National Institute of Justice

Developing Knowledge about What Works for Make Schools Safe: Implementation and Evaluation of Tools for Life in Jackson Public School District in Mississippi

Draft Final Summary Overview submitted by:

Gabriella C. Gonzalez, Senior Social Scientist, Principal Investigator
ggonzal@rand.org 1-412-683-2300 x 4426

RAND Corporation
1776 Main Street
Santa Monica, CA 90401
DUNS: 00-691-4071
TIN: 95-1958142

Grant No 2015-CK-BX-0005
Project/Grant Period: September, 2019-December 31,2019

Report Submitted: April 10, 2020

Tonya Bordonaro
Contract Administrator
RAND Corporation
4570 Fifth Avenue
Pittsburgh, PA 15213-2665
412 683-2300
tbordano@rand.org

RAND Social and Economic Well-being

A. Purpose of the Project

Teachers and administrators across the United States struggle to maintain safe, civil, community- and achievement-oriented schools. Disciplinary actions (such as school suspensions) and security measures (such as school resource officers/security guards and security cameras) are some approaches commonly used in schools to try to improve school climate and safety. Another approach is to improve the socioemotional skills of students and educators to foster a more positive school climate and build positive relationships among all individuals in schools.

The purpose of this project was to implement a social and emotional learning (SEL) program in a school district and evaluate its effects on school climate and safety throughout the district. Jackson Public School District, in Jackson, Mississippi, partnered with the Tools for Life® Corporation to implement its SEL curriculum in randomly selected elementary and middle schools in the district in the 2016–2017 and 2017–2018 school years. The RAND Corporation served as evaluators for the project, documenting the implementation of Tools for Life and its costs and assessing the program's impact on school climate and safety. This study represents the first district-wide randomized controlled trial of the Tools for Life program.

This study contributes to the education field's understanding of the extent to which SEL can improve school climate, and thus school safety, by empirically testing this relationship. JPSD was an ideal district within which to test TFL: The Jackson Public School District recognized its need to improve school climate and school safety, as demonstrated in the timeline in Table 1.

Table 1. Timeline of Jackson Public School District School Safety Issues (2015–2017)

	Date	Event
Prior to TFL implementation	2015–16	Mississippi Department of Education (MDE) goves JPSD an "F" performance rating based on a statewide accountability system that takes into account (a) student achievement, (b) student growth, (c) graduation rates, (d) participation rates, and (e) other outcome measures.
	April-August 2016	MDE conducts a limited audit of JPSD (based on a state law that requires schools with a C, D, or F rating to be audited periodically). Based on what was deemed to be concerning findings, JPSD's accreditation was placed on "probation."
Year 1 of TFL Implementation (2016–2017)	September 6, 2016–July 31, 2017	MDE conducts a second, more comprehensive audit, which finds JPSD in violation of 24 out of 32 process standards, including those related to school safety and student discipline.
	November 2016	JPSD superintendent Cedrick Gray resigns, and the JPSD school board selects Freddrick Murray, chief academic officer of high schools in JPSD and long-time educator and administrator in the district, as interim superintendent. JPSD files a Corrective Action Plan (CAP) with MDE to address the limited audit and is rejected by the State Board of Education due to lack of specifics.
	December 2016	JPSD files a revised Corrective Action Plan with MDE and is approved by the State Board of Education.
Year 2 of TFL Implementation (2017–2018)	September 13–14, 2017	The Commission on School Accreditation (a statewide representative body that makes recommendations to the state school board; Mississippi Code 37–17-3) meets to review the results of the audit and decide whether to designate emergency status and recommend that the district be placed under state control. Instead, the Commission of School Accreditation, and following them, the MDE declared an "extreme emergency situation" due to safety and educational concerns and requested the state appoint their own interim superintendent.
	September	A federal lawsuit is filed by an attorney on behalf of 23 JPSD parents, alleging that

	18, 2017	the decision to take over the district violated constitutional due-process rights. The lawsuit asserts that JPSD did not have adequate time to follow through on the CAP before the second audit was conducted, saying that some CAP items had a one year timeline. It also asserts that "because the audit was contemporaneous with the Corrective Action Plan already prepared and executed by the District, a number of the audit's 'findings' describe conditions existing before corrective action was undertaken."
	September 20, 2017	The JPSD school board, facing the resignation of four out of seven members, decides not to initiate the search for a new superintendent, and instead keep the interim superintendent in charge of JPSD, Freddrick Murray.
	October 2017	Mississippi Governor Phil Bryant does not sign the resolution and instead selects a third option—one that would be neither a state takeover or allowing JPSD to maintain full control. This third option was to form a 15-member Better Together Commission (BTC) to oversee the transformation of JPSD, with five seats appointed by the governor, 5 by Jackson city and five by W.K. Kellogg Foundation.

B. Project Design, Methods, and Data Analysis

This study had three objectives:

1. Describe the implementation of TFL in JPSD.
2. Assess the degree to which TFL is effective in improving school climate and school safety outcomes.
3. Measure the costs of implementing TFL in JPSD.

To meet these objectives, RAND researchers conducted a cluster randomized trial for students in grades 3 through 8 attending traditional public elementary and middle schools in JPSD. Twenty-three schools were randomly assigned to implement TFL in academic year 2016–2017 (treatment schools), and 22 schools were randomly assigned to continue to conduct business-as-usual and to delay implementation of TFL until academic year 2017–2018 (control schools).

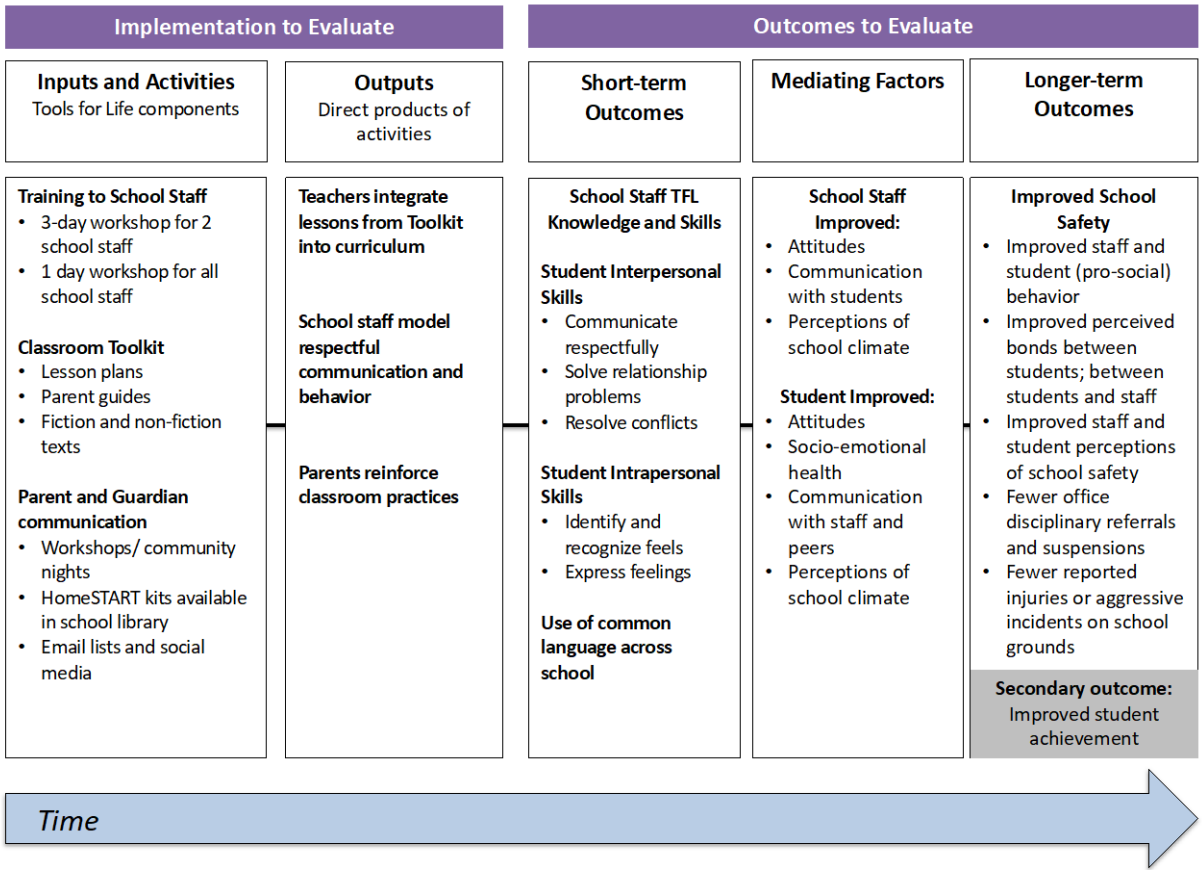
To provide a descriptive portrait of the implementation of TFL in JPSD in the 2016–2017 and 2017–2018 academic school years, we relied on data from interviews with JPSD district and Mississippi Department of Education (MDE) officials, district-level TFL implementers, and Tools for Life Corporation coaches; results from surveys administered to students in grades 3 through 8; and interviews with school leaders and focus groups with instructional staff collected from six purposefully selected treatment and control schools. To assess the impact of TFL on school and students' outcomes across both years of the study, we used data from the student surveys and JPSD administrative data. We also documented the costs of TFL in JPSD, relying on JPSD administrative data and financial records from TFL and JPSD.

Analytic Approach

The theory of change illustrated in Figure 1 served as the study's framework for analysis. This theory of change posits that TFL inputs and activities (school staff training provided by TFL, classroom toolkits with complementary practices for home) and outputs (the implementation of TFL) will increase short-term outcomes (school staff and students' knowledge and skills). School staff will develop an in-depth comprehension of TFL techniques, and students will increase their intrapersonal (identify and recognize feelings, express feelings) and interpersonal skills (communicate respectfully, solve relationship problems, resolve conflicts). All members of the school community will start using a common language vis-à-vis relationship-building. With increases in TFL-related school staff and student knowledge and skills, we hypothesize that

mediating factors (staff and student attitudes and students' socioemotional health, such as self-regulation and empathy) will change. Ultimately, short-term outcomes and mediating factors produce positive longer-term outcomes: increased perceptions of school safety among students and staff; stronger perceived bonds and improved school climate; increases in teacher and student attendance; increases in school safety as measured by observations of student and staff behavior and reduction in rates of office disciplinary referrals, out-of-school suspensions, and reported injury and aggressive incidents. Though not a primary outcome of interest, we also investigated the effect of TFL on student achievement (as measured through results on Mississippi state assessments) as a secondary, exploratory, analysis. We hypothesized that with a greater sense of community and connection among students and teachers, students would become more engaged in instruction, leading to gains in achievement. Decades of research have demonstrated that when students are more engaged in instruction they tend to have higher academic achievement (Fredericks et al., 2004).

Figure 1. Theory of Change for Implementation of TFL to Improve School Climate and Safety



This study took place over two academic years (2016–2017 and 2017–2018). In the first year of implementation, 45 schools were randomly assigned to treatment and control conditions. We first stratified the participating sample into elementary schools and middle schools. Then, the stratified schools were randomly assigned to treatment and control groups. Beginning in October

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

2016, treatment schools began implementing TFL programming in grades 4 through 8. After the first year (2016–2017), the 22 schools in the control group implemented TFL.

Of the 2,941 students whose parents or guardians consented that they could participate in data-collection activities associated with the study, 2,740 were enrolled in schools that were randomly assigned to treatment or control conditions. 1,422 were enrolled in treatment schools at the start of the study, and 1,318 were enrolled in control schools at the start of the study. Before conducting any analyses, we examined the extent to which the study sample was *balanced* (i.e., had similar background characteristics across intervention and control conditions) and *representative* (i.e., those who consented to participate in data collection had similar background characteristics as those who did not). We found that, though the sample was balanced and there were no significant differences in the characteristics of students in the experiment by treatment group, the sample was not representative: Students whose parents or guardians did not consent to data-collection activities tended to have lower achievement and were also more likely to have discipline issues and higher absenteeism.

To evaluate the impact of TFL, we conducted a cluster-randomized trial in two stages. Stage I examined impacts after one academic year of exposure to TFL. Stage II examined trajectories for students' outcomes from over the course of two academic years using a two-phase multiple baseline design, which is a within-person repeated-measures design that can be used to appraise the extent to which an intervention alters growth trajectories. In the focus groups, we allowed for open discussion among participants. To document the implementation of TFL, we relied on data from interviews with key JPSD, MDE, and TFL staff; analysis of the student survey; and data from the study of three treatment and three control schools, which involved focus groups with school instructional staff and interviews with school leaders.

Data Sources

1. *Interviews with Key Jackson Public School District, Tools for Life, and Mississippi Department of Education Officials*

RAND researchers conducted in-person and phone interviews with personnel in JPSD, TFL, and MDE in two data collection windows (spring 2017 and spring 2018). We spoke with all personnel who were involved in the direct implementation of TFL at the district level. This included the one JPSD TFL lead and six TFL implementation coaches. We also spoke to the three consultants from Tools for Life Corporation who were embedded in the schools to offer supports and guidance to school leaders and instructional staff. We purposefully selected JPSD central office staff from the superintendent's office and those who were in charge of guidance counseling and curriculum standards to gain insights on the vision of TFL and how TFL could be integrated into broader district processes and practices. We purposefully selected the one JPSD school board member who had recently started their term to gain insights on their role, expectations for the district, and for TFL. We purposefully selected officials from MDE who could provide a deeper understanding of the relationship between the state and district. Table 2 lists which stakeholder we interviewed in which window.

Table 2. Jackson Public School District, Tools for Life, and Mississippi Department of Education Officials Interviewed for Implementation Analysis

Data Collection Window	Interviewee	N
Spring 2017		
	JPSD TFL lead	1
	JPSD TFL implementation coach	6 (group interview)
	JPSD central office staff	2
	JPSD school board member	1
	Mississippi State Department of Education administrator TFL consultant	4 3 (group interview)
Spring 2018		
	JPSD TFL lead	1
	JPSD TFL implementation coach TFL consultants	6 (group interview) 3 (group interview)

Interview protocols sought to obtain data on the implementation of TFL to document its design features and to provide context for explaining any outcomes. We used semistructured interviews that included open-ended questions with supplemental probes to examine specific topics. The protocols differed for each type of interviewee and were modified in each data collection window in order to inquire about changes to the program or implementation through time.

2. Study of Six Focal Schools

In spring 2017, fall 2017, and spring 2018, we conducted site visits at six elementary and middle schools (three treatment and three control schools), conducting interviews with school administrators, focus groups with instructional staff members.

Interviews. Each site visit consisted of a 45–60-minute semistructured interview with a school leader/administrator (i.e., principal or assistant principal) and 45–60-minute interviews with designated lead TFL implementers or their counterparts in control schools (i.e., elementary school counselors and middle school social studies teachers). Protocols addressed such topics as perceptions of school climate and students' SEL needs, response to the program, and facilitators and challenges to implementation.

Focus Groups. Each site visit also included a 60-minute focus group with about six to ten staff members. We held focus groups with elementary staff after school; at middle schools, we held focus groups during the school day and selected among staff that had planning period during the chosen time. Prior to the start of each focus group, we assigned participants numbers for identification purposes so as to ensure confidentiality of participants in our notes. Protocols addressed similar topics as the interviews.

3. Observations of Districtwide TFL Events

We observed and took field notes during training sessions for school leaders and/or staff, year-end conferences jointly held by JPSD and Tools for Life Corporation, and bring-and-brag sessions at which designees from implementing schools across the district (not just the focal schools) shared examples of how they integrated the program into their schools and classrooms.

4. Student Surveys

To inform the implementation and outcomes analyses, RAND researchers fielded online surveys to students in grades 3 through 8. Questions on the survey inquired about students' self-assessments of their interpersonal and intrapersonal relationships and about the school's climate.[1] The student survey was adapted from several existing instruments, including the Alaska School Climate and Connectedness Survey, the 5Essentials My Voice, My School Survey and the Social Skills Improvement System (SSIS) Student Form. Intact scales were taken from each source, and scores were computed for each student (or school) based on these scales. Internal consistency estimates from the baseline survey administrations are reported in Table 3, as are estimates of the intraclass correlation coefficients (ICCs), which describe the extent to which survey constructs varied across schools. Most of the internal consistency estimates for the student level constructs are above 0.70 and show adequate reliability. For the school level constructs, the ICCs imply that the school-mean scale scores were also mostly reliable, with average school-mean reliabilities between 0.38 and 0.88. The Safety scale had the lowest school-level reliability. Because these constructs were considered central to TFL's theory of action, we included all these scales in our analyses. Also included in the survey was a set of items related to TFL programming. These items asked students whether they had seen TFL materials in their schools and whether they had participated in any lessons related to these materials.

Table 3. Student Survey Sources and Constructs

Source	Construct	Items	Cronbach's Alpha	Range	ICC
Student level					
My Voice, My School	Emotional Health	4	0.67	1-4	0.02
Alaska School Climate and Connectedness Survey	SEL	7	0.70	1-4	0.01
SSIS	Communication	6	0.81	6–24	0.08
SSIS	Empathy	6	0.86	6–24	0.07
SSIS	Self-Control	6	0.86	6–24	0.09
SSIS	Engagement	7	0.86	7–28	0.07
School level					
My Voice, My School	School Safety	3	0.64	1-4	0.11
My Voice, My School	Safety	5	0.68	1-4	0.01
My Voice, My School	Student-Teacher Trust	5	0.85	1-4	0.04
My Voice, My School	Student-Peer Relationships	3	0.44	1-4	0.03

Surveys were administered before and after the implementation of the TFL, in fall 2016, spring 2017, fall 2017, and spring 2018. Student surveys were administered to students whose parents actively consented that they could participate, using an online survey platform in treatment and control schools. Just prior to each administration, principals received emails with specific instructions for administering surveys. Follow-up with coordinators were made to check on data collection progress and respond to any challenges. Additional phone-calls, emails, and, in a few

[1] The original research plan was designed to include surveys of students and instructional staff. However, the initial fielding of the staff survey in fall 2016 showed very low staff participation, even with multiple follow ups and reminders. Response rates averaged around 16 percent per school for these surveys. Thus, after the first administration, the data collection plan was revised to focus energy and resources on maximizing student participation in the surveys.

instances, on-site assistance were provided to coordinators and teachers as necessary to ensure that these data were collected in a timely manner and for as many of the targeted students as possible.

5. Jackson Public School District Administrative Data

JPSD provided RAND with de-identified administrative data for all students whose parents or guardians had consented to their participation. These data included demographic information (race/ethnicity and gender), as well as indicators of suspension, attendance, and annual performance on the Mississippi Academic Assessment Program (MAAP) tests in English Language Arts and Mathematics. MAAP scores are not vertically scaled. For the purposes of this analysis, all test score data were standardized within grade level based on state means and standard deviations. We obtained administrative data from JPSD for both years of the study (2016–2017 and 2017–2018). The complete list of administrative data is in Table 4.

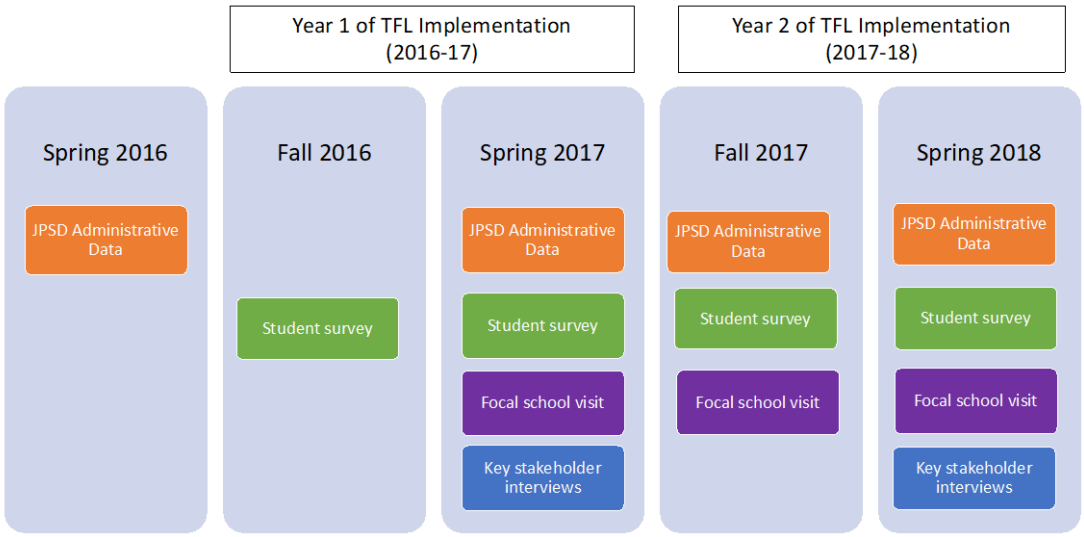
Table 4. JPSD Administrative Data Used in The Study

Student Level	School Level
Demographic information	
• School of enrollment • Grade level • English language learner/limited English proficient indicator • Free or reduced-price lunch indicator • Indicator of homelessness or housing instability • Special education status • Race/ethnicity • Gender	• Title I eligibility • Magnet school indicator • Total enrollment • Total male enrollment • Total special education enrollment • Total ELL enrollment • Total enrollment by race/ethnicity
Achievement information (MAAPS scores)	
• Language arts score • Math score	• Average language arts score (by grade) • Average math score (by grade)
Attendance, enrollment, and disciplinary data	
• Number of days attended • Number of days enrolled • Out-of-school suspension • In-school suspension	• Average daily attendance • Out-of-school suspension rate • In-school suspension rate

Source: JPSD administrative data.

The data collection timeline for these sources is presented in Figure 2

Figure 2. Data Collection Timeline



Limitations to Analysis

There are several limitations to this study that warrant consideration. First, the conclusions we can draw about program impacts should not be generalized, given the relatively small proportion of eligible students that consented to participate in data-collection activities. As previously noted, despite repeated efforts to secure parental consent from students, less than one-third of the eligible students in the district opted to participate in data collection. The remaining students either did not return consent forms or returned consent forms indicating that their parents or guardians did not wish to have their children participate in data-collection activities. The low participation rate poses a serious threat to validity of inferences about program impacts because of the possibility of selection bias. It is not clear whether the consenting students are different from other students in some way, and therefore it is not clear whether the treatment estimates are broadly applicable to all students in the district or only to the subset of students that provided consent to participate. Although our sensitivity analyses suggested that consenting and nonconsenting students were similar demographically, we cannot rule out the possibility of selection bias.

Second, while our initial study design incorporated data-collection activities involving both students and instructional staff, initial fielding of the staff survey showed very low staff participation, even with multiple follow-ups and reminders. Response rates averaged around 16 percent for these surveys, rendering the data unusable for analytic purposes. After one administration of the instructional staff surveys, we revised our data-collection plan to focus energy and resources on maximizing student participation in the surveys.

A third limitation was small sample sizes. Initial a priori power analysis showed that with over 10,000 students from 52 schools, we have 80 percent power to detect an effect size of 0.14–0.26 on student outcomes, assuming a range of 0.05-0.15 for the intra-class correlation and a range of 30–50 percent for the proportion of variance explained by the covariates. However, our

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

actualized sample size was smaller at both the student and school level. Post hoc power calculations suggest that, for the same ICCs and proportion of variance explained by covariates, we have 80 percent power to detect effects between 0.17 and 0.30. This is important for two reasons. First, the minimum detectable effect sizes are *larger* than were initially planned for. Second, most of the estimated effects from the study were considerably smaller in magnitude than this. For student-level outcomes, effect sizes averaged around two-tenths of a standard deviation. While larger sample sizes may have resulted in enough power to detect effects of this size, it is also important to consider the extent to which these effects are practically significant (e.g., representing substantively meaningful program impacts on student outcomes).

A fourth limitation is that our survey measures of SEL and school climate outcomes, while based on widely used instruments with evidence supporting their validity, are limited to student self-perceptions. There is some evidence that self-report measures of SEL and school climate may be biased. In particular, acquiescence bias, whereby respondents have a tendency to agree with all items, and halo effects, whereby positive perceptions about one domain influence perceptions of other domains, pose a particular threat to validity with such measures. This is also the case for all interview and focus group data, as those are self-reports. It is therefore possible that participants provided socially desirable responses about their perceptions of school safety and climate or about TFL implementation.

Finally, the data we collected and analyzed with respect to the focal schools have several limitations. First, among all the elementary and middle schools participating in the evaluation, we randomly selected three schools (two elementary and one middle school) in the treatment group and then selected three control schools that were similar in student population size, state assessment score results, and demographic breakdown of the student population. The aim was to include a range of school "types" and to compare implementation in the treatment and control schools. However, there was little variation of treatment and control schools in terms of observable selection characteristics (all schools are about the same size and have similar state assessment scores, and demographic populations). Given this and the relatively small sample size of six schools, findings based on collected data may not generalize to *all* schools in the district. Second, as described further in the report, given the small sample of schools, we cannot make credible comparative inferences about subgroups, such as elementary versus middle schools, or even control versus treatment schools, given that any differences may be confounded with principal and staff turnovers, among other factors. Third, we interviewed some, not all, staff at the focal schools. The data from the study of the six focal schools, then, may not represent the views of all stakeholders at each school. Despite these limitations, the case study offers qualitative insights into implementation successes and challenges and the perceptions of interview respondents on TFL.

C. Key Findings

TFL Had Little Impact on Students' Social and Emotional Learning, Behavior, or Academic Performance

In analyzing JPSD administrative records and student survey responses, we found that after one year of implementation of TFL, there were no practically or statistically significant differences between schools that implemented TFL (treatment schools) and those that did not (control

schools) in measures of students' social and emotional, school climate, behavioral, or achievement outcomes. We also conducted exploratory examinations of differences in outcomes among subgroup populations, specifically comparing middle school with elementary school students and students who were homeless with those who were not. We did not find any discernable differences across subgroups. We conducted exploratory analyses of the impact of full TFL exposure on student outcomes. We did find a positive association between students' self-reported exposure to TFL materials and lessons and their self-reported social and emotional outcomes. However, we found significant negative associations between TFL exposure and academic achievement, as measured on scores on state assessments.

Key Stakeholders Reported Positive and Negative Views About TFL Implementation

Interviews with JPSD implementation coaches (district staff who were to support TFL implementation) and TFL consultants, and interviews with school leaders and focus groups with instructional staff in the six focal schools, revealed varied perspectives about TFL. Teachers with whom we spoke noted that materials and resources were reportedly easy to use, yet they also noted that materials and resources were not age-appropriate (they seemed too young or basic) for middle schoolers. Interviewees and focus group participants also noted that dosage was low: School administrators and instructional staff wore lanyards with the basic concepts of TFL on them, hung posters about feelings in classroom and school hallway walls, and created calm-down corners in classrooms so that students had quiet places to go to if they felt they needed to, but the ongoing instruction about TFL concepts was left to guidance counselors (in elementary schools) and to social studies teachers (in middle schools). Moreover, focus group participants noted that the TFL lessons were reportedly not clear; staff had little idea about how exactly to implement the lessons, if they had chosen to make time to do so.

Implementation of TFL Was Uneven Across Schools and, in Many Cases, Reportedly Shallow

Interviews with JPSD implementation coaches (district staff who were to support TFL implementation) and with TFL consultants revealed that implementation of TFL varied across schools in both years of the study. Interviewees also reported that this variation occurred even in the second year for treatment schools, as well as across the control schools, which implemented TFL in 2017–2018 only. Student surveys indicated that exposure to various TFL materials and lessons varied widely across treatment schools, and also between treatment and control schools (in year 1). Our analysis in six focal schools corroborated the students' survey results: Interviews and focus group discussions with school staff and leaders in each school reported different ways in which TFL materials and structure was provided to students. TFL was rarely, if at all, implemented across a whole school as it was designed. Thus, implementation of TFL in JPSD over the first two years remained shallow.

A Variety of Factors Facilitated and Impeded Implementation

Interviews with key stakeholders and the analysis in six focal schools unveiled factors that could have affected the implementation of TFL and led to the reported uneven implementation across schools and varied exposure of TFL concepts to students. These factors included attitudes, district-level structures, and broader contextual factors:

- **Attitudes:** The extent to which parents, teachers, and especially school leadership had bought-in to the positive-behavior concepts promoted by TFL stakeholders' desire for a positive behavior approach.
- **District-level structure:** Respondents reported that JPSD and TFL support facilitated implementation for them. However, they also noted situations in which communication lines faltered and that there were few opportunities for collaboration. Parental engagement was minimal, and in those schools that experienced principal turnover there was a lack of clear leader advocacy or attention to the program.
- **Contextual factors:** MDE audited JPSD concurrently with implementation of TFL in 2016–2017, which reportedly distracted many educators from implementing TFL. Moreover, teachers and school leaders were reportedly "burned-out" on having to implement new programs that they perceived might not last in the district beyond the one year.

TFL Was Relatively Expensive to Implement

In our cost analysis, we found that TFL, as implemented in JPSD, cost approximately $2,797,728 over the course of the 2016–2017 and 2017–2018 school years. Breaking this figure down by year and by student, we calculate a participant cost of $174.31 per year, per student. We conducted sensitivity analyses in which we included alternative sets of assumptions, such as opportunity costs for teacher time for program implementation and attendance at events. These sensitivity checks increased the costs to as much as $177.79 per student, per year, bringing the total potential annual cost to $352.10 per student.

Although there are only a few examples of cost analysis of whole-school interventions that focus on improving school culture and student behavior from which we may draw comparisons, TFL appears to be relatively expensive compared with programs studied in other cost analyses. In particular, Blonigen et al. (2008) found that a schoolwide positive behavioral support program would cost approximately between $17,732 and $20,705 per school. If we take the average school in JPSD, which as of 2017 has 235 students, the average per-school, per-year cost of TFL implementation comes out to be $40,962.85. Furthermore, if we include the results of our sensitivity analyses by including in-kind opportunity cost of teachers' time commitment, that per-school, per-year cost increases to $82,743.50.

D. Policy Implications

Although the results from the outcomes analysis suggest that the offer of TFL was generally not associated with better outcomes, and thus that TFL might not have been effective or aligned with the specific needs of JPSD school population, it is clear from the documentation on implementation struggles and facilitators that there are several other plausible explanations for these findings that merit consideration.

First, TFL was not the only program implemented in JPSD that was designed to improve student SEL and school climate. For example, the district was also implementing other programs, including a Positive Behavioral Interventions and Supports (PBIS) program. To the extent that the benefits of participating in programs such as the PBIS program benefited all students in the district, the estimated effects from this study represent only the additional impact of the TFL program. This additional impact is likely to be smaller than would be anticipated had the control

condition been defined so that students in control schools received no additional SEL programming.

Second, issues of implementation may have compromised the potential of the program to help students. We found some evidence that implementation of TFL was not optimal: It varied across schools and was, at best, shallow. This is evident from both the data from focal schools and from the heterogeneity in program exposure (presented in Chapter Three)—there were students enrolled in schools that were implementing TFL programming in year 1 who reported low exposure to posters and lessons. Educators' contextualized critique of the program and reasons for not fully buying into it, described in Chapter Three, may provide further insight into the lack of detected effect in the outcomes. Moreover, at the time of TFL implementation, JPSD was dealing with superintendent turnover, the dissolution of its school board, and MDE's audit of its performance. These stressors reverberated among the stakeholders with whom we spoke and most likely affected educators' capacity to implement TFL as it was originally designed. Finally, there are some methodological issues to consider. First, a large number of parents and guardians in JPSD did not consent to have their children participate in the study, and it is unclear whether the students who did consent were the students who would benefit most from participating in the program. It is unclear whether the program would be more effective for these students and to what extent that would influence the overall estimation of program impacts. Second, the SEL and school climate outcomes rely on student-self report. There is some evidence that self-report measures of SEL and school climate may be biased. In particular, acquiescence bias, whereby respondents have a tendency to agree with all items, and halo effects, whereby positive perceptions about one domain influence perceptions of other domains, pose a particular threat to validity with such measures. This is also the case for all interview and focus group data, as those are self-reports. It is therefore possible that participants provided socially desirable responses about their perceptions of school safety and climate or about TFL implementation. It is also plausible that, as students participated in the program, their awareness of their own social and emotional competencies changes and they become less lenient in their judgments about their social and emotional competencies. This could even potentially explain the finding that self-reported self-awareness, social awareness, and self-management trajectories were negatively affected through TFL participation.

It is impossible to discern exactly which factor produced the results of this study, suggesting that additional research is needed to better understand the extent to which the findings reported here are driven by limitations of the data and sample, implementation challenges, or other factors.

E. Scholarly Publications

This study has three publications:

- Final report

Gonzalez, Gabriella C., Jennifer L. Cerully, Elaine Lin Wang, Jonathan Schweig, Ivy Todd, William R. Johnston, and Jessica Schnittka, *Social and Emotional Learning, School Climate, and School Safety: A Randomized Controlled Trial Evaluation of Tools for Life® in Elementary and Middle Schools*, Santa Monica, Calif.: RAND Corporation, RR-4285-NIJ, 2020. As of April 09, 2020: https://www.rand.org/pubs/research_reports/RR4285.html.

- Technical Appendix

Gonzalez, Gabriella C., Jennifer L. Cerully, Elaine Lin Wang, Jonathan Schweig, Ivy Todd, William R. Johnston, and Jessica Schnittka, *Social and Emotional Learning, School Climate, and School Safety: A Randomized Controlled Trial Evaluation of Tools for Life® in Elementary and Middle Schools—Appendixes*, Santa Monica, Calif.: RAND Corporation, RR-4285z1-NIJ, 2020. As of April 09, 2020: https://www.rand.org/pubs/research_reports/RR4285.html.

- Research Brief

Gonzalez, Gabriella C., Jennifer L. Cerully, Elaine Lin Wang, Jonathan Schweig, Ivy Todd, William R. Johnston, and Jessica Schnittka, *Four Lessons Learned from Implementing a Social and Emotional Learning Program to Enhance School Safety*, Santa Monica, Calif.: RAND Corporation, RB-10101-NIJ, 2020. As of April 09, 2020: https://www.rand.org/pubs/research_briefs/RB10101.html