October 2024

Impact Evaluation of TRAILS Tier 1 Social and Emotional Learning Curriculum in Two Intermediate School Districts in Michigan

AUTHORS

Katja Robinson

Robin Jacob

EXECUTIVE SUMMARY

Social Emotional Learning (SEL) is "the process through which all young people and adults acquire and apply the knowledge, skills, and attitudes to develop healthy identities, manage emotions and achieve personal and collective goals, feel and show empathy for others, establish and maintain supportive relationships, and make responsible and caring decisions."[1] Schools play an integral part in supporting this process by intentionally adopting SEL programs into their schools. Research on the effectiveness of school-based universal SEL interventions show that these programs generally have a positive effect on social and emotional skills, improve classroom behavior, increase academic performance, and strengthen overall well-being.[2]

During the 2022–2023 academic year, the Youth Policy Lab carried out a comprehensive evaluation of TRAILS (Transforming Research into Action to Improve the Lives of Students) Tier 1 Social and Emotional Learning (SEL) Curriculum in two intermediate school districts in Michigan. We conducted a 1-year school-level randomized control trial to measure the impact of one year of exposure to TRAILS SEL curriculum on various student outcomes related to social-emotional well-being and mental health. This report begins with a comprehensive portrait of the TRAILS SEL curriculum, the evaluation design, and school environments pre-implementation to contextualize the effect of the program on student outcomes, and then explores the impact of the program on students.

During the 2022-2023 school year, 27 schools were assigned to receive TRAILS SEL programming, while 22 schools were placed in the delayed control condition. The study experienced a considerable amount of attrition, with a notably higher dropout rate among the program group. This higher non-response rate from program schools could introduce bias into our results.

Key Implementation Findings

- At the outset of the study, students in the program and control schools were similar across all primary outcomes.
- Schools in the control group were offering a substantial amount of SEL related activities and interventions.
    - 61% of staff in control schools reported that Tier 1 Universal Mental Health Interventions (e.g., SEL) were available in their schools, and 80% had participated in professional development activities focusing on SEL in the previous year.
- Most staff trained in the TRAILS SEL curriculum (74%) reported feeling sufficiently prepared to deliver the curriculum.
- However, both staff and student surveys indicated low levels of TRAILS program implementation.
    - Around 21% of program schools delivered an average of at least 10 lessons and approximately 28% of students in the program group indicated that they recognized two salient TRAILS program components in the end-of-year student survey.
- Barriers to curriculum implementation included students being uncomfortable, resistant, or not meaningfully engaging with the curriculum (noted by 36% of staff), lack of time due to other teaching duties (26%), and lack of class time for SEL instruction (25%).

1 Collaborative for Academic, Social, and Emotional Learning (2024, May 16). Fundamentals of SEL. https://casel.org/fundamentals-of-sel/

2 Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D. and Kriston B. Schellinger. "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions," Child Development 82, no. 1 (2011): 405–32, https://doi.org/10.1111/j.1467-8624.2010.01564.x; Durlak, J.A., Mahoney, J.L., and Alaina E. Boyle, "What We Know, and What We Need to Find out about Universal, School-Based Social and Emotional Learning Programs for Children and Adolescents: A Review of Meta-Analyses and Directions for Future Research," Psychological Bulletin 148, nos 11-12 (2022): 765–82, https://doi.org/10.1037/bul0000383.

Key Outcome Findings

- There were few statistically significant impacts on student outcomes. However, trends favored the program group and higher levels of implementation, or a greater contrast between SEL programming in program and control schools might have yielded more statistically significant results.
    - Students in the program group reported statistically significantly higher levels of self-awareness—one of the five core SEL competency.
    - In addition, students in the program group felt slightly more competent with respect to overall SEL skills, used effective coping skills slightly more often, and had slightly lower levels of depression and anxiety, although none of these findings were statistically significant.
- TRAILS effects on overall SEL skills were larger among Black students, particularly those who identified as Black girls.
- In program schools, there was a notable increase in staff reporting the availability of Tier 1 and Tier 2 interventions.
    - Around 49% of staff in program schools reported the availability of Tier 1 interventions and 53% reported the availability of Tier 2 intervention before program rollout while ~75% mentioned the availability of these interventions at the end of the study year.
- We saw an increase in staff participation in professional development activities centered on social and emotional learning, with 66% of staff reporting on PD participation before and 83% after program rollout.

Despite the low levels of implementation and the prevalence of SEL programming in the control schools, we saw positive trends in student outcomes for students in schools implementing the TRAILS Tier 1 curriculum, suggesting the potential for the program to have positive impacts on students when implemented widely.

Table of Contents

Program Description

Transforming Research Into Action to Improve the Lives of Students (TRAILS) is a school-based program designed to help improve social and emotional learning (SEL) skills, student well-being, and mental health.

The TRAILS program intends to empower the whole school community (e.g., teachers, student support staff, and administrators) by connecting them with professional development opportunities, programming, and specialized mental health and social and emotional learning resources.

TRAILS programming is organized around a 3-tiered conceptual framework: universal education for the whole school community (Tier 1), early intervention targeting students with identified mental health concerns (Tier 2), and crisis intervention to help schools actively manage risk among their highest-needs students (Tier 3).

This impact evaluation focused on the TRAILS Tier 1 Social and Emotional Learning (SEL) curriculum, which aims to help students learn the life skills that they need to thrive. There are three components to TRAILS SEL programming: 1) staff training, 2) access to TRAILS' SEL curriculum, and 3) continuous access to practice workshops, monthly newsletters, and other implementation tools (see Figure 1).

Figure 1: Three Components of TRAILS SEL Curriculum

1. Comprehensive Training in the Curriculum

TRAILS training in the SEL curriculum includes participation in a 2.5-hour live webinar or watching a recorded training with one of two types of live question and answer (Q&A) (see Figure 2)[1]. The content includes an introduction to SEL, an explanation of the theoretical foundation, some skills and lesson demonstration, and some live practice with feedback.

Figure 2: Comprehensive Training in the Curriculum

Mode of Delivery	Length	Content
One of the following: • a live SEL webinar • a recorded training with live Q&A in the same session, or • a recorded training with live Q&A scheduled at a different time.	• 2.5 hours	• Introduction to SEL (rationale and evidence) • Theoretical foundation (CBT) • Skills and lesson demonstrations • Live practice with feedback

[1] The impact evaluation excluded both asynchronous module-based training and facilitated training, which are also provided by TRAILS.

2. TRAILS' Social and Emotional Learning (SEL) (Tier 1) Curriculum

TRAILS' curriculum is geared towards enhancing social and emotional learning (SEL) skills, specifically focusing on the five core competencies identified by the Collaborative for Academic, Social, and Emotional Learning (CASEL). The five core competencies are self-awareness, self-management, social awareness, relationship skills, and responsible decision-making.

TRAILS offers a series of 20 concise lessons[2], each estimated to last about 30 minutes, intended to be delivered to the whole classroom over the course of a school year, and tailored to distinct grade levels: kindergarten through second grade, third through fifth grade, sixth through eighth grade, and ninth through twelfth grade.

Each lesson describes the lesson objectives, the learning target, the SEL core competency foci, and provides a list of necessary resources that are available for download from the TRAILS website. Staff who have participated in TRAILS training have access to these resources via a special log in (see Table 1 for an example lesson outline for grades 6-8).

Individual lessons all follow a similar structure, which includes the following sections: (1) mindful check-in, (2) a review of the previous lesson, (3) introduction to core content, (4) discussion and practice related to core content, and (5) wrap-up and mindful check-out. During the mindful check-in students identify their current feelings and the intensity of their feeling by completing a grade-band-specific check-in sheet. For example, students in grades 9-12 identify their current feelings by circling up to three of 27 displayed animated faces and rating their intensity. Section 2 of each lesson focuses on reviewing previously learned content and revisiting learned strategies or skills. In section 3, the teacher introduces and discusses the core content of the lesson (see Table 1 for sample content). In section 4, children can explore, engage in, and discuss the new topic and associated strategies. The final section focuses on a brief reflection on learned SEL skills. For example, students in grades 6-8 complete a "Coping Skills" worksheet after each lesson. This worksheet consists of 20 distinct coping skills used by students to keep track of skills they have learned.

Table 1: Example Lesson Outline for Grades 6-8

Lesson Title	Lesson Objectives	SEL Competency*
Introduction to SEL	Introduce social and emotional learning (SEL) • Understand SEL is a set of tools that can be used to help promote well-being • Understand universal human needs and how needs are or aren't being met	1, 2, 3, 4, 5
Mindfulness	Define mindfulness • Practice using mindfulness skills to describe an experience without judgment	1, 2, 3, 4, 5
What is CBT?	Understand the relationship between thoughts, feelings, and behaviors using the CBT Model (Think-Feel-Do Cycle)	1, 2, 3, 4, 5
Emotions	Understand emotions can change and vary in intensity • Understand that all emotions serve a purpose • Recognize that skills can help us manage strong emotions	1, 3
Thoughts	Recognize automatic negative thoughts (ANTs) • Identify common thinking traps	1, 2
Unhelpful Thoughts	Develop specific skills for challenging automatic negative thoughts (ANTs) • Learn how to generate more helpful thoughts	2

[2]Program schools in the evaluation study used the curriculum that consists of 20 lessons. The curriculum was updated to include 25 lessons after the study year.

Managing Strong Emotions	Understand that intense emotions can be uncomfortable but not dangerous • Understand that we can reduce the intensity of emotions by changing behavior • Identify situations when doing the opposite of one's behavior urge is helpful	2
Getting Active	Describe the cycle of inactivity • Understand how being active can improve mood	2
Relaxation	Understand how stress impacts our thoughts, feelings, and behaviors • Practice relaxation techniques	2
Self-Care	Define self-care and its impact on physical and mental health Create a self-care plan	2, 5
Identifying My Supports	Identify supportive people in our lives • Understand the importance of using supports when facing difficult situations	3
Empathy	Understand that others have perspectives that differ from our own • Understand how empathy can change our thoughts and feelings about a situation and that this can affect our behavior Understand that empathy can improve relationships	3, 4
Diversity	Define diversity, inclusion, and prejudice • Identify and consider stereotypes • Understand that diversity and inclusion add value to relationships	3
Establishing Relationships	Understand how negative thoughts can impact the ability to establish relationships • Learn and practice skills for interacting with new people	4
Clear Communication	Understand that listening effectively is part of communication • Describe ways in which one's thinking affects our ability to listen effectively • Identify and practice behaviors that demonstrate active listening	2, 4
Dealing with Conflict	Understand that conflict is normal, and when handled appropriately, can help strengthen relationships • Develop specific strategies for managing conflict effectively	2, 4
Maintaining Relationships	Describe qualities of a healthy relationship • Understand how choosing effective behaviors in relationships positively impacts our own thoughts and feelings	2, 3, 4
Problem-solving	Use specific strategies to inform decision-making • Recognize the CBT model in decision-making	5
Values	Identify personal values • Notice how personal values influence behaviors and decision-making • Notice how we feel when we behave in ways that align and don't align with our values	3, 5
Goals	Identify steps and skills that lead to goal achievement • Identify a long-term goal and several short-term goals that work toward the larger goal	3, 5

NOTES: SEL Competency (1= self-awareness, 2= self-management, 3=social awareness, 4= relationship skills, 5=Responsible Decision-Making)

3. Access to Best Practice Workshops and Informational Newsletters

Staff who have participated in TRAILS training have ongoing access to best practice workshops and informational newsletters. Best practice workshops are offered via Zoom throughout the year and are free of cost. Workshop topics include, for example, how to use TRAILS lessons flexibly in the classroom or how to build trauma-informed classrooms (see Figure 3). The newsletter is distributed every other month, is free of charge, and includes information about future workshops, any modifications to the curriculum, and suggestions about SEL integration in the school year (see Figure 3).

Figure 3: Ongoing TRAILS Offerings

Best Practice Workshops	Informational Newsletter
Requirements: Completion of TRAILS SEL curriculum training prior to attending a workshop **Costs**: Free **Frequency:** Workshops on different topics are offered throughout the year **Continuing Education Credits Offered**: Yes **Mode of Delivery:** Workshops take place via Zoom (Recordings of past workshops are available—no education credits available for viewing recordings) **Example Workshop Topics:** • Coordinating TRAILS SEL at the building level • Using TRAILS lessons flexibly • Incorporating TRAILS SEL throughout the school day • Building trauma-informed classrooms • SEL in High Schools: Challenges and Opportunities (How to make SEL lessons more relevant to high school students)	**Costs:** Free **Frequency:** Every two months **Content:** Any of the following: • Suggestions about SEL integration aligned with the school year (e.g., suggestions for establishing a learning environment that is conducive to SEL are included in a newsletter at the beginning of the year) • Information about future workshops and how to register • Any updates or modifications to the curriculum

Impact Study Design

This impact study used a 1-year school-level randomized control trial design to assess the effectiveness of TRAILS Tier 1 social and emotional learning (SEL) curriculum on 4th to 12th grade students in two intermediate school districts (ISDs) in Michigan. Forty-nine schools participated in the evaluation across the two ISDs. Twenty-seven were randomly assigned to receive TRAILS SEL curriculum, training, and support in the 2022-2023 school year and 22 were assigned to a delayed control condition who did not receive TRAILS training and support until the 2023-2024 school year. By randomly assigning schools to implement in two cohorts (the first in 2022-2023 and the second in 2023-2024), we were able to assess the impact of one year of TRAILS SEL curriculum on student behavioral and mental health outcomes.

Primary Research Questions:

1. What is the impact of TRAILS SEL curriculum on students' social and emotional learning skills compared to students in a business-as-usual comparison group?
2. What is the impact of TRAILS SEL curriculum on students' use of effective coping skills compared to students in a business-as-usual comparison group?
3. What is the impact of TRAILS SEL curriculum on students' perceived embarrassment in relation to help-seeking behavior compared to students in a business-as-usual comparison group?

Exploratory Research Questions:

1. What is the impact of TRAILS SEL curriculum on students' symptoms of depression compared to students in a business-as-usual comparison group?
2. What is the impact of TRAILS SEL curriculum on students' symptoms of anxiety compared to students in a business-as-usual comparison group?
3. How does the impact of TRAILS SEL curriculum on proximal and distal outcomes vary by:
    a. Gender?
    b. Race/ethnicity?
    c. Students who experience homelessness?
    d. Students who speak exclusively another language at home?

Supplementary Descriptive Analysis:

1. How did staff who were trained in TRAILS SEL curriculum perceive student growth in SEL skills?
2. How did staff who were trained in TRAILS SEL curriculum report on their own growth in teaching SEL?
3. How did staff in the program group report on available mental health interventions, trainings and support, and participation in professional development activities before and after program rollout?

Measures:

We examined the impact of exposure to TRAILS SEL programming on changes in the following proximal outcomes 1) social and emotional learning skills, 2) use of effective CBT coping skills, and 3) perceived embarrassment related to help-seeking behavior (see Appendix Table A. 1 & 2).

To measure changes in social and emotional learning skills, students responded to a set of 26 items reflecting the following SEL domains: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. Individual items were rated on a scale from 1 (very difficult) to 4 (very easy) indicating how easy it was for the student to exhibit specific skills (e.g. Knowing ways to make myself feel better when I'm sad). These items were adapted from the open-source Washoe County School District Social and Emotional Competency Assessment

Long Form (WCSD-SECA Long)[1]. The WCSD-SECA Long was developed in collaboration by WCSD, the Collaborative for Academic and Social and Emotional Learning (CASEL), and the University of Illinois at Chicago. Thorough statistical analyses have been conducted to demonstrate that the instrument is reliable and valid.[2]

Changes in knowledge and use of evidence-based CBT coping skills were measured using an original measure developed by TRAILS. This self-report measure includes a list of five coping skills (behavioral activation, mindfulness, listening to music, cognitive coping, and exposure) that are rated on a scale from 1 (I don't know what this is) to 5 (I use it very often).

Change in perceived embarrassment was measured using one standalone item that asked students whether they would feel embarrassed if other students knew they had received support for a personal problem.

We also explored changes in distal outcomes, including changes in 1) symptoms of depression, and 2) symptoms of anxiety. Changes in symptoms of depression were measured using the two-item validated Patient Health Questionnaire (PHQ-2)[3]. This self-report measure includes the first two items of the longer PHQ-9 validated measure and reflects two symptoms "depressed mood" and "loss of interest or pleasure in daily activities" from the DSM-V. Students rated how frequently they have been bothered by symptoms of depression in the last two weeks on a 4-point scale from 0 (not at all) to 3 (nearly every day).

Changes in symptoms of anxiety were measured using the Generalized Anxiety Disorder 2-item (GAD-2) screener that is recommended as an initial first step to screening for generalized anxiety disorder.[4] This short measure includes the first two items from the longer GAD-7, a seven-item validated diagnostic tool for generalized anxiety disorder. Students were asked how frequently they have been bothered by symptoms of anxiety in the last two weeks and rated the two items on a 4-point scale from 0 (not at all) to 3 (nearly every day).

In addition, we explored findings from implementation surveys that were developed and administered by TRAILS to staff trained in TRAILS SEL curriculum. We particularly focused on changes in staff perceptions between the beginning and end of the year regarding 1) student growth in SEL skills, and 2) their own growth in confidence teaching SEL skills. To measure changes in student growth in SEL skills, staff were asked to rate the percentage of their current students who consistently demonstrate age-appropriate competency in a list of nine social and emotional skills. To measure any changes in confidence in teaching SEL skills, staff were asked to rate a list of 17 items relating to teaching skills in self-awareness, social awareness, self-management, interpersonal communication, decision-making, and CBT on a 4-point scale from 1 (not confident at all) to 4 (completely confident).

Finally, we analyzed changes in practices and interventions among the program group, focusing on changes in 1) available mental health interventions, 2) available trainings and supports, and 3) participation in professional development. To measure changes in the availability of interventions, trainings, and supports before and after program rollout, staff were asked to rate the availability of four types of mental health interventions and three types of training or support used in their school to support student mental health. An intervention, training, or support was classified as available when staff checked any of the following: a) available but not implemented, b) available but need support, or c) available and implemented well. To measure changes in participation in professional development before and after program rollout, we asked staff to rate how often they participated in any professional development on the following topics: a) student mental health, b) social emotional learning, and c) trauma or post-traumatic stress disorder.

Survey Approach

The Youth Policy Lab, with input from TRAILS, developed web-based surveys for students, instructional staff in program and control schools, and staff who were trained in TRAILS SEL curriculum.

The student survey included questions about students' well-being, coping behaviors, social and emotional learning skills, mental health stigma, and demographic questions. It was open to students in grades 4-12 who were slated to either receive TRAILS SEL curriculum in the school year 22-23 (program schools) or the school year 23-24 (control schools). We developed two grade-band-specific survey versions (4th-6th grade and 6th-12th grade)[3] that differed in length and content based on grade level and age of students. To enhance accessibility and inclusion, we translated the student survey into the following languages, per schools' preference: Albanian, Arabic, Farsi, French, Spanish, Russian, and Vietnamese.

The survey for all instructional staff in program and control schools included questions about existing interventions, practices and trainings, professional development participation, screening and referral protocols, self-reported burnout, stigma as it relates to help-seeking, and perception of students' classroom behavior.

Prior to survey administration, schools identified one or multiple survey coordinators who managed survey logistics and served as the point of contact for the study team. Survey coordinators received a comprehensive student and staff survey administration manual. The administration manuals included suggestions for survey administration, eligibility criteria for students and staff to participate, sample emails and backpack letters to be shared with parents of eligible children, the list of survey questions, the link to the student and staff surveys, resources for students, an FAQ document, and other resources.

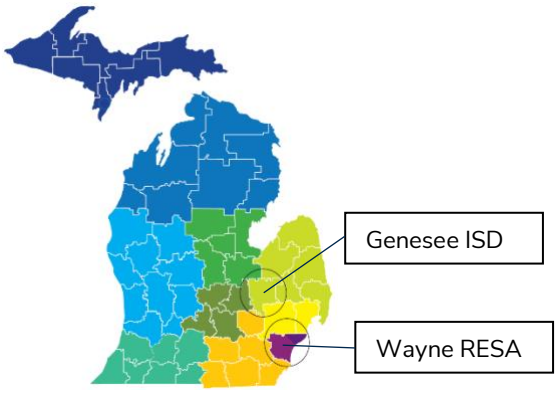
The study team strongly encouraged schools to administer the surveys to eligible students and staff before implementing the TRAILS SEL curriculum in program schools and again at the end of the 22-23 school year. The period for student and staff survey administration at the start of the school year varied by school and lasted from August 2022 to January 2023. Survey administration at the end of the 22-23 school year lasted from April to June 2023.

[3] Schools were able to select whether they wanted their 6th grade students to take the shorter survey intended for 4-6th grade students or the longer version intended for 6-12th grade students.

Where Our Data Came From

The study took place in two intermediate school districts (ISDs) in Michigan: Genesee ISD and Wayne RESA[4]. In the 22-23 school year, Genesee ISD enrolled ~60,000 students and Wayne RESA enrolled ~260,000. As shown in Table 2, the racial/ethnic distribution of students in both districts varied. The proportion of White students in Genesee was ~19 percentage points higher and that of Black/African American students about 16 percentage points lower than in Wayne. The percentage of students eligible for free or reduced lunch was comparable across both ISDs—63% in Genesee and 67% in Wayne. Wayne had a higher percentage of students who were classified as English Learners.

Figure 4: Map of Michigan's Intermediate School Districts



Source: Michigan Association of Intermediate School Districts

Table 2: Overview of Genesee ISD and Wayne RESA (Selected Characteristics)

	Genesee ISD	Wayne RESA
Overall Population		
Number of enrolled students	59,613	259,453
Race / Ethnicity (%)		
American Indian	.24	.24
Asian	.79	4.19
African American	25.62	41.49
Hispanic/Latino	5.9	9.21
Native Hawaiian or Other Pacific Islander	.07	.06
Two or More Races	6.68	2.91
White	60.71	41.90
Gender (%)		
Female	48.50	48.83
Male	51.50	51.17
Economic Status (%)		
Economically Disadvantaged Students[a]	63.27	67.41
Language (%)		
English Learners	<5.00	13.73

NOTES: Data Source: MI School Data. Information presented represents all grades in public and charter schools.

[a]The proportion of students who were eligible for free or reduced lunch was used as a proxy for economically disadvantaged students.

[4] RESA=Regional Educational Service Agency; sometimes ISDs are also called RESAs.

School Selection

Both ISDs provided a list of schools that were willing to participate in the impact evaluation: 15 schools in Genesee ISD and 34 schools in Wayne RESA. Prior to randomization, YPL stratified the list of schools based on ISD and grade levels served (K-6 or 7-12). Stratification ensured a balanced mix of younger and older students within each group. Randomization in Genesee ISD resulted in nine schools assigned to the program group and six schools assigned to the control group. Randomization in Wayne RESA resulted in 18 schools[5] assigned to the program group and 16 schools assigned to the control group (see Table 3).

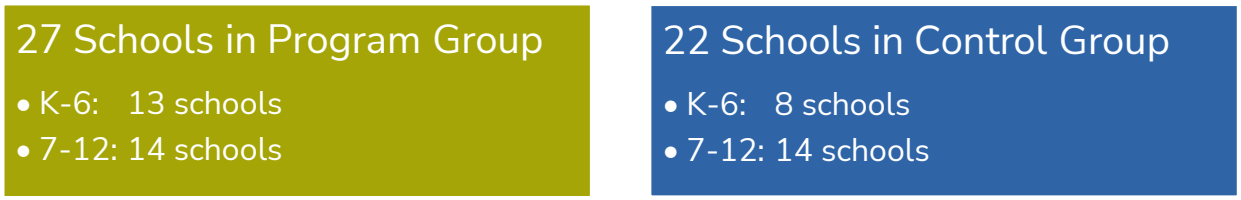
Table 3: Sample Selection by Selected Characteristics

	Total	Program Group	Control Group
Total Number of Schools Randomized	49	27	22
ISD			
Genesee ISD	15	9	6
Wayne RESA	34	18	16
Grade Band			
K-6	21	13	8
7-12	28	14	14

NOTES: K-8 schools (n=3) were classified in the K-6 category. K-12 schools (n=2) were classified in the 7-12 category.

Overall, 27 schools were assigned to receive TRAILS SEL programming in the school year 2022-2023 (program group) and 22 to the delayed control group who would not receive TRAILS SEL until the 2023-2024 school year (see Figure 5). The program group included 13 schools encompassing grades K-6 and 14 schools spanning grades 7-12; the control group included 8 schools in grades K-6 and 14 schools in grades 7-12.

Figure 5: Final Sample



[5] The randomization for Wayne RESA resulted in 17 schools assigned to the treatment and 17 schools assigned to the control condition. One school that was assigned to the control mistakenly assumed that it was a program school and thus, participated in TRAILS training and implemented the curriculum. The number reflects this adjustment.

Student Surveys

Below, we show an overview of survey data we received from students in both the program and control groups at the end of the study year. Overall, we analyzed study survey data from 19 schools in the program group (70% of eligible schools) and from 21 schools (95% of those eligible) in the control group (see Figure 6). This represents a substantial amount of attrition in the program group and a differential attrition of 25 percentage points between the program and control groups. According to What Works Clearinghouse[5], this level of attrition could introduce bias into the results. Both groups were comparable to one another in terms of grade level, gender, and race/ethnicity (see Figures 7 & 8, Tables 4 & 5). See Appendix Table B. 1 for additional information on student demographics.

Figure 6: Distribution of Student Survey Participation by Program Status

Program Group

19	schools represented
70%	of eligible schools represented
2244	student surveys analyzed

Control Group

21	schools represented
95%	of eligible schools represented
3175	student surveys analyzed

Figure 7: Grade Level, Program Group (n=2,244)

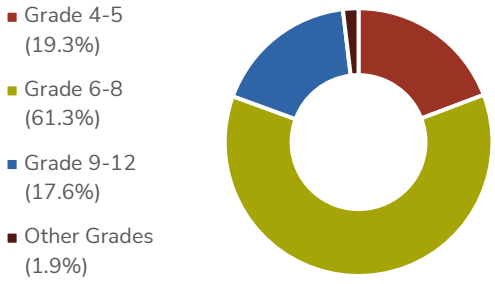


Figure 8: Grade Level Control Group (n=3,175)

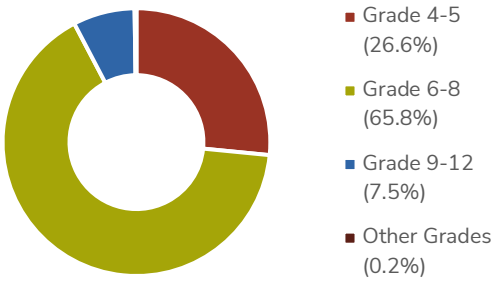


Table 4: Distribution of Gender, EoY 22-23

	Program Group n=1,982	Control Group n=2,855
Gender (%)		
Girl/Woman	44.6	47.1
Boy/Man	50.0	47.2
Gender non-binary	1.6	2.0
Questioning/unsure	0.5	0.8
Prefer not to say	1.6	1.8
I don't know	0.4	0.3
My gender is not listed here	1.5	0.8

Table 5: Distribution of Race/Ethnicity, EoY 22-23

	Program Group n=1,978	Control Group n=2,870
Race/Ethnicity (%)		
American Indian	1.2	1.8
Asian	1.4	1.1
Black or African American	18.2	21.3
Hispanic/Latinx	2.9	3.1
Middle Eastern/North African	6.8	4.5
Native Hawaiian/Pacific Islander	0.4	0.3
White	36.1	35.5
My race/ethnicity is not described here	2.4	3.0
Prefer not to say	15.0	14.3
Multiracial	15.7	15.3

Staff Surveys

Below, we show an overview of survey data we received from staff in both the program and control groups at the end of the study year. Overall, we analyzed survey data from 15 schools in the program group (56% of eligible schools) and from 17 schools (77% of those eligible) in the control group (see Figure 9). This represents a substantial amount of attrition in the program group and the differential attrition (21 percentage points between the program and control groups) and could introduce bias into the results. Both groups were comparable to another in terms of their professional roles, years of experience in their role, gender, and race/ethnicity (see Figures 10-13, Tables 6 & 7). See Appendix Table B. 2 for additional information on staff demographics.

Figure 9: Distribution of Staff Survey Participation by Program Status

Program Group

15	schools represented
56%	of eligible schools represented
278	staff surveys analyzed

Control Group

17	schools represented
77%	of eligible schools represented
355	staff surveys analyzed

Figure 10: Professional Role, Program Group (n=274)

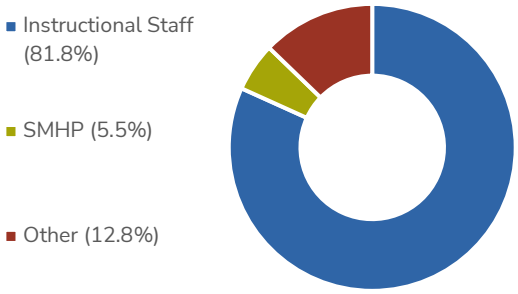
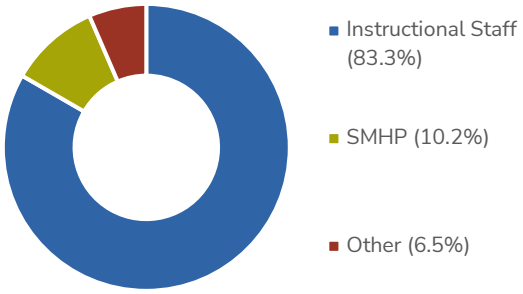


Figure 11: Professional Role, Control Group (n=354)



NOTE: SMHP=School Mental Health Professionals

Figure 12: Years of Experience, Program Group (n=240)

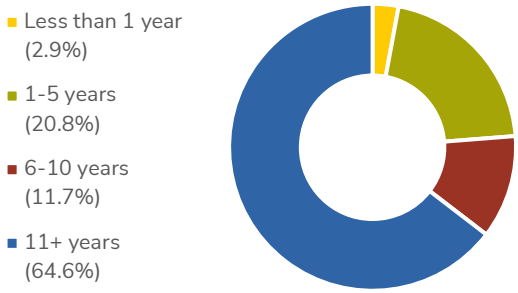


Figure 13: Years of Experience, Control Group (n=353)

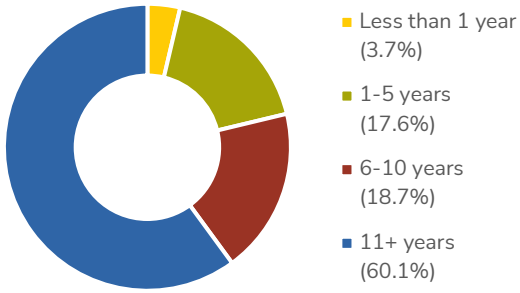


Table 6: Distribution of Gender, EoY 22-23

	Program Group n=230	Control Group n=332
Gender (%)		
Female	76.5	74.7
Male	14.4	16.0
Gender non-binary	0.9	0.3
Prefer not to answer	7.8	9.0
My gender is not listed here	0.4	0.0

Table 7: Distribution of Race/Ethnicity, EoY 22-23

	Program Group n=229	Control Group n=332
Race/Ethnicity (%)		
American Indian	0.0	0.3
Asian	1.8	0.0
Black or African American	3.1	6.3
Hispanic or Latinx	2.2	0.0
Middle Eastern / North African	5.2	0.9
Hawaiian or Pacific Islander	0.0	0.6
White	71.2	78.0
My race/ethnicity is not described here	0.4	0.0
Prefer not to say	9.6	11.1
Multiracial	6.6	2.7

School Context Prior to Implementation

Students' Behavioral Skills and Mental Health

To ensure a comprehensive understanding of the context in which the evaluation took place we present a detailed comparison of key student measures between the program and control group before the implementation of TRAILS SEL curriculum. Our analysis confirmed that both groups were similar across all primary measures prior to program rollout. Table 8 shows that, on average, students reported similar baseline levels of social and emotional learning skills, use of effective coping skills, and feelings of embarrassment. For example, on a scale ranging from 1=very difficult to 4=very easy, students had an average score of 2.79 for self-awareness in the program group and 2.75 in the control group. Similar shares of students presented with symptoms of depression and anxiety at baseline: ~39% of students in the program group and ~40% of students in the control group scored three or higher on the PHQ-2, indicating possible symptoms of depression. The percentage of students reporting symptoms of anxiety was identical in both groups (~45%). Based on this analysis, we conclude that both groups were comparable on key student measures, and any differences in these outcomes at the end of the year can be attributed to the effect of the TRAILS SEL curriculum with greater confidence[6].

Table 8: Comparison of Program Group and Control Group Prior to Implementation of TRAILS SEL Curriculum

	Program Group		Control Group	
	mean	std. dev	mean	std. dev
Student Survey:				
Social Emotional Learning				
Self-Awareness (6 items)	2.79	(0.58)	2.75	(0.61)
Self-Management (4 items)	2.40	(0.65)	2.36	(0.65)
Social Awareness (5 items)	2.84	(0.54)	2.85	(0.54)
Relationship Skills (6 items)	2.66	(0.56)	2.67	(0.55)
Responsible Decision-Making (5 items)	2.80	(0.59)	2.81	(0.58)
Overall SEL (26 items)	2.72	(0.45)	2.71	(0.45)
Effective Use of Coping Skills				
Combined Coping Skills Measure	3.33	(0.70)	3.30	(0.70)
Perceived Embarrassment (%)				
At least somewhat embarrassed if other friends knew about help receipt from a counselor	40.83		38.54	
Mental Health Measures (Grades 6-12 Only) (%)				
PHQ-2 Composite 3+	38.98		40.25	
GAD-2 Composite 3+	44.74		44.47	
Sample Size				
Students[a]	2936		3356	

NOTES: We report the standard deviation for mean outcomes in parentheses. The denominator for each question varies.

[a] The total number of students responding to each question varies. Data are available for at least 85% of the sample for all items related to Social Emotional Learning, Effective Use of Coping Skills, and Perceived Embarrassment. The program group denominator was 1,932 for PHQ-2 and 1,949 for GAD-2. The control group denominator was 1,558 for both PHQ-2 and GAD-2.

[6] We account for any small differences on the primary measures by including these as covariates in the HLM regression analysis.

Practices and Interventions in the Control Group

In the study, we compared the added value of implementing TRAILS SEL curriculum to a business-as-usual (BAU) condition in which schools continued their usual practice. In many cases this may have included other SEL programming or a different SEL curriculum. We captured the BAU condition by describing the control group on key indicators relating to practices and interventions prior to program rollout (see Table 9). Our analysis particularly focused on the availability of interventions, trainings and supports, and participation in professional development activities (PD) relating to mental health.

Our findings revealed that in the control group, approximately 61% of staff reported the availability of a Tier 1 universal mental health intervention (e.g. social and emotional learning curriculum). Around 42% reported the availability of mandatory training for instructors in student mental health and access to coaches and consultants. Furthermore, 80% of staff participated in professional development that focused on SEL.

Thus, the usual practices in the control group likely included a substantial amount of SEL-related practices and interventions and we are assessing the effectiveness of TRAILS SEL curriculum compared to a control group that presumably had considerable non-TRAILS SEL practices in place. If these other SEL practices were effective, we might expect to see a smaller difference between the program group, who implemented TRAILS, and schools in the control group who were implementing other, effective SEL interventions.

Usual Practices in Control Group Included a Substantial Amount of Non-TRAILS SEL Related Practices

- 61% Staff Reported that Tier 1 Universal Mental Health Intervention were Available
- 42% Staff Said Mandatory Training in Student mental Health and SEL was available
- 80% Staff Have Participated in Professional Development Activities Focusing on SEL

Table 9: Practices and Interventions in Control Group at Beginning of School Year 22-23

	Control Group
Interventions (%)	
Universal, school wide screening for student mental health	25.95
Tier 1 universal mental health interventions (e.g. SEL)	60.87
Tier 2 targeted mental health interventions for students with mild to moderate need	60.20
Tier 3 intense interventions for students with severe need	51.54
Training and Support (%)	
Voluntary training in student mental health or social and emotional learning	37.66
Mandatory training in student mental health or social and emotional learning	41.45
Access to consultants or coaches to support staff in implementing school-based mental health interventions	41.58
Professional Development (%)	
Student mental health	66.58
Social emotional learning	80.15
Trauma or post-traumatic stress disorder	54.95
Sample Size	
Staff[a]	404

NOTES: The table reflects the proportion of staff who reported that any of the listed intervention and training and supports were either a) available, but not implemented, b) available, but need support to implement it well, or c) available and implemented well. For professional development, we report staff who have participated in one of the listed PDs at least once in the last 12 months.

[a] The denominator for each question varies. For all variables in the table, data are available for at least 94% of the sample.

Analysis and Results

Analytic Approach

The impact evaluation compared behavioral outcomes between students in program schools that were slated to receive TRAILS SEL curriculum during the 2022-23 school year and students in control schools that did not receive TRAILS until the 2023-24 school year. Given the randomized design, a simple comparison between the two groups should capture the causal impact of the intervention at the end of the 2022-23 school year. We used hierarchical multilevel (HLM) regression models to assess whether TRAILS SEL curriculum had a significant impact on students' behavioral and mental health outcomes. We included the following covariates in our statistical models to increase statistical precision: baseline aggregate measures of school-level demographic characteristics including race, poverty, locale, the proportion of students experiencing homelessness, and proportion who were not born in the United States. We also included baseline aggregate measures of SEL skills, effective use of coping skills, perceived embarrassment, depression, and anxiety. At the individual level, we controlled for student race, gender, grade, district, whether the student experienced homelessness and whether the student was born outside the United States. We show both intent-to-treat estimates, which measure the effect of offering the program to a school and treatment on the treated estimates, which account for the degree to which students were actually exposed to the program. Missing values for covariates were imputed using mean imputation.

Criteria for Inclusion in Analytic Sample

For the regression analyses, students who completed at least 50%[7] of the follow-up survey were included in the analytic sample. Schools with fewer than five students responding to the follow-up survey were excluded from the analysis. Five out of 21 program schools and three out of 19 control schools had less than five students and were excluded from the regression analysis.[8]

The analytic sample for the subgroup regression analyses was further restricted. We excluded any schools with fewer than 20 student responses on the outcome of interest from analysis to ensure a comparable number of schools in each subgroup analysis. The final subgroup analyses included between 13 and 27 schools. Therefore, findings should be considered suggestive and not as definitive evidence.

Understanding Impact

In presenting results, we focus on both statistical significance and the magnitude of the results. Statistical significance indicates how likely we are to have found the observed difference by chance. The magnitude assesses the size of the effect; like a ruler that measures the differences between the group that received the program and the group that did not. An effect can be statistically significant but quite small or a very large effect may not be statistically significant. Understanding the impact of a program requires taking both into account.

Effect Size as a Measure of Magnitude

To measure the magnitude of effects in this study we use effect sizes. An effect size is calculated by dividing the difference between the program and control group by the standard deviation of the mean of the control group for the outcome of interest. One benefit of using effect sizes is that we can compare TRAILS SEL curriculum on various outcomes independent of their scoring. For example, the composite scales of the SEL core competencies are based on a different number of items and therefore the scoring range of the outcome varies by core competency. By focusing on the effect size, we can directly compare the impact of TRAILS across the different SEL core competencies.

[7] We used the progress variable in Qualtrics that indicates the exact percentage of how far a respondent got in the survey.
[8] The regression analysis with the PHQ-2 and GAD-2 measures as outcome variables excluded a total of 15 schools (10 program and 5 control schools) because only 6th to 12th grade students were eligible to receive the PHQ-2 and GAD-2 questions.

Interpreting Effect Size

Effect sizes can be positive or negative and they help us understand the strength of the program's effect. According to Jacob Cohen[6], effect sizes of 0.2 might be considered small and 0.8 might be considered large. We can also say that an effect size of 0.2 means that an additional 8% of students in the program group are doing better than the average student in the control group and an effect size of 0.8 indicates that an additional 29% of students are doing better than the control group's average. Overall, the greater the number of students in the program group scoring above the average student in the control group, the greater the effect size. There is a limit to the additional proportion of students scoring above or below the average of the control group. No more than an additional 50% of students can score above or below the mean of the control group.

Table 10: Interpreting Effect Size

Effect Size	Interpretation
0.0	Groups have the same mean score. There are no additional students in the program group scoring above the mean of the control group.
0.2	Additional 8% of students in the program group scored above the average student in the control group. This is considered a small effect.
0.5	Additional 19% of students in the program group scored above the average student in the control group. This is considered a medium effect.
0.8	Additional 29% of students in the program group scored above the average student in the control group. This is considered a large effect.
1.0	Additional 34% of students in the program group scored above the average student in the control group.
2.0	Additional 48% of students in the program group scored above the average student in the control group.

Estimating the Treatment on the Treated:

Our analyses begin by exploring estimates of the "intent to treat" (ITT) impact of the program. The intent-to-treat measures the impact of ***offering*** the program to schools, regardless of whether they implemented it. However, in this study not all teachers or schools who were offered the opportunity to implement TRAILS SEL curriculum did so. Therefore, we also include an adjusted effect size that shows the magnitude of the program effect ***had everyone who was offered the program implemented it***.[9] We call this adjusted effect size "treatment on the treated (TOT)". The TOT effect size is calculated by scaling up the ITT estimate by the proportion of the program group who implemented it. In our study, we used the proportion of the students in the program schools who indicated that they recognized two components of the TRAILS SEL curriculum (Feelings Thermometer/Check-In Sheet and CBT Model) to make this adjustment (see Appendix A for more information). Note that scaling up the estimate does not change the statistical significance of the findings, since standard errors are also scaled. If an ITT estimate is not statistically significant, the TOT results will also not be statistically significant, regardless of the magnitude of the TOT estimate.

Statistical Significance

Statistical significance helps indicate how much confidence we should have in our results—specifically how likely it is that the impacts we observe could have occurred by chance. The larger the observed impact, the less likely it is that it occurred by chance. Results that are not statistically significant suggest that the observed difference may be due to chance and we should not place too much confidence in them. In this study, most impacts that are larger than 0.10 standard deviations (for the full sample) are statistically significant.

[9] The TOT was calculated for the main student outcomes only.

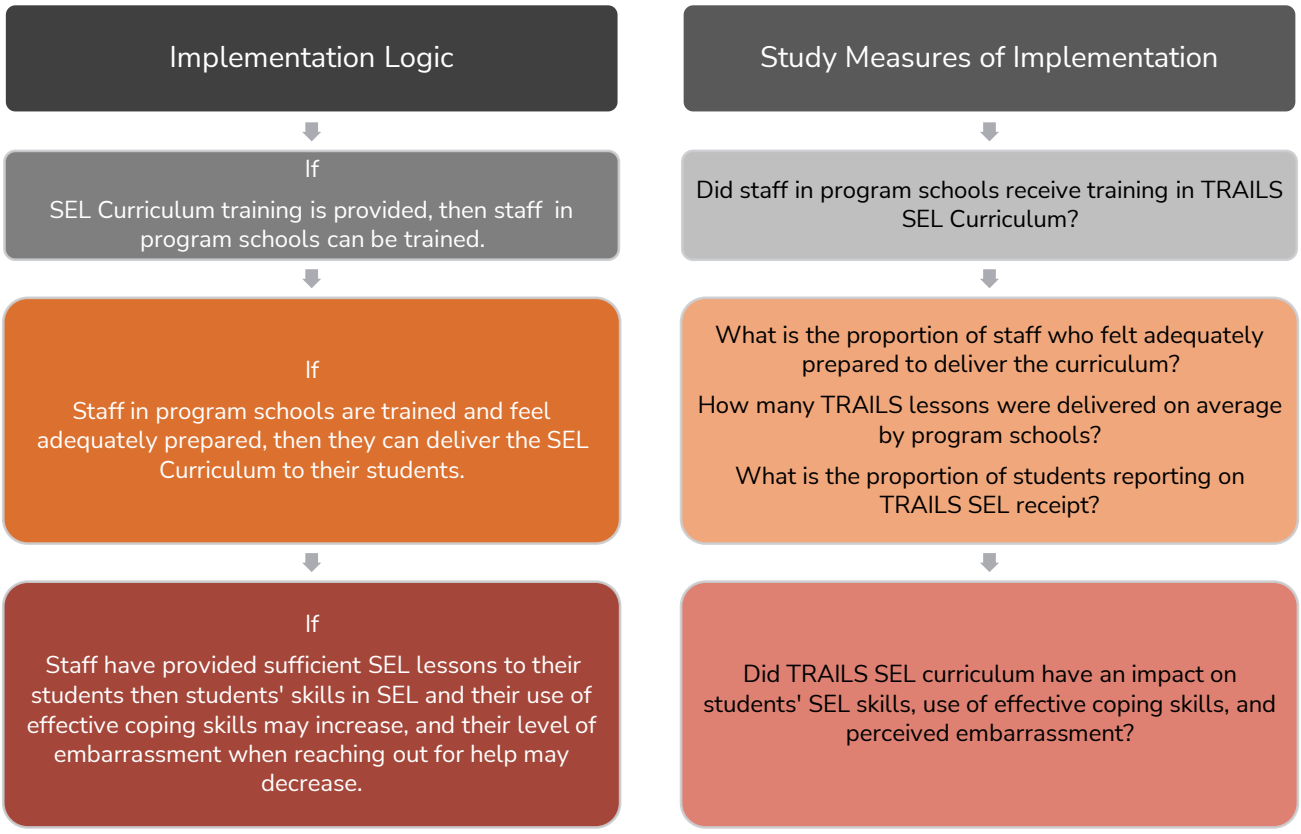
From Implementation to Outcomes

The effectiveness of any program depends on whether it is implemented as intended. Figure 14 outlines the logic model for the TRAILS SEL Curriculum. First, staff must be offered the opportunity to receive training on the TRAILS SEL curriculum. Second, staff need to participate in the training and feel adequately prepared to deliver the SEL curriculum effectively. Third, staff need to deliver SEL lessons to their students. We would only expect to see effects of the program on behavioral and mental health outcomes if all three conditions are met.

Understanding the level of program implementation is essential for gauging the potential impact on student outcomes. To this end, we explore on the following questions:

- Did staff in program schools receive TRAILS SEL curriculum training?
- What is the proportion of staff who felt adequately prepared to deliver TRAILS SEL curriculum?
- How many TRAILS lessons were delivered on average by program schools?
- What is the proportion of students reporting on familiarity with TRAILS?

Figure 14: Simplified Logic Model for TRAILS SEL Curriculum Evaluation



Level of Program Implementation, Adequacy of SEL Training, and Student Reach

Participation in Training

TRAILS reported that all schools in the program group either received a live SEL curriculum training, a recorded training with live Q&A in the same session, or a recorded training with live Q&A scheduled for at a different time. Staff who signed up for TRAILS SEL curriculum training received an invitation to participate in a TRAILS administered pre-training survey—a total of 528[10] staff responded to this survey. We infer that the number of staff who participated in the pre-training survey approximates the number of staff who were trained.

Adequacy of SEL Training and Other Factors that May Influence Program Delivery

TRAILS SEL trainees reported on a series of questions focusing on their perceived preparedness to deliver TRAILS SEL curriculum to their students, any barriers they encountered to program delivery, and their satisfaction with TRAILS programming overall. We report only on the 69 staff who participated in both the pre-training and post-implementation survey and indicated they teach in grades 4 to 12.[11]

Figure 15 shows that 74% of staff who had been trained in TRAILS SEL felt they had been adequately prepared by TRAILS, while 26% disagreed with this statement. Staff were most likely to report that they wished they had received "More training on the cognitive & behavioral skill concepts embedded in the curriculum" (26.1%) and "Opportunities to practice or observe example lessons or activities" (23.2%). See Appendix Table C. 1 for more information on training adequacy.

Figure 15: % of Staff Agreeing/Disagreeing with the Statement: "The Training I Received Adequately Prepared me to Deliver the TRAILS SEL Curriculum" (n=50)

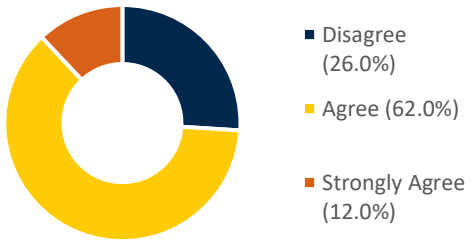
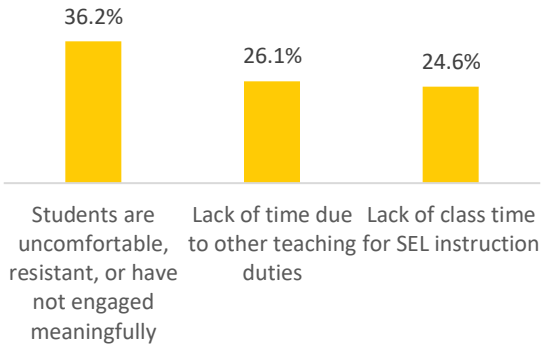


Figure 16 presents the top three barriers that staff said were a limitation to their SEL instruction. More than a third of TRAILS trainees reported that students are uncomfortable or resistant to the SEL lessons (36.2%). About a quarter of staff reported lack of time due to other teaching duties (26.1%), and lack of class time for SEL instruction (24.6%). See Appendix Table C. 3 for more information on barriers.

Figure 16: Top Three Reported Barriers to Providing SEL Instruction (n=69)

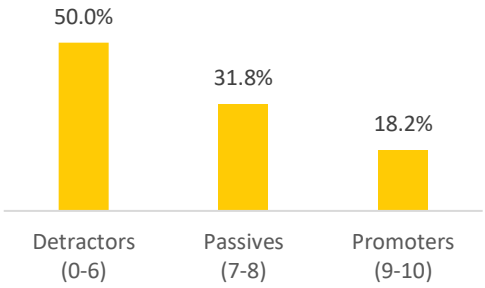


[10] 438 staff participated in only the pre-training survey, 39 in only the post-implementation survey, and 90 staff participated in both surveys.

[11] The denominator for individual questions varies.

We asked staff how likely they were to recommend TRAILS SEL curriculum to a friend or colleague on a scale from 0-10, where zero indicated extremely unlikely and 10 indicated extremely likely (see Figure 17). We used these responses to calculate the Net Promoter Score (NPS), a widely used metric measuring customer satisfaction. Respondents who scored between zero and six were classified as Detractors, those with scores of seven or eight were classified as Passives, and those with scores of nine and 10 were classified as Promoters. The findings indicate a negative NPS of -31.8, with a higher proportion of Detractors (50.0%) than Promoters (18.2%). This suggests potential issues with staff satisfaction regarding the TRAILS SEL curriculum. While a positive score is desirable, other satisfaction metrics and qualitative feedback should be considered to fully understand staff satisfaction. See Appendix Table C. 3 for more information on staff satisfaction.

Figure 17: Likelihood of Recommending TRAILS SEL Curriculum to a Friend or Colleague on a Scale of 0-10 (n=66)



Level of Program Implementation

We estimate the level of program implementation by using self-reported information from students and staff. To determine the number of lessons implemented by program schools, we focus on data from SEL trainees that responded to the post-implementation survey administered by TRAILS at the end of the 2022-23 school year. Table 11 shows the weighted average number of lessons that program schools reported.[12] We identified a threshold of implementing at least 10 out of 20 TRAILS SEL lessons as sufficient to detect gains on students' social emotional learning competencies.[13]

Only four schools (21%) reached the benchmark of delivering at least 10 lessons on average. Less than a quarter of program schools from which we have student outcome data sufficiently implemented TRAILS SEL curriculum.

Table 11: Average Number of Lessons Delivered

Number of Schools in Program Group that on Average Delivered:	
10+ Lessons	4
6-9 Lessons	4
1-5 Lessons	4
Number of Program Schools Considered	19

NOTES: Staff from six program schools from which we have student outcome data did not participate in the post-implementation survey. One other school participated in the survey but since this school only serves K-1 students, it is not considered here.

Evaluations of SEL interventions conducted before COVID-19 generally showed higher implementation rates than observed in this study. For example, MDRC's evaluation of Multi-tiered Systems of Support for Behavior found that

[12] The weighted average was calculated by using the total number of students served and the number of implemented lessons. This approach was used so that respondents who delivered lessons to a greater number of students received greater importance than staff who delivered lessons to fewer students.

[13] We are aware that this specification is limited as it only focuses on the quantity and not the quality of program delivery, and it relies on self-report data.

70% of schools met the implementation threshold, with only one school falling significantly short. However, some studies align with our observed lower implementation levels.[7] One classroom-based SEL program evaluation found that 41% of teachers did not return feedback forms, and at least 27% did not meet the minimum implementation threshold.[8] In addition, few studies conducted post-pandemic have been published to date, but anecdotally we know that many evaluators have been struggling with implementation since the pandemic. At the same time, the challenges encountered implementing SEL curricula in the schools in this study are not new. A 2018 report from the Harvard Graduate School of Education noted that "SEL lessons and other programs activities are often abridged or skipped due to tight schedules and competing priorities [...] and that "school administrators and staff sometimes perceive structured programs to be too 'top-down', and as a result, staff lack a sense of ownership and trust" (p. 4).[9]

In addition to these survey findings, we gathered feedback from the TRAILS Tier 1 implementation team members about their experiences in relation to the implementation of TRAILS SEL curriculum in program schools. We sought their views on the level of implementation and whether they had any knowledge about barriers to implementing the curriculum. Additionally, we requested feedback about how satisfied staff and administration in implementing schools were.

Their perceptions corroborated the findings from the staff surveys. The TRAILS implementation team cited issues related to communication as the biggest challenges for implementation. Implementation was most successful when there was a clear expectation of when lessons would be delivered and by whom. Capacity constraints and concerns relating to the age-appropriateness of the lessons may also have influenced program delivery. For example, a "high school teacher mentioned that lessons seemed more juvenile and it was harder to keep older students engaged with the material."

> "A couple of schools experienced admin changes midway through that stalled planning and rollout efforts of the curriculum. In other schools communication between admin and staff was not clear re: expectations with implementation [...]

Qualitative feedback from the implementation team showed mixed satisfaction with TRAILS SEL training. They said, "Generally speaking, staff were satisfied with the level of training they received." However, some schools expressed dissatisfaction with TRAILS. According to the implementation team "[their] contact at the district admits their dissatisfaction with training may have 'dimmed the TRAILS light' in their district". More specifically, some schools would have found it beneficial to "see a video of an SEL lesson being delivered to feel more comfortable delivering lessons on their own."

Student Reach

Finally, we explored the number of students who self-reported recognizing relevant curriculum components. Specifically, we explored the proportion of students who reported recognizing two important components of the TRAILS curriculum: a) the "Feelings Thermometer and/or Check-In Sheet" and b) the CBT Model. Here we drew from the student survey administered by TRAILS at the end of the 2022-23 school year.

Table 12 shows the proportion of students in the program and control group that indicated they recognized these two salient components of the TRAILS program. Students were first asked if they had seen or used the "Feeling Thermometer or the Check-In Sheet", and if they indicated in the affirmative, they were asked if they had seen or used the "CBT Model". The survey questions were displayed next to a picture of these components as they are used in the classroom.

Analysis showed that 40% of students in the program group said they have seen or used the Feelings Thermometer or Check-In Sheet. 34% reported having seen the CBT Model and 28% knew about both components. Some students in the control group also reported having seen or used TRAILS' components but did so to a lesser extent. 9% of students in the control group reported having seen the Feelings Thermometer and 4.5% said they knew both components.

Table 12: % of Students Recognizing TRAILS Components in End of Year Survey

Proportion of Students in Program Group Endorsing TRAILS Components	Program Group	Control Group
Feelings Thermometer/Check-in Sheet	40.2%	9.0%
CBT Model	34.3%	8.7%
Feelings Thermometer/Check-In Sheet & CBT Model	28.0%	4.5%
Number of Students Considered	2386	3354

Finding 1: Does TRAILS Have an Impact on Students' SEL Skills?

Students indicated how easy it was for them to make use of 26 social emotional skills that align with the following five core SEL competencies: self-awareness, social awareness, self-management, relationship skills, and responsible decision-making. The scores for the separate competencies were obtained by adding the scores within the respective domains.

Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on the combined SEL measure, for four of the five core SEL competencies: social awareness, self-management, relationship skills, and responsible decision-making (see Appendix Table D. 1 for detailed information).

Students who received TRAILS curriculum rated self-awareness skills easier to do in comparison to students in the control group.

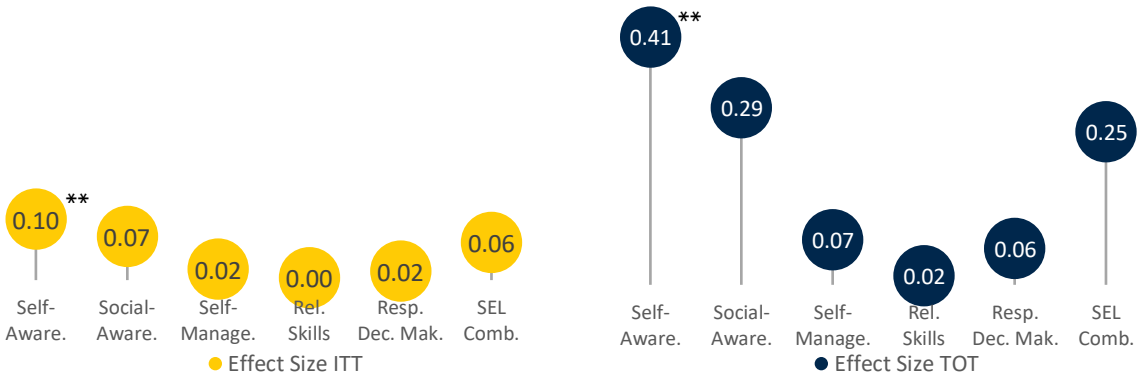
Analysis indicated a statistically significant difference for self-awareness. On a scale from six (very difficult to do this skill) to 24 (very easy to do this skill) the average score on the self-awareness scale was 17.07 for students in the program group and 16.70 for students in the control group (see Appendix Table D. 1). For reference, a 1-point increase on this scale corresponds to students rating one of the six skills comprising the self-awareness measure as easier to do. For example, students in the program group may have rated the skill "Know the emotions I feel" as "very easy" while students in the control group rated the same skill as "easy".

Figure 18 shows the unadjusted (ITT) and adjusted (TOT) effect sizes for all SEL measures. TRAILS SEL curriculum had an effect size of 0.10 (ITT) and 0.41 (TOT) for self-awareness. This means that an additional 4% (ITT) or 16% (TOT) of students in the program group rated self-awareness skills easier to do than the average student did in the control group on this measure.

Effect Size Interpretation

- 0.0 (program and control group have the same mean, no effect)
- 0.2 (additional 8% of program group above mean of the control group)
- 0.5 (additional 19% of program group above mean of control group)

Figure 18: Comparison of Effect Sizes for five SEL Core Competencies and the combined SEL measure for Intent to Treat (ITT) and Treatment on the Treated (TOT) Condition



Note: Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Finding 2: Does TRAILS Have an Impact on Students' Use of Effective Coping Skills?

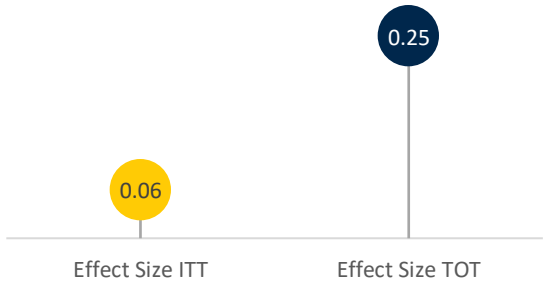
Students rated whether they knew and how often they used the following five effective coping skills to help them feel better when they are stressed or worried: Mindfulness, Cognitive Coping, Behavioral Activation, Listening to Music, and Exposure on a scale from 1="I don't know what this" is to 5="Often." The combined measure that reflects the use of these five effective coping skills was obtained by adding the score for each of the five items.

Students in the program group reported utilizing effective coping skills slighly more often than students in the control group. The difference between the two groups is not statistically significant.

Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on the combined measure reflecting use of effective coping skills. On a scale from five (I don't know what this is) to 25 (I use this skill often), the average score was 16.9 for students in the program group and 16.7 for students in the control group (see Appendix Table D. 1). Being in the program group increased the use of effective coping skills by 0.22 points on this scale. For reference, a 1-point increase on this scale corresponds to students utilizing one of the five coping skills comprising the combined measure more often. For example, students in the program group may have said they "sometimes" utilize the skill "Mindfulness" while students in the control group may have said they "rarely" used this skill.

Figure 19 shows the unadjusted (ITT) and adjusted (TOT) effect size for the combined measure reflecting use of five effective coping skills. TRAILS SEL curriculum had an effect size of 0.06 (ITT) and 0.25 (TOT) on this measure. In other words, an additional 2% (ITT) or 10% (TOT) of students in the program group used effective coping skills more frequently than the average student did in the control group (see Appendix Table D. 1).

Figure 19: Comparison of Effect Sizes for the Combined Measure of Effective Coping Skills for Intent to Treat (ITT) and Treatment on the Treated (TOT) Condition



Finding 3: Does TRAILS Have an Impact on Students' Perceived Embarrassment in Relation to Help-Seeking Behavior?

We asked students if they would be embarrassed if their friends knew they were getting help from a counselor for an emotional problem.

The proportion of students who report they would feel embarrassed is equivalent in both groups.

Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on the measure reflecting perceived embarrassment about receiving help from a counselor.

Analysis shows that 25% of students in both the program group and the control group said that they would be embarrassed if their friends knew they were getting help from a counselor (see Appendix Table D. 1). The effect size for this measure was 0.005 (ITT) and 0.02 (TOT), meaning that there was no meaningful difference on this outcome measure between the two groups.

Finding 4: Does TRAILS have an Impact on Symptoms of Depression or Anxiety?

Students in grades 6 to 12 responded to the two-item validated Patient Health Questionnaire (PHQ-2) measuring symptoms of depression and the Generalized Anxiety Disorder 2-item (GAD-2) screener measuring symptoms of anxiety on a scale of 0="Not at all" to 3="Nearly every day." A PHQ-2 and GAD-2 score was obtained by adding the scores for each of the two questions for the respective screener.

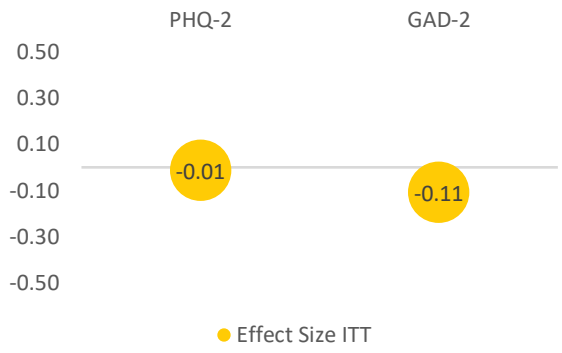
Students in the program group had lower levels of anxiety compared to students in the control group. Findings are not statistically significant.

Results indicated no statistically significant differences between the TRAILS program group and the business-as-usual control group on symptoms of depression or anxiety. While not significant, the results showed that students in the TRAILS program group had somewhat lower levels of anxiety, while levels of depression were nearly identical. On a scale from zero (no symptoms of depression) to six (major depressive disorder likely), the average depression score was 2.10 in the program group and 2.12 in the control group. On a scale from zero (no symptoms of anxiety) to six (major anxiety disorder likely), the average score for anxiety was 2.48 in the program group and 2.72 in the control group. Being in the program group reduced symptoms of anxiety by 0.23 points on this scale (see Appendix Table D. 2). For reference, a 1.0 decrease corresponds to students rating one of the two items on the depression or anxiety screener one point lower. For example, students in the program group may have said that they have been bothered by "Feeling nervous, anxious, or on edge" for "several days" while students in the control group may have rated the same statement with "more than half the days."

Figure 20 shows the unadjusted (ITT) effect sizes for the PHQ-2 depression and the GAD-2 anxiety measure. A negative effect size indicates a reduction in symptoms of depression (PHQ-2) or symptoms of anxiety (GAD-2). The effect size for PHQ-2 was -0.01 (ITT), meaning that there was no meaningful difference between the treatment and control group.

The effect size for GAD-2 was -0.11 (ITT), meaning that an additional 4% (ITT) of students were below the mean of the control group. Overall, we see that effect sizes are larger for the GAD-2 anxiety measure (see Appendix Table D. 2 for detailed information).

Figure 20: Comparison of Effect Sizes for PHQ-2 and GAD-2 Measure for Intent to Treat (ITT)



Finding 5: Does TRAILS' Impact on SEL Skills, Use of Effective Coping Skills, and Perceived Embarrassment Differ for Different Groups of Students?

Students responded to questions about their gender, their racial/ethnic identity, whether they spoke another language besides English at home, and their living arrangements. We analyzed the impact of TRAILS on SEL skills, effective use of coping skills, and embarrassment for the following groups: boys, girls, White students, Black students, intersectional groups (e.g. White Girls), and students who are experiencing homelessness or speak exclusively a different language at home.

TRAILS curriculum had the most profound effect on SEL skills for students who identified as Black, particularly Black girls.

Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on the combined SEL skills for boys, for students who identify as White, who are experiencing homelessness, or speak exclusively a different language than English at home. We did not find any statistically significant differences between the program and control group for any subgroups on the combined measure reflecting use of effective coping skills, nor the embarrassment measure (see Appendix Table D. 3 for detailed information).

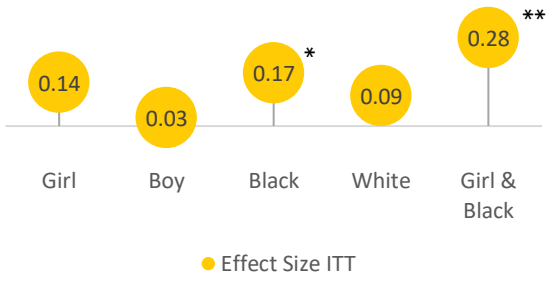
Analysis indicated two statistically significant differences for the combined SEL skills measure for students identifying as Black, as well as for the intersectional group of students identifying as Black girls.

Figure 21 shows the unadjusted (ITT) effect sizes for the combined SEL Skills measures for selected subgroups. The effect size for the combined SEL Skills was 0.17 (ITT) for Black students. This means that an additional 7% (ITT) of Black students in the program group rated skills easier to do than the average Black student in the control group on the combined SEL skills measure. The effect size for Black girls on the same measure was 0.28 (ITT), meaning that an additional 11% (ITT) of Black girls in the program group rated these things easier to do than the average Black girl in the control group (see Appendix Table D. 3 for detailed information).

Effect Size Interpretation

- 0.0 (program and control group have the same mean, no effect)
- 0.2 (additional 8% of program group above mean of the control group)
- 0.5 (additional 19% of program group above mean of control group)
- 0.8 (additional 29% of program group above mean of control group)

Figure 21: Comparison of Effect Sizes for the Combined SEL Skills Measure for Intent to Treat (ITT) for Selected Subgroups



Note: Statistical significance levels are indicated as follows:
*** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Finding 6: Does TRAILS' Impact on Symptoms of Depression Differ for Different Groups of Students?

We analyzed the impact of TRAILS on symptoms of depression for the following groups: boys, girls, White students, Black students, intersectional groups (e.g. White Girls), and students who are experiencing homelessness or speak a different language at home.

Several subgroups of students in the program group had lower levels of depression but the findings are not statistically significant.

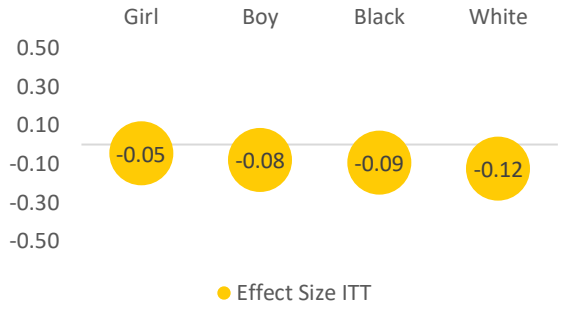
Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on the measures of depression for any of the subgroups (see Appendix Table D. 6 for detailed information). While not statistically significant, the results show that several subgroups of students in the TRAILS program group had lower levels of depression compared to students in the BAU control group. However, students in the program group who spoke exclusively another language at home had higher levels of depression compared to the control group.[14] The effect of TRAILS programming was most profound for White students. On a scale from zero (no symptoms of depression) to six (major depression likely), the average depression score was 1.75 in the program and 1.98 in the control group for White students. Being in the program group reduced symptoms of depression by -0.23 on this scale.

Figure 22 shows the unadjusted (ITT) effect sizes for the PHQ-2 depression screener for selected subgroups. A negative effect size indicates a reduction in symptoms of depression (PHQ-2). While not statistically, significant we see that the effect size of -0.12 (ITT) for the PHQ-2 was largest for students who identified as White, meaning that an additional 5% (ITT) of White students in the program group scored below the average White student in the control group (see Appendix Table D. 6 for detailed information).

Effect Size Interpretation

- 0.0 (program and control group have the same mean, no effect)
- -0.2 (additional 8% of program group below mean of the control group)
- -0.5 (additional 19% of program group below mean of control group)
- -0.8 (additional 29% of program group below mean of control group)

Figure 22: Effect Sizes for the PHQ-2 Measure for Intent to Treat (ITT) and Treatment on the Treated (TOT) for Selected Subgroups



[14] Results for students speaking exclusively another language indicated statistical significance ($p < 0.05$). However, caution should be exercised in interpreting the findings due to the small sample size for this group.

Finding 7: Does TRAILS' Impact on Symptoms of Anxiety Differ for Different Groups of Students?

We analyzed the impact of TRAILS on symptoms of anxiety for the following groups: boys, girls, White students, Black students, intersectional groups (e.g. White Girls), and students who are experiencing homelessness or speak a different language at home.

Students in the program group had lower levels of anxiety regardless of their background characteristics. Effects for anxiety were most profound for White students in the program group, but are not statistically significant.

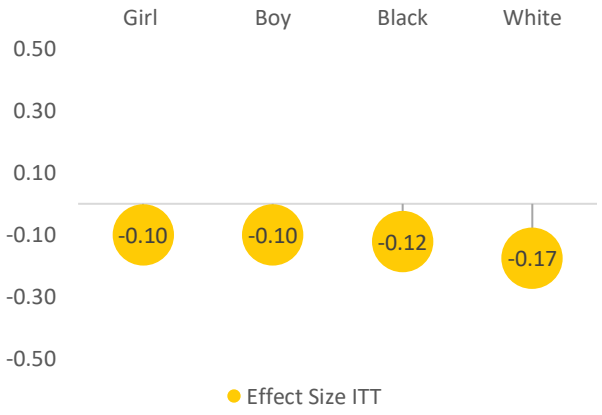
Results indicated no statistically significant difference between the TRAILS program group and the business-as-usual comparison group on symptoms of anxiety for any of the subgroups (see Appendix Table D. 7 for detailed information). While not significant, the results show that students in the TRAILS program group had lower levels of anxiety regardless of their gender, race, language, or housing status. The effect was most profound for White students. On a scale from zero (no symptoms of anxiety) to six (major anxiety disorder likely), the average anxiety score for this group was 2.45 in the program and 2.83 in the control group. Being in the program group reduced symptoms of anxiety by -0.39 on this scale for students who identified as White.

Figure 23 shows the effect sizes for the GAD-2 anxiety screener. A negative effect size indicates a reduction in symptoms of anxiety (GAD-2). While not statistically significant we see that the effect size -0.17 (ITT) for the GAD-2 anxiety screener was largest for White students. In other words, an additional 7% (ITT) of White students in the program group scored below the average White student in the control group (see Appendix Table D. 7 for detailed information).

Effect Size Interpretation

- 0.0 (program and control group have the same mean, no effect)
- -0.2 (additional 8% of program group below mean of the control group)
- -0.5 (additional 19% of program group below mean of control group)
- -0.8 (additional 29% of program group below mean of control group)

Figure 23: Effect Sizes for the GAD-2 Measure for Intent to Treat (ITT) and Treatment on the Treated (TOT) for Selected Subgroups



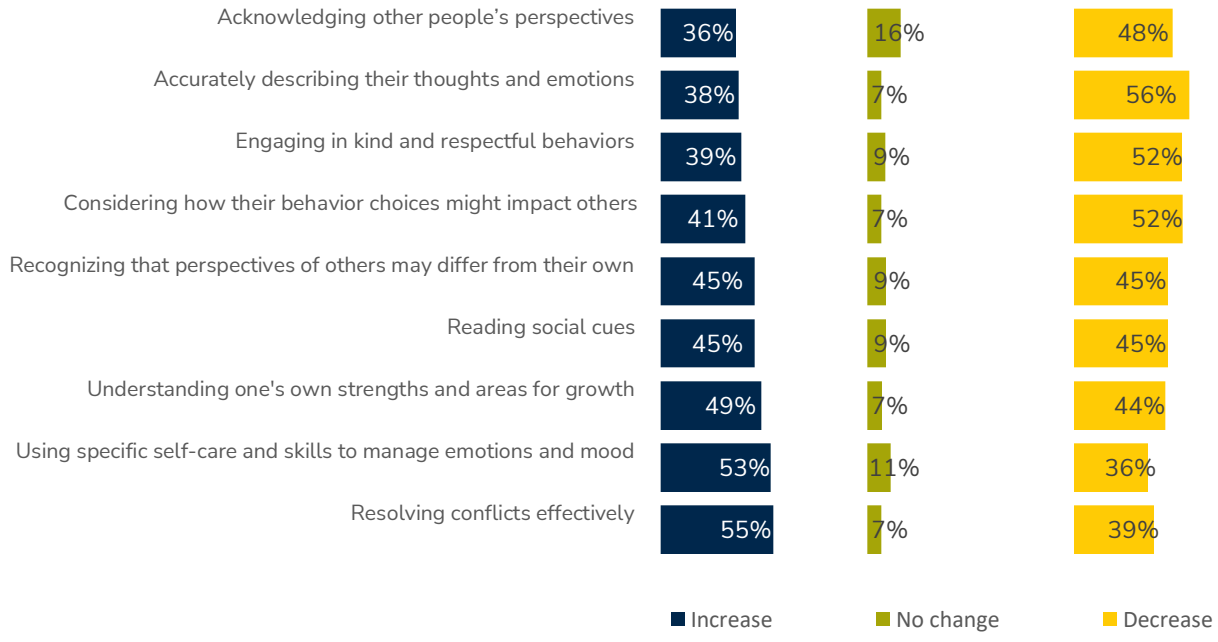
Additional Findings

In this section, we summarize some additional findings related to staff perspectives on students' growth in social and emotional skills and their viewpoint on their own growth in confidence teaching SEL skills. We also report on changes in staff reports on the availability of interventions and trainings in their schools, and their participation in professional development activities focusing on mental health related topics.

What was TRAILS SEL Implementers' Perspective on Students' Growth in SEL Skills at the End of the 2022-23 School Year?

TRAILS conducted a survey before training staff in their SEL curriculum and a survey after curriculum rollout at the end of the 2022-23 school year. Both surveys included questions about staff perceptions of various students' social and emotional competencies. Staff were asked to indicate the percentage of their current students who consistently demonstrate age-appropriate competency in a list of social and emotional skills. By matching the data of staff participating in the pre-training and post-implementation surveys, we calculated the proportion of staff who observed an increase, a decrease, or no change in the percentage of their students demonstrating age appropriate SEL competencies (see Figure 24). We saw that about half of staff who delivered TRAILS SEL lessons to students reported an increase in students consistently showing the following skills: "Resolving conflict effectively" (55%), "Using specific self-care skills to manage emotions and mood" (53%), and "Understanding one's own strengths and areas for growth" (49%).

Figure 24: SEL Implementers Perspective on Proportion of Students Who Consistently Show Age-Appropriate SEL Competencies

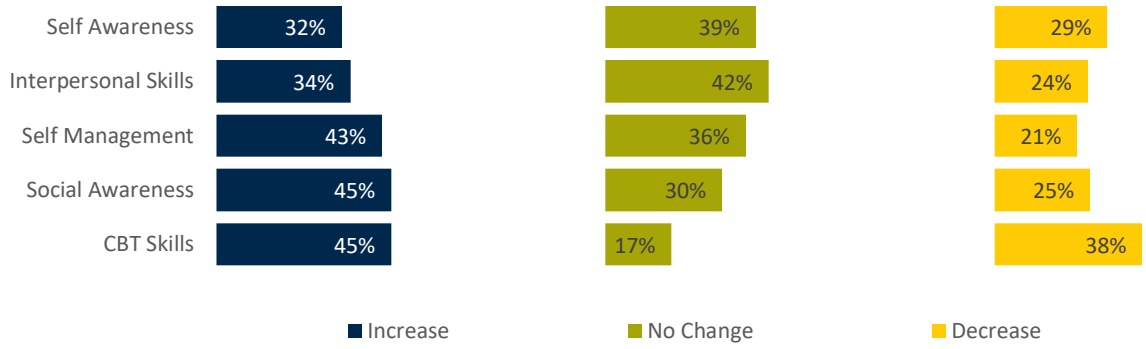


NOTES: The figure represents staff who answered the listed questions in the pre-training and post-implementation survey. The number of matched responses varied by question (n=43 to n=46). Percentages may not add up to 100% because of rounding.

What was TRAILS SEL Implementers' Perspective on Their Own Growth in Confidence in Teaching SEL Skills?

Prior to receiving TRAILS SEL training and after program rollout, staff were asked about their level of confidence teaching a list of skills relating to self-awareness, interpersonal skills, social awareness, self-management, and cognitive and behavioral skills. We calculated a composite score for the set of skills relating to these main SEL competencies and explored the proportion of staff whose scores increased, decreased, or did not change from before to after program rollout. We saw that the greatest proportion of trainees reported growth in CBT skills (45%) and social-awareness (45%), followed by self-management (43%) (see Figure 25). A notable 38% of staff reported a decreased level in confidence teaching CBT skills. This finding may reflect an initial overestimation of skills prior to training. Staff may have been more confident in teaching certain skills due to limited knowledge in relation to that skill. Expanding one's knowledge about these skills may recalibrate self-perception of expertise. Overall, we saw more people reporting growth or no change in confidence in teaching SEL skills than reporting a decrease in confidence teaching these skills.

Figure 25: Proportion of SEL Implementers Who Reported an Increase, No Change, or Decrease in Confidence Teaching SEL Skills



NOTES: The figure represents staff who answered the list of questions in the pre-training and post-implementation survey. The number of matched responses varied by skill (n=58 to n=62). Percentages may not add up to 100% due to rounding. We generated a composite scale for each of these five skills for the pre-training and post-implementation survey and analyzed whether there was an increase, decrease, or no change from time one to time two.

How did Staff in the Program Group Report on Indicators Relating to Practices and Interventions Prior to Program Rollout and Post-Implementation?

We asked staff in program schools to report on the availability of mental health related interventions, trainings and supports in their school, and on their participation in professional development in the last 12 months. Our findings revealed an increase in staff reporting the availability of all interventions, all trainings and supports, and participation in all professional development activities (see Table 13). We saw the greatest increase in staff reporting on the availability of Tier 1 universal and Tier 2 targeted mental health interventions. There was a 25.7 percentage point increase in staff reporting on the availability of Tier 1 and a 21.6 percentage point increase on Tier 2 interventions. Additionally, we saw an increase in the proportion of staff reporting on their participation in professional development activities focusing on social and emotional learning. Prior to program rollout, 66% of staff reported they participated in such activities, whereas at the end of the school year nearly 83% reported they had participated.

Table 13: Proportion of Staff in Program Group Reporting on Available Interventions, Trainings & Support, and Participation in Professional Development Activities Before and After Program Roll Out in School Year 22-23

	Prior to Program Rollout	After Program Rollout	Difference
Available Interventions (%)			
Universal, school wide screening for student mental health	27.2	36.2	+9.0
Tier 1 universal mental health interventions delivered to all students (e.g., SEL)	49.1	74.8	+25.7
Tier 2 targeted mental health interventions for students with mild to moderate need (e.g. counselor-led programs)	52.7	74.3	+21.6
Tier 3 intense interventions for students with severe need (e.g., protocols for evaluating and responding to suicide risk)	52.4	61.4	+9.0
Available Training and Support (%)			
Voluntary continuing education opportunities for instructors in student mental health or social and emotional learning	36.3	42.4	+6.1
Mandatory training for instructors in student mental health or social and emotional learning	43.2	47.8	+4.6
Access to consultants or coaches to support staff in implementing school-based mental health interventions	46.5	49.6	+3.1
Professional Development (%)			
Student mental health	52.3	62.7	+10.4
Social emotional learning	66.2	82.9	+16.7
Trauma or post-traumatic stress disorder	49.3	49.6	+0.3
Sample Size			
Staff[a]	287	251	

NOTES: The table reflects the proportion of staff who reported that any of the listed interventions and trainings and supports were either a) available, but not implemented, b) available, but need support to implement it well, or c) available and implemented well. For professional development, we report staff who have participated at least once in the last 12 months in one of the listed PDs. Responses for these questions could not be matched and thus, the percentages do not necessarily reflect the same respondents. Rounding may cause slight discrepancies in sums and differences.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 95% of the sample.

Conclusion

The findings of the evaluation underscore the challenges of implementing SEL programs in schools, especially in a post-pandemic environment. Teachers have substantial demands on their time, including trying to compensate for lost instructional time during the COVID-19 pandemic, and many schools continue to face staffing challenges.[10] Further, students and parents may be resistant to SEL programming which can pose an additional barrier to implementation.[11] Without support and encouragement by school leaders, including principals and superintendents, SEL programming is not likely to be a priority for teachers in the classroom. This is true despite the widely reported struggles students are having with both mental health and social emotional skills post-pandemic. Making the link between SEL skills and other student outcomes, including student behavior and academic achievement, could help teachers see benefits of SEL programming.

Despite this, we observed some positive trends in this data. Even with low levels of implementation and a high degree of SEL programming in the comparison schools—both of which have the potential to dilute impacts—we find some positive and statistically significant impacts on students. The positive impacts for Black students, and Black girls in particular, are especially encouraging. In environments where teachers and other school staff have the support they need to implement the program fully, the program has the potential to positively impact students' behavioral health.

Limitations

Although these results are encouraging, the study also has several limitations and so the findings should be interpreted with caution.

First, there was a considerable amount of attrition from the study and relatively high differential attrition with only 19 of 27 schools in the program group responding to the end of year student survey.[15] This higher non-response rate from program schools could lead to bias into our results. For example, if the non-responding program schools had students who were particularly struggling with SEL, while all control schools responded, it could make the TRAILS program appear more effective than it actually is.

Second, survey findings indicated that a considerable amount of SEL programming was taking place in the control schools. If students in these schools were receiving high quality SEL, equivalent to TRAILS, we are unlikely to observe any significant differences between the two sets of schools. The results must be contextualized to account for the fact that we are comparing the TRAILS program to an environment where substantial SEL efforts may have already been in place.

Finally, we are not able to link individual student responses on baseline and follow-up surveys, meaning that we can only control for baseline skills at the school level and not at the individual level. This may mean that we have not adequately accounted for any baseline differences between the program and control groups at the beginning of the study.

[15] Differential attrition differed by type of regression analysis. However, non-response rate from program groups was consistently higher compared to program schools regardless of regression analysis. Thus, the implication are the same.

Recommendations

Given the study findings, we have identified the following recommendations.

Recommendation 1: Consider providing more support for program implementation

To date, TRAILS has focused on providing high quality training and materials but has placed more limited emphasis on implementation support. TRAILS may want to explore barriers to program implementation more deeply and potentially provide more implementation support for programming, including exploring ways to increase buy-in and communication across different levels of the school organization—i.e., school building staff, principals, and administrators; district superintendents; and intermediate school districts.

Recommendation 2: Explore whether there are ways to increase the relevance of the curriculum

Staff reported that some students were uncomfortable, resistant, or not meaningfully engaged with the SEL curriculum. TRAILS might explore the causes of this by engaging students from different grade levels and/or teachers through discussion, focus groups, or anonymous surveys to better understand their perspectives and identify ways to make the curriculum more relevant and engaging for students.

Recommendation 3: Review training approach and materials

Staff who participated in the Tier 1 training had mixed reactions. TRAILS may want to explore the aspects of the training and material that worked well and those that were less effective, including whether some delivery modes were more effective than others were. Considering those findings TRAILS could potentially revise some aspects of the training program.

Recommendation 4: Identify measurable goals around SEL skills and align measures with goals

The analysis indicated that students showed greater confidence in skills associated with self-awareness compared to skills in domains like responsible-decision making. This discrepancy may be a result of a variety of factors including a greater emphasis on skills that focus on self-awareness within the curriculum, an introduction of self-awareness concepts earlier in the sequence of the curriculum, or simply that the outcome measure was more aligned with evaluating self-awareness compared to other SEL concepts. TRAILS should continue to explore outcome measures that are well-aligned with their curriculum. We also suggest exploring the frequency and repetition of certain activities within the curriculum. Skills that are introduced early and reinforced often likely lead to greater confidence among students in those areas.

Appendix A: Methodology

Survey Instrument

The Youth Policy Lab with input from TRAILS developed web-based survey instruments for students and staff that were administered prior to program rollout at the beginning of the 22-23 school year and after program implementation at the end of the school year.

All students in grades 4-12 who either received the SEL curriculum (program group) or were slated to receive the SEL curriculum in the following year (control group) were eligible to participate in the survey. Two grade-band specific student survey versions were developed that differed in length and content based on grade level, age of students, and time of administration of the survey. Schools could choose whether their 6th grade students would receive the longer or the shorter survey. Five schools chose the longer survey for their 6th grade students. In addition, schools could also choose if students in their school would receive the default demographic or the simplified demographic questions. All students in schools that choose the default demographic questions would receive the multiple choice gender question. In addition, students in grade 9 through 12 in these schools would also receive the sexual orientation and transgender question. Students in schools who chose the simplified version would only answer the following open-ended question: "What is your gender? Please describe here (for example, 'girl') _____." 26 schools in our sample preferred the simplified demographic questions. The full list of survey questions can be found in the student survey administration manual.

Staff surveys were open to all instructional staff and School Mental Health Professionals (SMHP) in all grades in participating schools. All staff were asked questions about existing school programming and screening/referral protocols, self-reported burnout/exhaustion, stigma as it relates to help-seeking, perception of the school's support for mental health programming, professional development participation, and perception of students' classroom behavior and coping skills. The End-of-Year (EOY) survey version is slightly longer than the Beginning-of-Year (BOY) version. The full list survey questions in the staff survey can be found in the staff survey administration manual.

Data Collection and Analysis

Prior to survey administration schools were asked to identify one or multiple point-persons in schools that would help with student survey and staff survey administration, as well as managing opt-outs and other technical aspects of the survey administration process. A comprehensive survey administration manual was developed and shared with the schools prior to survey administration and before the 22-23 school year started. The administration manuals included suggestions for survey administration, eligibility criteria for students and staff to participate, sample emails and backpack letters to be shared with parents of eligible children, the list of survey questions, the link to the student and staff surveys, resources for students, a FAQ document, and other resources.

The study team strongly encouraged schools to administer the surveys to eligible students and staff before implementing the TRAILS SEL curriculum in program schools and again at the end of the 22-23 school year. The period for student and staff survey administration at the start of the school year varied by school and lasted from August 2022 to January 2023. Survey administration at the end of the 22-23 school year lasted from April to June 2023.

When the BOY student survey closed there were 7,385 recorded responses of which 830 were less than 50% completed and dropped from analysis. At the end of the school year we collected 5,929 responses from students of which 510 were less than 50% completed and dropped from analysis. When the BOY staff survey closed there were 941 recorded responses of which 49 were less than 50% completed and automatically dropped from analysis. At the end of the school year we collected 731 responses from staff of which 81 were less than 50% completed and dropped from analysis.

Survey items were analyzed using Stata by staff at the Youth Policy Lab. Open-ended text was first analyzed and then recoded into existing categories whenever possible. Guidelines for coding open-ended text responses and for cleaning, appending, and analyzing surveys were shared with TRAILS. Where possible, the survey used already validated measures and survey items. This included measures related to social and emotional learning competencies (WCSD SECA), mental health screeners (Patient Health Questionnaire-2 and General Anxiety Disorder-2), and measures of staff burnout (Oldenburg Burnout Inventory).

Treatment on the Treated Analysis

To calculate the factor by which to scale up the ITT estimate for our main outcomes, we identified the proportion of students in both the program and control group who recognized both TRAILS components. We restricted the sample to students in schools with at least five responses on the end of year survey and to those who responded to all 26 SEL items. A total of 665 students in the program group and 150 students in the control group indicated that they recognized the two TRAILS components. We adjusted the total number of students in the program group by subtracting the number of students in the control group who also indicated they had seen the TRAILS components. Based on this calculation about 25% of students in the program group likely received the TRAILS program. To calculate the Treatment on the Treated (TOT) effect size we divided the ITT effect size by the proportion of program students who received the program. For example, the ITT effect size of outcome "self-awareness" was 0.10. We divided the ITT effect size of 0.10 by ~25% to obtain a TOT effect size of 0.41.

Student Outcome Measures

Table A.1 shows the list of the questions and response options of student outcome measures as used in the regression analysis and Table A.2 provides additional details, such as psychometric properties.

Appendix Table A. 1: List of Student Outcome Measure and Related Survey Questions

Outcome	Individual Questions	Response Options
SEL Competence: Self-Awareness	How easy is it for you to____: • Know when my feelings are making it hard for me to focus • Know the emotions I feel • Know ways to make myself feel better when I am sad • Notice what my body does when I am nervous • Know when my mood affects how I treat others • Know ways to calm myself down	1= Very difficult, 2= Difficult, 3= Easy, 4= Very easy
SEL Competence: Social-Awareness	How easy is it for you to____: • Learn from people with different opinions than me • Know what people may be feeling by the look on their face • Know when someone needs help • Know how to get help when I'm having trouble with a classmate • Know how my actions impact my classmates	1= Very difficult, 2= Difficult, 3= Easy, 4= Very easy
SEL Competence: Self-Management	How easy is it for you to____: • Get through something even when I feel frustrated • Be patient even when I am really excited • Stay calm when I feel stressed • Work on things even when I don't like them	1= Very difficult, 2= Difficult, 3= Easy, 4= Very easy

SEL Competence: Relationship Skills	How easy is it for you to____: • Respect a classmate's opinions during a disagreement • Get along with my classmates • Share what I am feeling with others • Talk to an adult when I have problems at school • Be welcoming to someone I don't usually eat lunch with • Get along with my teachers	1= Very difficult 2= Difficult 3= Easy 4= Very easy
SEL Competence: Responsible Decision-Making	How easy is it for you to____: • Think about what might happen before making a decision • Know what is right or wrong • Think of different ways to solve a problem • Say "no" to a friend who wants to break the rules • Help make my school a better place	1= Very difficult 2= Difficult 3= Easy 4= Very easy
TRAILS CBT Skills Checklist	When you are stressed or worried, how often do you use the following skills to help you feel better? • Mindfulness or relaxation strategies (ex. mindful eating or meditation) • Listening to calming or happy music • Behavioral Activation (doing something active for at least 10 minutes (ex. sports, dancing, walking, running, bicycling, etc) • Cognitive Coping (questioning your automatic negative thoughts (ANTs)) • Exposure (overcoming avoidance by learning to face your fears)	1= I don't know what this is 2= Never 3= Rarely 4= Sometimes 5= Often
Perceived embarrassment in relation to help-seeking behavior	Would you be embarrassed if your friends knew you were getting help from a counselor for an emotional problem?	1= Yes 2= No
Symptoms of Depression	Over the past 2 weeks, how often have you been bothered by any of the following? • Little interest or pleasure in doing things • Feeling down, depressed, or hopeless	0= Not at all 1= Several days 2= More than half of the days 3= Nearly every day or every day
Symptoms of Anxiety	Over the past 2 weeks, how often have you been bothered by any of the following? • Feeling nervous, anxious or on edge • Not being able to stop or control worrying	0= Not at all 1= Several days 2= More than half of the days 3= Nearly every day or every day

Appendix Table A. 2: Description of Student Outcomes Included in Regression Analyses

Outcomes	Measure	Description	Previous Use/Psychometric Properties
Social Emotional Learning (SEL) Competence	WCSD SECA	The WCSD SECA is a set of instruments developed through a collaboration between Washoe County School District, the Collaborative for Academic and Social Emotional Learning (CASEL), and the University of Illinois at Chicago through an Institute of Education Sciences Research-Practitioner Partnership grant. The project resulted in the development of two instruments and a bank of items aligned to the CASEL 5 clusters and WCSD SEL standards. Developed in 2012. https://www.washoeschools.net/cms/lib/NV01912265/Centricity/Domain/231/WCSD_SECA_FAQ_Format_October_2018_FINAL.pdf. We have used 26 items form the WCSD SECA instrument that reflect five SEL competencies: Self-Awareness (6 items), Social-Awareness (5 items), Self-Management (4 items), Relationship skills (6 items), and Responsible-Decision Making (5 items). Students rated this list of competencies from a scale from 1 (very difficult) to 4 (very easy).	**Evidence based on content** Items were developed based on district SEL standards and test items were then placed into appropriate grade-level bands by research team members.[12] **Evidence based on response processes** Student focus groups were used to improve item readability and to ensure that items assessed high levels of SEL competencies.[13] **Evidence based on internal structure** Rasch modeling was used to show that differential item functioning by grade level, gender, and race/ethnicity was negligible.[14] **Evidence based on relations with other variables** Regression models show positive association among scales and academic achievement (standardized test scores and GPA) as well as lower rates of suspension and absenteeism.[15]
Utilization of effective coping skills	CBT Skills Checklist	The CBT skill Checklist is an original measure developed by TRAILS that identifies the five core CBT coping skills taught in TRAILS-EI (mindfulness, cognitive coping, behavioral activation, exposure, and listening to music). For each CBT skill listed, students are asked to use a Likert scale to indicate how likely they would be to use the skill if they felt down or depressed. The structure of the measure allows to detect anticipated utilization of targeted coping skills and whether the respondent has been exposed to the skill through teaching or clinical opportunities. The rating scale ranges from 1 (I don't know what this is) to 5 (Often).	Used with many TRAILS initiatives. For example, the CBT Skills Checklist was administered to nearly 11,000 DPSCD students in October 2019 as part of a community needs assessment.
Perceived embarrassment	Single Question	We use a stand-alone item to gather students' perception about outreach for help and related stigma. We asked	

in relation help-seeking behavior		students "Would you be embarrassed if your friends knew you were getting help from a counselor for an emotional problem?" [Response Options: Yes, No]	
Symptoms of Depression	Patient Health Questionnaire (PHQ-2)	This measure is a brief self-report tool, widely used in both research and a variety of clinical and community settings, to screen for and evaluate the severity of symptoms of depression. It includes the first 2 items of the longer PHQ-9 depression screener. A PHQ-2 score ranges from 0-6 with a score of 3 or higher indicating that a major depressive disorder is likely.	Richardson et al. report a sensitivity of 73% and specificity of 75% for detecting major depression among adolescents.[16]
Symptoms of Anxiety	Patient Health Questionnaire (GAD-2)	This measure is a brief self-report too, widely used in both research and a variety of clinical and community settings to screen for and evaluate symptoms of anxiety.	Plummer et al. report a sensitivity of 76% and a specificity of 81% for detecting anxiety disorders with a cutoff of 3.[17]

Appendix B: Demographics

Student and Staff Demographics at the End of the School Year 22-23

Appendix Table B. 1: Comparison of Student Characteristics at End of School Year 22-23

Characteristics	Program Group	Control Group	Difference
County (%)			
Genesee	63.10	42.80	20.30
Wayne	36.90	57.20	-20.30
Gender[a] (%)			
Female	44.60	47.08	-2.47
Male	49.95	47.22	2.73
Gender non-binary	1.56	2.03	-0.47
Questioning/unsure	0.45	0.81	-0.35
Prefer not to say	1.61	1.82	-0.21
I don't know	0.35	0.28	0.07
Gender not listed, other[b]	1.46	0.77	0.69
Race/Ethnicity (%)			
American Indian or Alaska Native	1.16	1.78	-0.61
Asian	1.42	1.05	0.37
Black or African American	18.15	21.25	-3.10
Hispanic or Latinx	2.88	3.07	-0.18
Middle Eastern or North African	6.77	4.53	2.24
Native Hawaiian or Pacific Islander	0.40	0.31	0.09
White	36.10	35.47	0.63
Race/Ethnicity not described, other[c]	2.38	2.96	-0.59
Prefer not to say	15.02	14.32	0.69
Multiracial	15.72	15.26	0.46
Grade (%)			
Grades 4-5	19.25	26.55	-7.30
Grades 6-8	61.32	65.76	-4.44
Grades 9-12	17.56	7.28	10.28
Other	1.87	0.22	1.65
Self or People They Live With Born Outside the US (%)			
Yes	26.25	23.97	2.28
No	60.10	60.95	-0.85
Prefer not to say	13.65	15.08	-1.43
Homelessness[d] (%)			
Students who experience homeless	15.46	13.62	1.84
Main Language Used at Home (%)[e]			
English Only	70.65	75.36	
Multilingual	24.92	20.76	
Other Language	4.42	3.88	0.54
Sample Size[f]			
Schools	19	21	
Children	2244	3175	

NOTES: Rounding may cause slight discrepancies in sums and differences.

[a]Schools could choose how the gender question was displayed in the survey. The choices included a multiple-choice question where students would check any of the listed gender options, or an alternative open-ended question

where students were asked to write in their gender. All redundant information was recoded into existing gender groups, novel information was included as "other", and non-meaningful information was classified as missing.
[b]This variable captures students who entered novel information that could not be recoded into existing groups.
[c]This variable captures students who entered novel information that could not be recoded into existing groups.
[d]McKinney Vento Act Students checked a list of items indicating their living arrangements in the past year. We followed the McKinney Vento Act recommendation for the definition of homelessness (see here: https://nche.ed.gov/wp-content/uploads/2018/10/conf-elig.pdf). A student was classified as homeless if they selected any of the following living arrangements: At someone's else's home, At a hotel/motel, At a shelter, In a car, van, or other vehicle, Outside, or Where I live changes frequently.
[e] We asked students to identify their main language used at their home. A student who checked multiple languages was classified as multilingual. A student whose main language was different from English was classified as ¨Other language". Other languages included Albanian, Arabic, Farsi, French, Russian, Spanish, Vietnamese
[f] The denominator for each question varies. For all variables in the table, data are available for at least 86% of the sample.

Appendix Table B. 2: Comparison of Staff Characteristics at End of School Year 22-23

Characteristics	Program Group	Control Group	Difference
Gender (%)			
Female	76.52	74.70	1.82
Male	14.35	15.96	-1.62
Gender non-binary	0.87	0.30	0.57
Prefer not to answer	7.83	9.04	-1.21
Gender not listed, other[a]	0.43	0.00	0.43
Race/Ethnicity (%)			
American Indian or Alaska Native	0.00	0.30	-0.30
Asian	1.75	0.00	1.75
Black or African American	3.06	6.33	-3.27
Hispanic or Latinx	2.18	0.00	2.18
Middle Eastern or North African	5.24	0.90	4.34
Native Hawaiian or Pacific Islander	0.00	0.60	-0.60
White	71.18	78.01	-6.83
Race/Ethnicity not described, other[b]	0.44	0.00	0.44
Prefer not to say	9.61	11.14	-1.54
Multiracial[c]	6.55	2.71	3.84
Grades (%)[c]			
Grade PK-2	27.85	17.33	10.52
Grade 3-5	24.05	26.42	-2.37
Grade 6-8	51.90	55.40	-3.50
Grade 9-12	17.30	22.44	-5.14
Professional Role (%)			
Instructional Staff	81.75	83.33	-1.58
School Mental Health Professionals	5.47	10.17	-4.70
Other[d]	12.77	6.50	6.28
Years of Experience in Professional Field (%)			
Less than 1 year	2.92	3.68	-0.77
1-5 years	20.83	17.56	3.27
6 years or more	76.25	78.75	-2.50
FTE (%)			
Full-time	97.49	97.74	-0.25
Sample Size			
Staff[e]	278[e]	355	

NOTES: Rounding may cause slight discrepancies in sums and differences.
[a]This variable captures staff who entered novel information that could not be recoded into existing groups.
[b]This variable captures staff who entered novel information that could not be recoded into existing groups.
[c]This variable captures grade bands staff taught. Staff could select multiple grade-bands.
[d]This variable captures other professional roles that could not be recoded into existing groups.
[e]The denominator for each question varies. For all variables in the table, data are available for at least 82% of the sample. Staff who self-identified in the survey to have received TRAILS training were redirected (n=147) to the combined survey. 129 staff started with the combined survey, 149 staff continued with the evaluation survey.

Appendix C: Implementation Findings

Data from TRAILS Post-Implementation Survey

Appendix Table C. 1: TRAILS Trainees Report: Preparedness to Deliver SEL Curriculum and Context of Lesson Delivery after Program Rollout at the End of the School Year 22-23

	Program Group %
Adequately Prepared to Deliver TRAILS SEL Curriculum (n=50)	
Staff Who agreed/strongly agreed to the questions that they felt adequately prepared to deliver TRAILS SEL curriculum	74.0
% of Staff Reporting on That Following Aspects Would Have Helped to Feel More Prepared to Deliver TRAILS Material to Students (n=69)	
More training on the cognitive & behavioral skill concepts embedded in the curriculum	26.1
More training on specific lessons	17.4
Opportunities to practice or observe example lessons or activities	23.2
Guidance on how to flexibly implement the curriculum	15.9
More training regarding the goals of SEL	13.0
Strategies to help me evaluate my lesson fidelity or adherence	5.8
More encouragement from administrators and/or supervisors in my school/district	8.7
Other	5.8
Nothing, I felt well prepared to deliver TRAILS SEL materials	7.3
Lesson Delivery Time (n=50)	
Average time spend on delivering a TRAILS lesson (Minutes)	14.5 (8.6)
Context of Delivering TRAILS SEL instruction? (n=69)	
One-on-one with individual students	5.8
With groups of students	21.7
In a homeroom or advisory class	42.0
In core content classes (e.g., Health, Social Studies)	5.8
Sample Size	
Staff[a]	69

NOTES: This table represents data from staff from TRAILS post-implementation survey. Staff that taught any grades between 4 and 12 and those that participated in the pre-training survey are considered for this analysis.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 72% of the sample.

Appendix Table C. 2: TRAILS Trainees Report: Satisfaction of TRAILS Program after Program Rollout at the End of the School Year 22-23

	Program Group %
Likelihood to Recommend the TRAILS SEL Curriculum to a Friend or Colleague on Scale from 0 (not very likely) to 10 (very likely). (n=66)	
0-3	12.1
4-6	37.9
7+	50.0
Overall Satisfaction with TRAILS Lessons Staff Delivered (n=46)	
Extremely dissatisfied/somewhat dissatisfied	8.7
Neither	34.8
Satisfied/extremely satisfied	56.5
% of Staff Reporting to Like the Following Aspects of TRAILS Lessons Since the TRAILS SEL Training (n=52)	
Lesson length	25.0
Included materials	53.9
Lesson format	34.6
Suggestions for incorporating content into classroom routines	17.3
% of Staff who Agreed/Strongly Agreed With the Following Statements About the Curriculum and Materials (n=41)	
My students find the lessons engaging.	51.2
Classroom behavior has improved since we began using the curriculum.	46.3
The lessons are positively impacting student wellness.	75.0
The lessons effectively teach self-care skills that are beneficial to students.	84.6
The lessons are developmentally appropriate.	87.5
The curriculum would be appropriate for students of diverse cultures, races, and ethnicities.	92.5
Delivering the lessons requires minimal prep time.	80.0
The lessons are easy to deliver.	95.0
Agreed/Strongly Agreed With the Following Statements About Satisfaction (n=51)	
Being able to teach the TRAILS SEL curriculum to my students makes me feel empowered as a teacher.	74.5
I would recommend the TRAILS SEL curriculum to others.	78.9
TRAILS curriculum is low burden.	78.4
TRAILS curriculum is flexible.	88.2
Sample Size	
Staff	66

NOTES: This table represents data from staff from TRAILS post-implementation survey. Staff that taught any grades between 4 and 12 and those that participated in the pre-training survey are considered for this analysis.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 62% of the sample.

Appendix Table C. 3: TRAILS Trainees Report: Barriers to Implementation After Program Rollout at the End of the School Year 22-23

	Program Group %
% of Staff Reporting on Barriers They Experienced Related to Providing SEL Instruction (n=69)	
Students are uncomfortable, resistant, or have not engaged meaningfully	36.2
Lack of time due to other teaching duties	26.1
Lack of class time for SEL instruction	24.6
Not enough time to prep SEL lessons	13.0
Limited administrative support or an unsupportive school climate	10.1
Lack of confidence in my ability to teach SEL	10.1
I have not experienced significant barriers related to providing SEL instruction	8.7
SEL instruction is inappropriate for my role	5.8
I don't believe SEL is necessary / important for students	1.5
Sample Size	
Staff	69

NOTES: This table represents data from staff from TRAILS post-implementation survey. Staff that taught any grades between 4 and 12 and those that participated in the pre-training survey are considered for this analysis.

Appendix D: Detailed Regression Analysis Results

Appendix Table D. 1: HLM Conditional Regression Model: Proximal Outcomes (Social and Emotional Learning Measures, Effective Coping Skills, Embarrassment)

Main Outcomes	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size (ITT)[a]	Effect Size (TOT)[b]
SEL Core Competencies									
Self-Awareness[c]	5064	17.07	16.70	.36	.006	**	.13	.10	.41
Social-Awareness[d]	4996	14.57	14.37	.20	.078		.11	.07	.29
Self-Management[e]	5131	9.65	9.60	.05	.616		.10	.02	.07
Relationship Skills[f]	4947	15.99	15.98	.01	.932		.16	.00	.02
Responsible Decision-Making[g]	4956	14.07	14.02	.05	.745		.14	.02	.06
SEL Combined Measure									
Combined SEL[h]	4625	71.61	70.84	.77	.140		.52	.06	.25
Effective Coping Skills									
Combined Effective Coping Skills[i]	5201	16.93	16.71	.22	.093		.13	.06	.25
Perceived Embarrassment									
Embarrassment[j]	5261	0.25	0.25	.00	.900		.02	.00	.02
Sample Size									
Schools	32	16	16						

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Impact were estimated by comparing outcomes for the group assigned to TRAILS SEL curriculum with corresponding outcomes for the business as usual control group, with an adjustment for selected individual- and school level characteristics. Schools with at least five responses were included in the regression analysis. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by outcome.

[a]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[b]Effect size (TOT) is calculated by dividing the ITT effect size by the proportion of students who endorsed to recognize two components of the TRAILS curriculum (~25%).

[c]The Self-Awareness score was calculated by adding the score of the six items reflecting this domain. The scale ranged from 6=very difficult to 24=very easy.

[d]The Social-Awareness score was calculated by adding the scores of the five items reflecting this domain. The scale ranged from 5=very difficult to 20=very easy.

[e]The Self-Management score was calculated by adding the scores of the four items reflecting this domain. The scale ranged from 4=very difficult to 16=very easy.
[f]The Relationship Skills score was calculated by adding the scores of the six items reflecting this domain. The scale ranged from 6=very difficult to 24=very easy.
[g]The Responsible-Decision Making score was calculated by adding the scores of the five items reflecting this domain. The scale ranged from 5=very difficult to 20=very easy.
[h]The Combined SEL score was calculated by adding the scores of all 26 items reflecting all five domains. The scale ranged from 26=very difficult to 104=very easy.
[i]The Combined Effective Coping Skills score was obtained by adding the scores of the five effective coping skill (Mindfulness, Cognitive Coping, Behavioral Activation, Listening to Music, and Exposure). The scale ranged from 5=I don't know what this skill is to 25=I use this skill often.
[j]The Embarrassment measure reflects the proportion of students who said they were embarrassed if their friends knew they were getting help from a counselor for an emotional problem.

Statistical significance levels are indicated as follows: *** = $P \leq 0.001$; ** = $P \leq 0.01$; * = $P \leq 0.05$.

Appendix Table D. 2: HLM Conditional Regression Model: Distal Outcomes (PHQ-2 Composite Score, GAD-2 Composite Score)

Distal Outcomes	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size (ITT)[a]
Measure of Depression Symptoms							
PHQ-2 Composite Score[c]	2827	2.10	2.12	-.03	.824	.12	-.01
Measure of Anxiety Symptoms							
GAD-2 Composite Score[d]	2850	2.48	2.72	-.23	.076	.13	-.11
Sample Size							
Schools	25	14	11				

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Students in grade 6-12 were eligible to respond to the PHQ-2 and GAD-2 measures. Impact were estimated by comparing outcomes for the group assigned to TRAILS SEL curriculum with corresponding outcomes for the business as usual control group, with an adjustment for selected individual and school level characteristics. Schools with at least five responses were included in the regression analysis. Rounding may cause slight discrepancies in sums and differences.

[a]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.
[b]Effect size (TOT) is calculated by dividing the ITT effect size by the proportion of students who endorsed to recognize two components of the TRAILS curriculum (~34%)[16].
[c]The PHQ-2 score was obtained by adding the scores for the two question of this depression screener. The score ranged from 0 to 6.
[d]The GAD-2 score was obtained by adding the scores for the two question of this anxiety screener. The score ranged from 0 to 6.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

[16] Only students in grades 6-12 were presented with the PHQ-2 and GAD-2 questions. Thus, this proportion reflects the number of students who recognized two TRAILS components of this adjusted sample and therefore differs from that presented in Table A.1.

Appendix Table D. 3: HLM Unconditional Regression Model Limited to Subgroups (Outcome: Combined SEL Measure)

Outcome: Combined SEL Measure[a]	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size (ITT)[b]
Gender								
Boy	2084	73.01	72.68	.33	.755		1.07	.03
Girl	2006	71.22	69.67	1.55	.116		.99	.14
Race/Ethnicity								
Black	834	72.79	70.93	1.87	.022	*	.82	.17
White	1617	72.07	70.95	1.12	.382		1.29	.09
I prefer not to answer	610	70.73	69.28	1.45	.178		1.08	.11
Vulnerable Groups								
Students who experience homelessness	571	68.43	67.43	1.00	.452		1.32	.08
Students who speak exclusively another language than English at home	177	72.98	70.25	2.73	.344		2.88	.19
Intersectional Groups								
White and Boy	763	73.09	72.85	.25	.853		1.34	.02
White and Girl	755	71.73	69.62	2.11	.121		1.36	.18
Black and Boy	393	73.23	72.39	.84	.531		1.34	.08
Black and Girl	399	72.72	69.86	2.86	.010	**	1.10	.28
Sample Size								
Schools	27[c]	13	14					

NOTES. The program group received TRAILS SEL curriculum and the control group continued with business as usual. Impact for subgroups was estimated by comparing the outcome (Combined SEL Measure) for the group assigned to TRAILS SEL curriculum with the corresponding outcome for the business as usual control group. No covariates were included in the subgroup regression analyses. Schools with at least 20 responses were included in the regression analysis to avoid large variation in group numbers for subgroup regression analyses. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by subgroup analysis.

[a]The Combined SEL score was calculated by adding the scores of all 26 items reflecting all five domains. The scale ranged from 26=very difficult to 104=very easy.

[b]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]Number of schools varied. Two subgroup analyses included 24 schools (White and Girl, Students who speak...) and one 25 schools (White and Boy). All other subgroup analyses included 26 or 27 schools.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Appendix Table D. 4: HLM Unconditional Regression Model Restricted to Subgroups (Outcome: Combined Effective Coping Skills Measure)

Outcome: Combined Effective Coping Skills Measure[a]	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size (ITT)[b]
Gender							
Boy	2242	16.75	16.77	-.02	.95	.26	.00
Girl	2163	17.15	16.92	.23	.318	.23	.07
Race/Ethnicity							
Black	923	17.25	17.26	-.01	.978	.28	.00
White	1697	16.89	16.58	.31	.058	.16	.10
I prefer not to answer	674	16.38	15.97	.41	.190	.31	.10
Vulnerable Groups							
Students who experience homelessness	642	16.55	16.59	-.04	.932	.45	-.01
Students who speak exclusively another language than English at home	211	16.95	15.99	.96	.224	.79	.23
Intersectional Groups							
White and Boy	802	16.88	16.46	.42	.092	.25	.13
White and Girl	784	16.99	16.75	.25	.271	.22	.08
Black and Boy	423	17.11	17.28	-.17	.682	.42	-.05
Black and Girl	447	17.60	17.33	.27	.374	.30	.09
Sample Size							
Schools	27[c]	13	14				

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual (BAU). Impact for subgroups was estimated by comparing the outcome (Combined Effective Coping Skills Measure) for the group assigned to TRAILS SEL curriculum with the corresponding outcome for the BAU control group. No covariates were included in the subgroup regression analyses. Schools with at least 20 responses were included in the regression analysis to avoid large variation in group numbers for subgroup regression analyses. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by subgroup analysis.

[a]The Combined Effective Coping Skills score was obtained by adding the scores of the five effective coping skill (Mindfulness, Cognitive Coping, Behavioral Activation, Listening to Music, and Exposure). The scale ranged from 5=I don't know what this skill is to 25=I use this skill often.

[b]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]Number of schools varied. Number of schools varied. One subgroup analysis included 24 schools (White and Girl) and one 25 schools (White and Boy). All other subgroup analyses included 26 or 27 schools.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Appendix Table D. 5: HLM Unconditional Regression Model Restricted to Subgroups (Outcome: Perceived Embarrassment)

Outcome: Perceived Embarrassment[a]	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size (ITT)[b]
Gender							
Boy	2275	0.23	0.23	-.01	.821	.03	-.01
Girl	2173	0.26	0.27	-.01	.865	.03	-.01
Race/Ethnicity							
Black	925	0.15	0.17	-.02	.588	.03	-.05
White	1708	0.28	0.27	.01	.846	.04	.02
I prefer not to answer	686	0.30	0.29	.01	.852	.05	.02
Vulnerable Groups							
Students who experience homelessness	656	0.32	0.34	-.02	.529	.04	-.05
Students who speak exclusively another language than English at home	220	0.30	0.35	-.06	.574	.10	-.12
Intersectional Groups							
White and Boy	807	0.25	0.24	.01	.752	.03	.02
White and Girl	787	0.27	0.29	-.02	.669	.05	-.05
Black and Boy	429	0.11	0.14	-.03	.482	.04	-.08
Black and Girl	446	0.18	0.19	-.01	.815	.05	-.03
Sample Size							
Schools	26[c]	12	14				

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Impact for subgroups was estimated by comparing the outcome (Embarrassment) for the group assigned to TRAILS SEL curriculum with the corresponding outcome for the business as usual control group. No covariates were included in the subgroup regression analyses. Schools with at least 20 responses were included in the regression analysis to avoid large variation in group numbers for subgroup regression analyses. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by subgroup analysis.

[a]The Embarrassment measure reflects the proportion of students who said they were embarrassed if their friends knew they were getting help from a counselor for an emotional problem.

[b]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]Number of schools varied. Two subgroup analyses included 24 schools (White and Girl, White and Boy). All other subgroup analyses included 25 or 26 schools.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Appendix Table D. 6: HLM Unconditional Regression Model Restricted to Subgroups (Outcome: PHQ-2 Composite Score)

Outcome: PHQ-2 Composite Score[a]	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value		Standard Error	Effect Size (ITT)[b]
Gender								
Boy	1259	1.52	1.65	-0.13	.418		0.17	-.08
Girl	1187	2.35	2.42	-0.08	.491		0.12	-.05
Race/Ethnicity								
Black	548	1.98	2.15	-0.17	.273		0.16	-.09
White	982	1.75	1.98	-0.23	.345		0.24	-.12
I prefer not to answer	257	2.08	2.44	-0.36	.15		0.25	-.17
Vulnerable Groups								
Students who experience homelessness	356	2.65	2.81	-0.16	.496		0.23	-.08
Students who speak exclusively another language than English at home	121	2.61	1.86	0.75	.02	*	0.32	.47
Intersectional Groups								
White and Boy	463	1.28	1.47	-0.19	.334		0.20	-.11
White and Girl	439	2.15	2.35	-0.21	.241		0.18	-.11
Black and Boy	273	1.59	1.82	-0.23	.331		0.24	-.13
Black and Girl	248	2.42	2.37	0.05	.844		0.24	.02
Sample Size								
Schools	17[c]	7	10					

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Impact for subgroups was estimated by comparing the outcome (PHQ-2 Score) for the group assigned to TRAILS SEL curriculum with the corresponding outcome for the business as usual control group. No covariates were included in the subgroup regression analyses. Schools with at least 20 responses were included in the regression analysis to avoid large variation in group numbers for subgroup regression analyses. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by subgroup analysis.

[a]The PHQ-2 score was obtained by adding the scores for the two questions of this depression screener. The score ranged from 0 to 6.

[b]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]Number of schools varied by subgroup analysis. One subgroup analysis included 13 schools (White and Girl) and one 15 schools (Black and Girl). All other subgroup analyses included 16 or 17 schools.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Appendix Table D. 7: HLM Unconditional Regression Model Restricted to Subgroups (Outcome: GAD-2 Composite Score)

Outcome: GAD-2 Composite Score[a]	Number of Students	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error	Effect Size (ITT)[b]
Gender							
Boy	1258	1.75	1.95	-0.20	.126	0.13	-.10
Girl	1185	3.09	3.29	-0.20	.208	0.16	-.10
Race/Ethnicity							
Black	554	2.27	2.53	-0.26	.148	0.18	-.12
White	980	2.45	2.83	-0.39	.178	0.29	-.17
I prefer not to answer	256	2.48	2.76	-0.28	.317	0.28	-.12
Vulnerable Groups							
Students who experience homelessness	361	3.44	3.47	-0.03	.911	0.28	-.02
Students who speak exclusively another language than English at home	124	2.46	2.30	0.16	.666	0.36	.08
Intersectional Groups							
White and Boy	462	1.71	2.05	-0.34	.194	0.26	-.16
White and Girl	438	3.30	3.54	-0.24	.229	0.20	-.12
Black and Boy	278	1.67	1.92	-0.25	.444	0.32	-.12
Black and Girl	246	2.89	2.98	-0.09	.734	0.26	-.04
Sample Size							
Schools	17[c]	7	10				

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Impact for subgroups was estimated by comparing the outcome (GAD-2 Score) for the group assigned to TRAILS SEL curriculum with the corresponding outcome for the business as usual control group. No covariates were included in the subgroup regression analyses. Schools with at least 20 responses were included in the regression analysis to avoid large variation in group numbers for subgroup regression analyses. Rounding may cause slight discrepancies in sums and differences. Number of schools in program and control group varies slightly by subgroup analysis.

[a]The GAD-2 score was obtained by adding the scores for the two questions of this anxiety screener. The score ranged from 0 to 6.

[b]Effect size (ITT) is calculated by dividing the impact of the program (the difference between the means for the program group and the control group) by the standard deviation for the control group.

[c]Number of schools varied by subgroup analysis. One subgroup analysis included 13 schools (White and Girl) and one 15 schools (Language other than English). All other subgroup analyses included 16 or 17 schools.

Statistical significance levels are indicated as follows: *** = P ≤ 0.001; ** = P ≤ 0.01; * = P ≤ 0.05.

Appendix Table D. 8: Unconditional Regression Analysis Model (Outcome: Student Educational Performance)

Student Educational Performance[a]	Program Group Mean	Control Group Mean	Difference (Impact)	P-Value	Standard Error
Measure of Educational Achievement					
Student Performance	0.19	0.25	-0.06	.176	0.05
Sample Size					
Schools	32				

NOTES: The program group received TRAILS SEL curriculum and the control group continued with business as usual. Schools with at least five responses were included in the regression analysis. Rounding may cause slight discrepancies in sums and differences

[a]Changes in student performance were measured by using the publicly available school data from the MI School Data website, the State of Michigan's official public portal for education data. The means represent the proportion of students meeting state academic standards (scoring "proficient" or "advanced") on state tests.

Appendix E: Staff Outcomes

Practices, Intervention, and Training

Appendix Table E. 1: Comparison of Staff Reports on Availability of School-Based Protocols and Supports at the Beginning of the School Year 22-23

	Program Group	Control Group	Difference
Protocols (%)			
Screening individual students	48.06	42.29	5.77
Referring students to school-based MH services	59.79	64.68	-4.89
Referring students to community-based MH services	53.90	58.00	-4.10
Interventions (%)			
Universal, school wide screening for student mental health	27.17	25.95	1.22
Tier 1 universal mental health interventions (e.g. SEL)	49.09	60.87	-11.78
Tier 2 targeted mental health interventions for students with mild to moderate need	52.73	60.20	-7.48
Tier 3 intense interventions for students with severe need	52.36	51.54	0.83
Training and Support (%)			
Voluntary training in student mental health or social and emotional learning	36.26	37.66	-1.40
Mandatory training in student mental health or social and emotional learning	43.17	41.45	1.72
Access to consultants or coaches to support staff in implementing school-based mental health interventions	46.52	41.58	4.94
Sample Size			
Staff[a]	283	402	

NOTE: Rounding may cause slight discrepancies in sums and differences. The table reflects the proportion of staff who reported that that any of the listed protocols and supports were available, but not implemented, available, but need support to implement it well, and available and implemented well.

[a] The denominator for each question varies. For all variables in the table, data are available for at least 95% of the sample.

Appendix Table E. 2: Comparison of Staff Reports on Availability of School-Based Protocols and Supports at the End of the School Year 22-23

	Program Group	Control Group	Difference
Protocols (%)			
Screening individual students	54.18	46.82	7.36
Referring students to school-based MH services	67.34	62.43	4.91
Referring students to community-based MH services	61.45	52.31	9.13
Interventions (%)			
Universal, school wide screening for student mental health	36.18	26.76	9.41
Tier 1 universal mental health interventions (e.g. SEL)	74.80	57.14	17.65
Tier 2 targeted mental health interventions for students with mild to moderate need	74.29	60.06	14.23
Tier 3 intense interventions for students with severe need	61.38	52.77	8.61
Training and Support (%)			
Voluntary training in student mental health or social and emotional learning	42.45	35.96	6.48
Mandatory training in student mental health or social and emotional learning	47.76	32.16	15.59
Access to consultants or coaches to support staff in implementing school-based mental health interventions	49.59	38.35	11.24
Sample Size			
Staff[a]	251	346	

NOTES: Rounding may cause slight discrepancies in sums and differences. The table reflects the proportion of staff who reported that that any of the listed protocols and supports were available, but not implemented, available, but need support to implement it well, and available and implemented well.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 96% of the sample.

Appendix Table E. 3: Comparison Staff's Participation in Professional Development, Knowledge About MH Resources at the Beginning of the School Year 22-23

	Program Group Group Mean	Control Group Group Mean	Difference	P-Value
Professional Development[a] (%)				
Student mental health	52.26	66.58	-14.32	.000
Social emotional learning	66.20	80.15	-13.95	.000
Trauma or post-traumatic stress disorder	49.30	54.95	-5.65	.144
Knowledge[b] (%)				
I know where to turn for support when I am concerned about a student's emotional well-being.	92.66	92.04	0.62	.645
I feel confident that I can connect students to services that will be beneficial, when I am concerned about their emotional well-being.	83.57	74.94	8.63	.007
Sample Size				
Staff[c]	287	404		

NOTES: Rounding may cause slight discrepancies in sums and differences We report the standard deviation for mean outcomes in parentheses. Statistical significance was assessed using independent t-test analyses and are reported as p-value.

[a]Staff reported how often they have participated in professional development activities focusing on student mental health, social emotional learning, or trauma in the last 12 months. We report on staff who have participated at least once in PDs in the last 12 months.
[b]Staff rated statement on knowledge of resources for student on a 4-point scale ranging from 1=strongly agree to 4=strongly disagree. We report on staff who agree or strongly agree on the respective statements.
[c]The denominator for each question varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table E. 4: Comparison Staff's Participation in Professional Development, Knowledge About MH Resources, Burnout, and Openness to New Programs at the End of the School Year 22-23

	Program Group Group Mean	Control Group Group Mean	Difference	P-Value
Professional Development[a] (%)				
Student mental health	62.70	61.36	1.33	.740
Social emotional learning	82.87	72.65	10.22	.003
Trauma or post-traumatic stress disorder	49.60	45.74	3.86	.349
Knowledge[b] (%)				
I know where to turn for support when I am concerned about a student's emotional well-being.	94.86	95.70	-0.84	.749
I feel confident that I can connect students to services that will be beneficial, when I am concerned about their emotional well-being.	84.98	79.94	5.04	.112
Sample Size				
Staff[c]	253	352		

NOTES: Rounding may cause slight discrepancies in sums and differences. We report the standard deviation for mean outcomes in parentheses. Statistical significance was assessed using independent t-test analyses and are reported as p-value.

[a]Staff reported how often they have participated in professional development activities focusing on student mental health, social emotional learning, or trauma in the last 12 months. We report on staff who have participated at least once in PDs in the last 12 months.
[b]Staff rated statement on knowledge of resources for student on a 4-point scale ranging from 1=strongly agree to 4=strongly disagree. We report on staff who agree or strongly agree on the respective statements.
[c] The denominator for each question varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table E. 5: Comparison Level of Burnout, and Openness to Change at the End of the School Year 22-23

	Program Group Group Mean	Control Group Group Mean	Difference	P-Value
Burnout[a]				
Level of exhaustion	2.75 (0.40)	2.64 (0.52)	0.11	.006
Openness to Change[b]				
Openness to Change	3.67 (0.85)	3.60 (0.95)	0.07	.399
Sample Size				
Staff[c]	237	353		

NOTES: Rounding may cause slight discrepancies in sums and differences. We report the standard deviation for mean outcomes in parentheses. Statistical significance was assessed using independent t-test analyses and are reported as p-value.

[a]Staff rated 8 items relating to "Exhaustion" from the Oldenburg Inventory from 1=strongly disagree to 4=strongly agree. The higher the score the greater the level of overall "exhaustion". More information about this scale can be found here: https://www.goodmedicine.org.uk/sites/default/files/assessment%2C%20burnout%2C%20olbi.pdf.[18]
[c]Staff 9 items relating to "Faculty openness to change" from the "Change Orientation Scale" from 1=strongly disagree to 6=strongly agree. The higher the score the greater the "openness to change". More information about this scale can be found here: https://www.waynekhoy.com/change-scale/[19]
[d]The denominator for each question varies. For all variables in the table, data are available for at least 93% of the sample.

Appendix F: Supplemental Findings

Supplemental Findings: Student Outcomes

Appendix F presents supplemental findings not referenced in the body of the report. We include additional findings on student outcomes related to individual SEL and coping skills items, depression and anxiety, student outreach, self-management skills as they relate to school work, bullying, and other measures.

Appendix Table F. 1: % of Students Reporting to Use the Following Effective Coping Skills Sometimes or Often at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
I Use this Coping Skill Sometimes or Often:			
Mindfulness	38.4	38.1	0.3
Listening to calming or uplifting music	69.4	68.4	0.9
Doing something active, like sports, dancing, walking, running, basketball, etc. (Behavioral Activation)	61.9	60.7	1.2
Questioning your automatic or negative thoughts (Cognitive Coping)	31.8	32.8	-1.0
Overcoming avoidance by learning to face your fears (Exposure)	44.0	43.6	0.4
Sample Size			
Students[a]	2192	3123	

NOTES: Students rated the question "When you are stressed or worried, how often do you use the following skills to help you feel better?" on a 5-point rating skill from 1=I don't know what this is to 5=Often.
[a]The denominator for each question varies. For all variables in the table, data are available for at least 98% of the sample.

Appendix Table F. 2: % of Students Reporting That they Don't Know The Following Coping Skills at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
"I don't know this" Coping Skill:			
Mindfulness	7.8	8.8	-1.1
Listening to calming or uplifting music	2.4	1.6	0.8
Doing something active, like sports, dancing, walking, running, basketball, etc. (Behavioral Activation)	3.4	3.3	0.1
Questioning your automatic or negative thoughts (Cognitive Coping)	15.3	18.3	-3.0
Overcoming avoidance by learning to face your fears (Exposure)	7.2	7.4	-0.2
Sample Size			
Students[a]	2192	3126	

NOTES: Students rated the question "When you are stressed or worried, how often do you use the following skills to help you feel better?" on a 5-point rating skill from 1=I don't know what this is to 5=Often.
[a]The denominator for each question varies. For all variables in the table, data are available for at least 98% of the sample.

Appendix Table F. 3: % of Students Reporting That SEL Skills are Easy/Very Easy to Do at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
Self Awareness			
Know when my feelings are making it hard for me to focus.	68.5	66.1	2.4
Know the emotions I feel.	69.0	67.2	1.8
Know ways to make myself feel better when I'm sad	58.1	52.6	5.5
Noticing what my body does when I am nervous	73.0	72.4	0.6
Know when my mood affects how I treat others.	72.0	69.4	2.6
Know ways I calm myself down.	58.5	55.9	2.6
Self Management			
Get through something even when I feel frustrated.	41.6	37.8	3.8
Be patient even when I am really excited.	54.1	53.8	0.3
Stay calm when I feel stressed.	43.1	40.4	2.7
Work on things even when I don't like them.	48.8	50.8	-2.0
Social Awareness			
Learn from people with different opinions than me	68.8	68.5	0.3
Know what people may be feeling by the look on their face.	77.8	77.0	0.8
Know when someone needs help.	76.5	75.3	1.2
Know how my actions impact my classmates.	61.5	60.1	1.5
Know how to get help when I'm having trouble with a classmate.	73.3	72.6	0.7
Relationship			
Respect a classmate's opinions during a disagreement	66.4	66.3	0.0
Get along with my classmates.	75.8	73.6	2.3
Share what I am feeling with others.	37.0	32.6	4.3
Talk to an adult when I have problems at school.	45.8	44.8	1.1
Be welcoming to someone I don't usually eat lunch with.	65.6	64.2	1.4
Get along with my teachers.	80.0	81.3	-1.2
Responsible Decision-Making			
Think about what might happen before making a decision.	58.0	55.3	2.7
Know what is right or wrong	79.5	78.5	1.0
Think of different ways to solve a problem	69.6	68.1	1.5
Say "no" to a friend who wants to break the rules.	73.0	69.5	3.5
Help to make my school a better place.	61.0	60.5	0.5
Sample Size			
Students[a]	2140	3078	

NOTES: Students rated the list of items on a 4-point rating skill from 1=very difficult to 4=very easy.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 96% of the sample.

Appendix Table F. 4: % of Students Scoring 3+ on the Depression (PHQ-2) and Anxiety (GAD-2) Screener at the Beginning of the School Year 22-23

	Program Group %	Control Group %	Difference
PHQ-2			
Score of 3 or higher	39.0	40.3	-1.3
GAD-2			
Score of 3 or higher	44.7	44.5	0.2
Sample Size			
Students[a]	1,996	1,581	

NOTES: Only students in grades 6-12 were eligible to respond to the PHQ-2 and GAD-2 screener.

[a]The denominator for each variable varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table F. 5: % of Students Scoring 3+ on the Depression (PHQ-2) and Anxiety (GAD-2) Screener, and % of Students Reporting about Trauma at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
PHQ-2			
Score of 3 or higher	36.8	38.8	-2.0
GAD-2			
Score of 3 or higher	47.1	50.5	-3.4
Trauma[a]			
Yes	24.5	29.0	-4.5
YRBS[b]			
Experienced persistent feelings of sadness or hopelessness	38.8	42.1	-3.3
Sample Size			
Students[c]	1,490	1,366	

NOTES: Only students in grades 6-12 were eligible to respond to the PHQ-2, the GAD-2 screener, and the question about trauma.

[a]Students responded to the following questions with yes, no, I don't know: "Sometimes people have violent or very scary or upsetting things happen to them. This could be something that happened to you or something you saw. It can include being badly hurt, someone doing something harmful to you or someone else, or a serious accident or serious illness. Has anything like this ever happened to you?
[b]The survey at the end of the school year included the following standalone item from the Youth Risk Behavioral Survey (YRBS) "During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing some usual activities?" Students answered this item with yes or no.
[c]The denominator for each variable varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table F. 6: % of Students Reporting to Have Talked to the Following People about a Personal or Emotional Problem at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
% of Students Who Talked to the Following People About a Personal or Emotional Problem			
Teacher	15.5	19.3	-3.7
School Social Worker, School Nurse, or School Counselor	15.1	17.8	-2.7
Support group at school	6.5	4.7	1.8
Other people	3.3	2.7	0.6
None of these	37.1	39.8	-2.6
Sample Size			
Students	1590	1459	

NOTES: Rounding may cause slight discrepancies in sums and differences. Students in grades 6-12 responded to the following questions: "Which of the following people have you talked to in the past year about a personal or emotional problem?"

Appendix Table F. 7: % of Students Reporting That Self-Management Skills Related to School Work are Easy/Very Easy to Do at the End of the School Year 22-23

	Program Group %	Control Group %	Difference
% of Students Saying that it is Easy/Very Easy to:			
Do my schoolwork even when I don't not feel like it	44.2	42.1	2.1
Work on assignments even when they are hard	47.4	45.7	1.7
Plan ahead so I can turn a project in on time	52.4	50.8	1.6
Finish my schoolwork without reminders	52.6	52.1	0.5
Be prepared for tests	57.3	52.4	4.8
Sample Size			
Students[a]	1427	1340	

NOTES: Rounding may cause slight discrepancies in sums and differences. Students in grades 6-12 rated the list of items on a 4-point scale from 1=very difficult to 4=very easy to do.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table F. 8: % of Students Reporting to Agree Often or Always with Statements Related to Belonging and Safety at the End of the School Year 22-23

	Program Group	Control Group	Difference
% of Students Often or Always Agreeing with the Following Statements			
I get along with other students	64.07	61.61	2.46
Students treat each other well	38.24	35.22	3.01
There is an adult at school who will help me if I need to	56.31	62.53	-6.21
I feel safe at school	49.29	48.95	0.35
My teachers really care about me	53.04	60.42	-7.38
Sample Size			
Students[a]	1567	2141	

NOTES: Rounding may cause slight discrepancies in sums and differences. Students rated the list of items on a 4-point scale from 1=Never to 4=Always.

[a]The denominator for each question varies. For all variables in the table, data are available for at least 99% of the sample.

Appendix Table F. 9: % of 6-12th Grade Students Saying They Have Been Teased or Called Names at the End of the 22-23 School Year

	Program Group	Control Group	Difference
Students Who Have Been Teased or Called Names in the Last 12 Months Because of their: ____ ()			
Race	13.12	23.09	-9.98
Ethnic background or national origin	7.58	8.16	-0.58
Sexual orientation	11.66	15.09	-3.43
Gender identity	6.34	6.85	-0.51
Religion	8.31	8.39	-0.08
Disability status	8.67	10.55	-1.87
Physical appearance	36.30	40.65	-4.35
Students Who Have Not Been Teased or Called Names in the Last 12 Months ()	48.91	42.57	6.34
Sample Size			
Students	1,372	1,299	

NOTES: Rounding may cause slight discrepancies in sums and differences. Only students in grades 6-12 were eligible to respond to this question.

Appendix Table F. 10: Average Scores of Mean Outcomes at the Beginning of the School Year 22-23

	Program Group Group Mean	Control Group Group Mean	Difference	P-Value
Social Emotional Learning Skills[a]				
Self-Awareness (6 items)	2.79 (0.58)	2.75 (0.61)	0.04	.014
Self-Management (4 items)	2.40 (0.65)	2.36 (0.65)	0.05	.004
Social Awareness (5 items)	2.84 (0.54)	2.85 (0.54)	-0.01	.528
Relationship Skills (6 items)	2.66 (0.56)	2.67 (0.55)	-0.02	.221
Responsible Decision-Making (5 items)	2.80 (0.59)	2.81 (0.58)	-0.01	.592
Overall SEL (26 items)	2.72 (0.45)	2.71 (0.45)	0.01	.500
Use of Effective Coping Skills[b]				
Non-behavioral skills (Mindfulness, Cognitive Coping, Listening to Music)	3.26 (0.81)	3.21 (0.82)	0.05	.030
Non-behavioral skills (Mindfulness, Cognitive Coping)	2.89 (0.94)	2.84 (0.96)	0.05	.040
Behavioral skills (Behavioral Activation, Exposure)	3.42 (0.94)	3.42 (0.93)	0.00	.990
Overall Composite Coping Skills (All 5 items)	3.33 (0.70)	3.30 (0.70)	0.03	.120
Perceived Embarrassment (%)[c]				
Students who report embarrassment if their friends knew they were getting help from a counselor	40.83	38.54	2.29	.061
Sample Size				
Students[d]	3010	3356		

NOTES: We report the standard deviation for mean outcomes in parentheses. Statistical significance was assessed with t-test analyses and are reported as p-value. Rounding may cause slight discrepancies in sums and differences.

[a]Students rated a total of 26 items focusing on social-emotional learning competencies on a scale ranging from 1=very difficult to 4=very easy. We computed an average score for each competency reflecting a list of respective competencies. The SEL items were derived from from the Washoa County School District Social and Emotional Competency Assessments (WCSD-SECA). More information can be found here: https://www.washoeschools.net/Page/10932. Higher average ratings (closer to a 4.0) mean that students rated those types of skills as easier to do. Lower average ratings (closer to a 1.0) mean that students rated those types of skills as harder to do.

[b]Students rated how often they used 5 coping skills on a scale ranging from 1=I don't know what this is to 5=Often. We computed an average score for non-behavioral and behavioral skills separately, as well as generated an average score for all items. Higher average ratings (closer to 5.0) mean that students rated those coping skills to use more often.

[c]Students rated the following question "How embarrassed would you be if your friends knew you were getting professional?" on a scale ranging from 1=Very embarrassed to 5=Not at all embarrassed. We classified students who said "very embarrassed, embarrassed, or somewhat embarrassed" as students who reported embarrassment.

[d]The denominator for each question varies. For all variables in the table, data are available for at least 84% of the sample in both groups.

Appendix Table F. 11: Average Scores of Mean Outcomes at the End of the School Year 22-23

	Program Group Group Mean	Control Group Group Mean	Difference	*p*-Value
Social Emotional Learning Skills[a]				
Self-Awareness (6 items)	2.83 (0.58)	2.78 (0.59)	0.05	.006
Self-Management (4 items)	2.43 (0.65)	2.40 (0.66)	0.03	.151
Social Awareness (5 items)	2.88 (0.54)	2.87 (0.55)	0.01	.549
Relationship Skills (6 items)	2.68 (0.59)	2.66 (0.58)	0.01	.405
Responsible Decision-Making (5 items)	2.83 (0.60)	2.80 (0.60)	0.02	.161
Combined SEL Measure (26 items)	2.75 (0.46)	2.72 (0.47)	0.03	.067
Use of Effective Coping Skills[b]				
Non-behavioral skills (Mindfulness, Cognitive Coping, Listening to Music)	3.29 (0.82)	3.27 (0.80)	0.02	.340
Non-behavioral skills (Mindfulness, Cognitive Coping)	2.96 (0.92)	2.92 (0.93)	0.04	.170
Behavioral skills (Behavioral Activation, Exposure)	3.48 (0.93)	3.46 (0.92)	0.02	.310
Combined Effective Coping Skills Measure (All 5 items)	3.37 (0.72)	3.34 (0.70)	0.03	.210
Perceived Embarrassment (%)[c]				
Students who report embarrassment if their friends knew they were getting help from a counselor	25.50	25.10	0.00	.741
Sample Size				
Students[d]	2161	3,112		

NOTES: We report the standard deviation for mean outcomes in parentheses. Statistical significance was assessed with t-test analyses and are reported as p-value. Rounding may cause slight discrepancies in sums and differences.

[a]Students rated a total of 26 items focusing on social-emotional learning competencies on a scale ranging from 1=very difficult to 4=very easy. We computed an average score for each competency reflecting a list of respective competencies. The SEL items were derived from from the Washoa County School District Social and Emotional Competency Assessments (WCSD-SECA). More information can be found here: https://www.washoeschools.net/Page/10932. Higher average ratings (closer to a 4.0) mean that students rated those types of skills as easier to do. Lower average ratings (closer to a 1.0) mean that students rated those types of skills as harder to do.

[b]Students rated how often they used 5 coping skills on a scale ranging from 1=I don't know what this is to 5=Often. We computed an average score for non-behavioral and behavioral skills separately, as well as generated an average score for all items. Higher average ratings (closer to 5.0) mean that students rated those coping skills to use more often.

[c]Students responded to one question "Would you be embarrassed if your friends knew you were getting help from a counselor for an emotional problem?" with yes or no.

[d]The denominator for each question varies. For all variables in the table for the program group, data are available for at least 87% of the sample; for the program group, data are available for at least 89% of the sample.

Appendix Table F. 12: % of Students With a Traumatic Experience who Showed Symptoms of Depression or Anxiety at the End of the School Year 22-23

	Program Group	Control Group	Difference
% of students with Traumatic Experience:			
PHQ-2			
Score of 3 or higher	53.6	55.9	-2.3
GAD-2			
Score of 3 or higher	64.5	68.1	-3.6
Sample Size			
Students	346	386	

NOTES: This table represents the number of students who showed symptoms of depression and responded to the outreach question. Rounding may cause slight discrepancies in sums and differences. Only students in grades 6-12 were eligible to respond to these questions.

Appendix Table F. 13: % of Students With Symptoms of Depression who Reached Out to Certain People to Talk About a Personal Problem End of the School Year 22-23

	Program Group	Control Group	Difference
% of Students with Symptoms of Depression[a] who:			
talked to a teacher	17.6	20.4	-2.8
talked to a social worker	20.0	24.6	-4.6
talked to a support group	7.8	6.3	1.5
talked to none	33.2	32.6	0.6
Sample Size			
Students	540	521	

NOTES: This table represents the number of students who showed symptoms of depression and responded to the outreach question. Rounding may cause slight discrepancies in sums and differences. Only students in grades 6-12 were eligible to respond to these questions.

[a]Symptoms of Depression=Score of 3 or higher on the PHQ-2.

Appendix Table F. 14: % of Students With Symptoms of Anxiety who Reached Out to Certain People to Talk About a Personal Problem End of the School Year 22-23

	Program Group	Control Group	Difference
% of Students with Symptoms of Anxiety[a] who:			
talked to a teacher	18.4	19.4	-0.9
talked to a social worker	19.6	23.6	-4.0
talked to a support group	7.6	6.7	0.9
talked to none	31.4	31.9	-0.5
Sample Size			
Students	694	687	

NOTES: This table represents the number of students who showed symptoms of depression and responded to the outreach question. Rounding may cause slight discrepancies in sums and differences. Only students in grades 6-12 were eligible to respond to these questions.

[a]Symptoms of Anxiety=Score of 3 or higher on the GAD-2.

Appendix Table F. 15: Aggregated Data from MI School Data (Michigan's Official Source for Pre-, K-12, Postsecondary and Workforce Data) at the End of the 2022-23 School Year

	Program Group	Control Group	Difference	State Average (Michigan)
Student Data (%)				
Student Performance in English Language Arts[a]	23.8 (11.8)	32.5 (15.9)	-8.7	45.0
Student Performance in Mathematics[b]	12.2 (9.8)	18.1 (16.0)	-5.9	34.0
Grade Retention[c]	5.4 (7.0)	2.0 (4.6)	3.4	2.9
Graduation Rate[d]	69.1 (27.0)	64.8 (28.0)	4.3	82.0
Attendance[e]	87.3 (2.6)	88.2 (4.2)	-0.9	91.0
Economically Disadvantaged[g]	78.5 (12.4)	73.4 (19.6)	5.1	53.0
Staff Data (%)				
Qualified Teaching Staff[h]	91.9 (9.3)	90.4 (14.6)	1.5	91.0
Teacher Retention[i]	69.9 (25.1)	70.1 (12.2)	-0.2	79.0
Sample Size				
Number of schools	21	19		

NOTES: This table represents aggregated data derived from the MI School Data website (https://www.mischooldata.org/). The aggregated school data was averaged over the number of schools by treatment status. We show the standard deviation in parentheses. We also include a state average for comparison reasons.

[a]The percent of students meeting state academic standards in English Language Arts (scoring "proficient" or "advanced") on state tests (M-STEP, MI-Access, and SAT; PSAT) where available.

[b]The percent of students meeting state academic standards in English Language Arts (scoring "proficient" or "advanced") on state tests (M-STEP, MI-Access, and SAT; PSAT) where available.

[c]The Retained in Grade report shows the percentages of students who did not move with their class to the next grade level.

[d]The percent represents all students who graduated within a four-year period. This information was only available from schools serving 12th grade students. The percentage encompasses 261 students across four control schools and 538 students across seven treatment schools.

[e]Attendance rates are calculated using the number of days a student is in attendance divided by the possible days attended.

[f]Percentage of students who were suspended or expelled. Suspensions include In-School and Out-of-School Suspensions.

[g] The percent of students who are eligible to receive free or reduced lunch.

[h]The percent of teachers, by content area, who are certified to teach the assignment they are teaching.

[i]The percent of teachers who are retained year over year at the same school.

Endnotes

[1] Washoe County School District (2024, May 16). *Social & Emotional Learning*. https://www.wcsddata.net/data-topics/sel/.

[2] Gordon, Rachel A., Marisa K. Crowder, Ariel M. Aloe, Laura A. Davidson, and Celene E. Domitrovich. "Student Self-Ratings of Social-Emotional Competencies: Dimensional Structure and Outcome Associations of the WCSD-SECA among Hispanic and Non-Hispanic White Boys and Girls in Elementary through High School," *Journal of School Psychology 93* (2022): 41–62, https://doi.org/10.1016/j.jsp.2022.05.002; Crowder, Marisa K, Rachel A Gordon, Randal D Brown, Laura A Davidson, and Celene E Domitrovich. "Linking Social and Emotional Learning Standards to the WCSD Social–Emotional Competency Assessment: A Rasch Approach," *School Psychology* 34, no. 3 (2019): 281–95, https://doi.org/10.1037/spq0000308.

[3] Richardson, Laura P., Carol Rockhill, Joan E. Russo, David C. Grossman, Julie Richards, Carolyn McCarty, Elizabeth McCauley, and Wayne Katon. "Evaluation of the PHQ-2 as a Brief Screen for Detecting Major Depression among Adolescents." *Pediatrics* 125, no. 5 (2010): e1097-1103. https://doi.org/10.1542/peds.2009-2712.

[4] Plummer, F. Manea, L., Trepel., D. and Dean McMillan, "Screening for Anxiety Disorders with the GAD-7 and GAD-2: A Systematic Review and Diagnostic Metaanalysis", *Gen Hosp Pscychiatry* 39: 24-31 (2016), https://doi.org/10.1016/j.genhosppsych.2015.11.005.

[5] What Works Clearinghouse, *What Works Clearinghouse procedures and standards handbook*, version 5.0. (NCEE, 2022), U.S. Department of Education, Institute of Education sciences, https://ies.ed.gov/ncee/wwc/Handbooks.

[6] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. (2nd ed.) (Hillsdale, NJ: Lawrence Erlbaum, 1988).

[7] Condliffe, Barbara, Zhu, P., Doolittle, F., van Dok, M., Power, H., Denison, D. and Anja Kurki. *Study of Training in Multi-Tiered Systems of Support for Behavior: Impact on Elementary School Students' Outcomes* (NCEE 2022-008), U.S. Department of Education, Institute of Education Sciences, http://ies.ed.gov/ncee

[8] Al-Jbouri, E., Andrews, N.C.Z., Peddigrew, E., Fortier, A, and Tracy Weaver. "Building elementary students' social and emotional skills: A randomized control trial to evaluate a teacher-led intervention," *School Mental Health*, no. 15, (2023): 138–150, https://doi.org/10.1007/s12310-022-09538-x

[9] Jones, S. Baily, R., Brush, K. & J. Kahn. *Preparing for effective SEL implementation*, (Harvard Graduate School of Education, EASEL Lab, 2018), https://wallacefoundation.org/sites/default/files/2023-09/Preparing-for-Effective-SEL-Implementation.pdf

[10] Doan, S., Steiner, E.D., and Rakesh Pandey. *Teacher well-being and intentions to leave in 2024. Findings from the 2024 State of the American Teacher Survey.* (Santa Monica, Calif: RAND Corporation, 2024), https://www.rand.org/pubs/research_reports/RRA1108-12.html

[11] Tyner, Adam, How to sell SEL: Parents and the politics of social-emotional learning, (Washington DC, Thomas B. Fordham Institute, 2021), https://sel.fordhaminstitute.org/

[12] Crowder, M.K., Gordon, R.A, Brown, R.D., Davidson, L.A., and Celene E. Domitrovich. "Linking Social and Emotional Learning Standards to the WCSD Social–Emotional Competency Assessment: A Rasch Approach," *School Psychology* 34, no. 3 (2019): 281–95, https://doi.org/10.1037/spq0000308.

[13] Davidson, L.A., Crowder, M.K., Gordon, R.A., Domitrovich, C.E., Brown, R.D., and Benjamin L. Hayes. "A Continuous Improvement Approach to Social and Emotional Competency Measurement," *Journal of Applied Developmental Psychology*, Social-Emotional Assessment to Guide Educational Practice, 55 (2018): 93–106, https://doi.org/10.1016/j.appdev.2017.03.002.

[14] Crowder, Linking Social and Emotional Learning Standards to the WCSD, 281–95

[15] Davidson, A Continuous Improvement Approach to Social and Emotional Competency Measurement, 93–106

[16] Richardson, Evaluation of the PHQ-2, 1097-1103.

[17] Plummer, Screening for Anxiety Disorders with the GAD-7 and GAD-2, 24-31.

[18] Demerouti, Evangelia, and Arnold B. Bakker, "The Oldenburg Burnout Inventory: A Good Alternative to Measure Burnout (and Engagement)," 2007, 25.

[19] Kearney, S., & Smith P. A. A theoretical and empirical analysis of change orientations in schools. In Wayne K. Hoy & Michael DiPaola (Eds.). *Studies in School Improvement*. (Greenwich, CT: Information Age, 2009)

Acknowledgements

The Youth Policy Lab would like to thank our partners at TRAILS for their leadership in efforts to improve student mental health. In particular, we would like to thank the TRAILS team who made this work possible by collaborating with schools and staff and surveying students in schools. This work would not have been possible without the support from our YPL team. We particularly want to thank Milagros Chocce, Yejae Kim, and Liz Reosti for contributing to this report.

Conflict of Interest Statement

The University of Michigan has a financial interest in the TIDES Center, with which the TRAILS program is affiliated.

Support the Youth Policy Lab's effort to use data for good.

University of Michigan Youth Policy Lab
5201 Institute for Social Research
426 Thompson St
Ann Arbor, MI 48104

734-647-8829
@YouthPolicyLab
www.linkedin.com/showcase/youth-policy-lab/
youthpolicylab.umich.edu

Photo by Kelli Tungay on Unsplash

About the Authors

Robin Jacob is a faculty co-director of the Youth Policy Lab and a research professor at the Institute for Social Research at the University of Michigan. Her research focuses on rigorously evaluating interventions designed to improve the educational and life outcomes of youth and their families.

Katja Robinson is a project manager with the Youth Policy Lab. Her work primarily focuses on early childhood education and youth mental health policy.

Youth Policy Lab

The University of Michigan Youth Policy Lab was launched in 2016 with a vision for reducing socio-economic disparities through improvements in education and other social policies affecting youth. By developing evidence-based, policy-relevant research in partnership with local and state agencies, practitioners, and policymakers, Dr. Robin Jacob and Dr. Brian Jacob sought to build upon their exemplary careers in social science research by taking research findings out of academic journals and putting them in the hands of decision-makers. With this aim in mind, they have spent the past seven years bringing the resources and expertise of one of the nation's leading public research universities to bear on some of Michigan's most pressing social challenges.

The Youth Policy Lab envisions a world where partner-driven research drives positive social change. Our mission is to inform public policy decisions by analyzing data and evaluating programs to help our partners answer their most pressing questions.