routines, adult behaviors, and student experiences that define life in schools.

References

Astor, R. A., Benbenishty, R., & Estrada, J. N. (2009). School violence and theoretically atypical schools: The principal's centrality in orchestrating safe schools. *American Educational Research Journal, 46*(2), 423–461.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).

Gottfredson, G. D., Gottfredson, D. C., Payne, A. A., & Gottfredson, N. C. (2005). School climate predictors of school disorder: Results from a national study of delinquency prevention in schools. *Journal of Research in Crime and Delinquency, 42*(4), 412–444.

Musu-Gillette, L., Zhang, A., Wang, K., Zhang, J., & Oudekerk, B. A. (2016). *Indicators of school crime and safety: 2015* (NCES 2016-079/NCJ 249758). National Center for Education Statistics, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.

Ricciardi, D., & Petrosko, J. M. (2000). *Perceptions of first-year administrators: Impact of responsibilities and preparation on professional growth needs*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

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. (2016). The role of teacher leadership in how principals influence classroom instruction and student learning. *American Journal of Education, 123*(1), 69–108.

Sebastian, J., Allensworth, E., & Stevens, D. (2014). The influence of school leadership on classroom participation: Examining configurations of organizational supports. *Teachers College Record, 116*(8), 1–36.

Timmons, S. L. (2010). *Principal perceptions of training needs in school safety in Virginia* (Doctoral dissertation, Virginia Polytechnic Institute and State University).