

The Impact of the Second Step Program on Elementary Students' Social and Emotional Skills, Exclusionary Discipline, and Attendance

Nicholas A Gage, PhD

April 2025

© 2025 WestEd. All rights reserved.



Gage, N.A. (2025). *The impact of the Second Step program on elementary students' social and emotional skills, exclusionary discipline, and attendance*. WestEd.

WestEd is a nonpartisan, nonprofit organization that aims to improve the lives of children and adults at all ages of learning and development. We do this by addressing challenges in education and human development, increasing opportunity, and helping build communities where all can thrive. WestEd staff conduct and apply research, provide technical assistance, and support professional learning. We work with early learning educators, classroom teachers, local and state leaders, and policymakers at all levels of government.

For more information, visit [WestEd.org](https://www.wested.org). For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at [WestEd.org/subscribe](https://www.wested.org/subscribe).

Executive Summary

This report describes and evaluation of the impact of the Second Step digital social-emotional learning (SEL) program on elementary student outcomes in a large, diverse North Carolina school district. Using data from over 23,000 students and rigorous statistical methods including propensity score weighting and multilevel modeling, the study finds that moderate-to-high-fidelity implementation of Second Step (completing over 60-80% of lessons) is associated with small but significant improvements in students' self-management skills and a meaningful reduction in out-of-school suspensions. No significant effects were found for other SEL domains, office discipline referrals, or in-school suspensions, though attendance showed a marginal improvement at high fidelity levels. The study highlights the importance of program fidelity and fostering a strong sense of belonging in schools to support positive behavioral and attendance outcomes. Limitations include lack of student demographic data and reliance on lesson completion as the sole fidelity measure. Overall, the findings suggest that well-implemented universal SEL programs like Second Step can modestly enhance key social-emotional competencies and reduce exclusionary discipline in elementary settings.

This study meets the What Works Clearinghouse evidence standards with reservations and the Collaborative for Academic, Social, and Emotional Learning (CASEL) Guide to Effective Social and Emotional Learning Programs design criteria by including a baseline equivalence comparison group and finding a significant effect on an outcome in the behavioral student outcome domain.

Table of Contents

The Impact of the Second Steps Program on Elementary Students' Social and Emotional Skills, Exclusionary Discipline, and Attendance	1
Method	2
Sample	2
Measures	2
School Environment and Social and Emotional Skills	2
Discipline	3
Attendance	4
Independent Variable	4
Data Analysis	5
Propensity Score Weighting and Covariate Balance	5
Multilevel Models	6
Results	7
School Environment and Social and Emotional Skills	8
Primary Treatment Effect Models	8
Different Levels of Fidelity	8
Discipline Outcomes	8
Primary Treatment Effect Models	8
Different Levels of Fidelity	8
Attendance	9
Primary Treatment Effect Model	9
Different Levels of Fidelity	9
Discussion	10
Implications for Research and Practice	10
Limitations	11
Conclusion	11
References	12

LIST OF TABLES

Table 1. Composite Scores for School Environment and Social and Emotional Skills in the Fall and Spring of the 2023-2024 School Year 14

Table 2. Unweighted and Weighted Covariate Balance 14

Table 3. School Environment and Social and Emotional Skills Models Intra-Class Correlation Coefficient 15

Table 4. Multilevel Model Estimates of Second Step Implementation School Environment and Social and Emotional Skill Change Scores 15

Table 5. Multilevel Model Estimates of SS3 Dosage on SEL Change Scores 16

Table 6. Fixed Effects from Logistic Mixed-Effects Models Predicting Discipline Outcomes with Dichotomous Treatment Indicator 16

Table 7. Fixed Effects from Logistic Mixed-Effect Models Predicting Discipline Outcomes with Varying Fidelity Levels for the Treatment Indicator 17

The Impact of the Second Steps Program on Elementary Students' Social and Emotional Skills, Exclusionary Discipline, and Attendance

Universal social–emotional learning (SEL) programs are increasingly recognized as essential components of whole-child education, with the potential to support students' emotional regulation, interpersonal skills, and academic engagement (Durlak et al., 2011; Jones & Kahn, 2017). Designed for delivery to all students, regardless of risk status, universal SEL programs aim to create positive school climates and equip children with foundational competencies such as self-awareness, social awareness, self-management, relationship skills, and responsible decision-making (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2023).

The Second Step® Elementary Digital Program, one of the most widely adopted universal SEL curricula, is a classroom-based program designed to promote social–emotional competencies in students from prekindergarten through middle school. The program includes scripted lessons and activities focusing on emotion management, empathy, problem-solving, and academic success. Although a growing body of research has linked SEL programs like Second Step to improved student outcomes, findings have been mixed—particularly regarding long-term behavioral change and schoolwide climate impacts (Sklad et al., 2012; Taylor et al., 2017).

Given the high implementation costs and time commitment associated with these programs, it is essential to assess not only whether Second Step is effective, but also under what conditions and with what dosage. Moreover, questions remain about the extent to which gains in student-level SEL outcomes translate to improvements in behavior and attendance, especially in elementary school settings.

This study evaluates the implementation and impact of Second Step in large, diverse school district in North Carolina. Specifically, we examine whether students in schools implementing Second Step with fidelity demonstrate greater gains in perceptions of the school environment

and social and emotional skills, reduced behavioral incidents, and improved attendance relative to students in lower-implementation schools. Using a combination of multilevel models and propensity score weighting, we estimate both overall and dosage-based treatment effects on student outcomes, contributing to the growing body of literature on SEL effectiveness in real-world conditions.

Method

Sample

Data for this study comes from a large school district in the Southeastern United States. The school district enrolls just over 52,000 students in kindergarten through grade 12 across 81 schools, including 42 elementary schools. Districtwide, approximately 33% of students are White, 29% Black, 29% Hispanic, and 3% Asian. About 71% of students in the district qualify for free or reduced-price lunch. For this study, we used data from the 42 elementary schools from the 2023–2024 school year. We received student attendance and discipline data from 23,634 students in grades kindergarten through grade five, and school environment and social and emotional skills from 10,657 students in grades three through five. The student-level data was restricted to only grade-level, attendance, discipline, and perceptions of school environment and social and emotional skills. No student-level demographic data was available.

Measures

School Environment and Social and Emotional Skills

The Panorama Student Survey – Elementary Version is a validated student-report instrument designed to capture key aspects of students' experiences, perceptions, and competencies within the school environment. Developed by Panorama Education, the survey is grounded in frameworks from SEL, school climate, and student engagement research. It is widely used in K–12 settings to support data-driven decision-making in schools and districts.

The elementary version of the survey is specifically adapted for students in grades three through five, using age-appropriate language and item structures. It typically includes 3–5-point Likert-type response options (e.g., "Almost Never" to "Almost Always"), with sections designed to be administered online or in paper–pencil formats, depending on the district's implementation context. The elementary version measures multiple research-backed domains

that align with widely accepted SEL and climate frameworks (e.g., CASEL, National School Climate Center), including

- Classroom teacher–student relationships: Perceptions of care, support, and respect from teachers.
- School climate: Perceptions of the overall emotional and social tone of the school.
- Sense of belonging: Feelings of inclusion and acceptance within the school community.
- School safety: Student-reported feelings of physical and emotional safety.
- Self-efficacy: Students' confidence in their academic capabilities.
- Self-management: Students' ability to regulate emotions, behaviors, and focus in academic settings.

These domains are often reported as composite scores derived from multiple items within each subscale. Internal consistency (Cronbach's alpha) has been reported as acceptable to strong across subscales in elementary samples ($\alpha \approx 0.70$ – 0.90 ; Panorama Education, 2020). We calculated the mean, standard deviation, and Cronbach's alpha for each administration (fall and spring) of the subscales using data in hand (see Table 1). We found that the school safety, school climate, and fall self-efficacy domain alpha values were below .70, suggesting low reliability. Therefore, results from these domains should be considered exploratory.

Discipline

The district provided the number of office discipline referrals (ODR), in-school suspensions (ISS), and out-of-school suspensions (OSS) for each student. An ODR is formal documentation of student misbehavior that violates the school's Code of Student Conduct. These referrals are typically initiated by school staff, often teachers, and can lead to various disciplinary actions, including ISS or OSS. The specific behaviors warranting an ODR and the procedures following a referral are determined by individual school districts. ISS involves removing a student from their regular classroom setting and placing them in a designated area within the school for disciplinary reasons. This action is taken for one-half of a school day or more. The student remains under supervision and is expected to continue academic work during this time. The important distinction between an ISS and an OSS is that an ISS is used as a disciplinary measure that keeps the student within the school environment. OSS is an exclusion of a student from school attendance for disciplinary purposes. If a student is sent home for at least half of the school day, it is considered an OSS.

The mean count of ODR, ISS, and ODR during the 2023–2024 school year was 0.52 (SD = 2.62), 0.06 (SD = 0.47), and 0.07 (SD = 0.46), respectively. Each outcome had many zeros and large standard deviations. Therefore, we converted each of these counts to a dichotomous indicator for never having a disciplinary incident and at least one disciplinary incident for each student. Overall, we found that 13.2% of students had at least one ODR, 3.6% of students had at least one ISS, and 3.9% of students had at least one OSS.

Attendance

The district also provided the percentage of days absent for each student. This was defined as the number of days absent divided by the number of days expected for each student. Days absent included both excused and unexcused. The average percentage of days absent per student was 8% (SD = 7%).

Independent Variable

Second Step, developed by Committee for Children, is a comprehensive, web-based SEL curriculum designed for use in kindergarten through fifth grade. It is intended to be implemented universally—that is, delivered to all students—during regular class time by classroom teachers. The program is structured to promote the development of social–emotional competencies, support positive classroom climates, and reduce behavior problems through proactive skill-building.

In elementary schools, the Second Step Program is typically

- delivered weekly via short, scripted lessons accessible through an online platform;
- facilitated by classroom teachers, with no requirement for additional staff or outside providers;
- integrated into the school day, often during morning meetings, designated SEL blocks, or homeroom periods;
- sequenced by grade level, ensuring developmental appropriateness from kindergarten through grade five; and
- aligned with CASEL's five core SEL competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making.

Each lesson includes engaging multimedia content, including animations and videos, as well as class discussion prompts, reflection activities, and opportunities for practice.

Key Features

1. Digital access: Lessons and teaching materials are fully accessible online, allowing for flexible delivery across classroom and remote settings.
2. Grade-specific content: Each grade level includes a sequenced set of lessons designed to scaffold SEL skill development.
3. Interactive elements: The program uses animated scenarios and student engagement prompts to support participation and comprehension.
4. Skill reinforcement: Lessons build on one another across the year and include opportunities for skill reinforcement and generalization.

5. Embedded assessment: Some versions of the program offer formative check-ins and educator feedback tools.

To promote high-quality delivery and sustainability, Second Step offers

- professional learning modules for educators, including onboarding and continuous training through an online learning hub;
- implementation guidance documents to assist school leaders with planning, scheduling, and fidelity monitoring;
- schoolwide integration resources, such as posters, reminder cues, and family engagement materials;
- progress monitoring tools, including teacher self-assessments and fidelity tracking forms; and
- technical and instructional support provided by the Committee for Children, including implementation coaching for districts.

We used the percentage of lessons completed by each school as the indicator of fidelity.

Following recommendations from the Committee for Children, we defined adequate implementation as 60% or greater lessons completed. As noted below, we also explored differential impacts of high fidelity, defined as completing 80% or more of the lessons.

Data Analysis

Propensity Score Weighting and Covariate Balance

To adjust for baseline differences between treatment groups, we estimated propensity score weights using a logistic regression model predicting treatment assignment from a set of baseline covariates. The treatment indicator, Second Step (SS), was a binary indicator for students being in schools using the Second Step program with fidelity (> 60% lessons completed) condition, with 1 indicating assignment and 0 indicating comparison (< 60% lessons completed).

We used all of the fall 2023 Panorama composite scores and student's grade-level as baseline covariates for propensity weights. The model was estimated using complete case analysis, resulting in a sample of 11,228 students. The logistic regression model estimated propensity scores for each student, from which we derived the inverse probability of treatment weights (IPTW):

$$w_i = 1 / \hat{e}_i \text{ if } SS_i = 1; w_i = 1 / (1 - \hat{e}_i) \text{ if } SS_i = 0$$

where $\hat{e}_i$ is the predicted probability of receiving treatment from the propensity score model. The What Works Clearinghouse (WWC) recognizes propensity score methods (matching,

stratification, weighting) as valid strategies for reducing bias in quasi-experimental designs. Specifically, propensity score weighting (e.g., inverse probability of treatment weighting) can be used to equate treatment and comparison groups

To evaluate covariate balance, we used the `bal.tab()` function from the `cobalt` package in R to compute standardized mean differences (SMDs) before and after weighting. Balance diagnostics were conducted using the average treatment effect (ATE) estimate. All standardized mean differences after weighting were < 0.01 , indicating excellent covariate balance across all baseline characteristics (see Table 2).

Multilevel Models

We estimated a series of two-level random intercept models to assess the association between Second Step (SS) implementation and change in students' perceptions of school environment and social and emotional skills. The dependent variable in each model was a change score, calculated as the difference between spring 2024 and fall 2023 scores for each domain (e.g., Self Efficacy spring 2024—Self Efficacy fall 2023). We used change scores to reflect the impact, or change, in student perceptions from fall (baseline) to spring (post) by treatment condition.

The primary predictor was a binary indicator of program fidelity: $SS = 1$ for schools with $> 60\%$ of Second Step lessons completed, and $SS = 0$ otherwise. Models included school-level random intercepts to account for nesting. The model was specified as

$$Y_{ij} = \beta_0 + \beta_1 SS_{ij} + u_j + \epsilon_{ij}$$

where $u_j \sim N(0, \tau^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$. The inverse probability of treatment weights was applied to all models to adjust for baseline differences.

We then extended the models to three treatment groups ($SS3$): Second Step $< 60\%$ fidelity ($n = 5,799$; 25 schools), 60-79% fidelity ($n = 2,846$; 10 schools), and 80% and above ($n = 2,583$; 9 schools). Again, each model was specified as a two-level random intercept model:

$$Y_{ij} = \beta_0 + \beta_1 SS3_{ij} + u_j + \epsilon_{ij}$$

where Y_{ij} is the change in outcome for student i in school j , $SS3_{ij}$ is the ordinal dosage indicator, u_j is the school-level random effect ($u_j \sim N(0, \tau^2)$), and ϵ_{ij} is the student-level residual ($\epsilon_{ij} \sim N(0, \sigma^2)$). Weights based on inverse probability of treatment were again applied using estimated propensity scores. Finally, we calculated standardized mean differences (d) for statistically significant treatment effects from any of the models using the following formula:

$$SMD = \hat{\beta} / \sqrt{(\sigma^2 + \tau^2)}$$

To evaluate the effects of SS implementation on student discipline outcomes, we estimated generalized linear mixed-effects models (GLMMs) with a binomial distribution and logit link function. These models used the full sample of students (grades kindergarten through five). We

used the same approach as above, modeling SS implementation as a binary indicator and as a three-level indicator (SS3). Each model included fixed effects for SS or SS3 (reference = low fidelity for both models), the school-level average of students' sense of belonging in fall 2023, and student grade level modeled as a factor, with kindergarten as the reference group. We included the school-level average of students' sense of belonging in fall 2023 for two important reasons. First, we wanted to include a baseline measure beyond grade-level but wanted to use the full sample of students (i.e., grades kindergarten through five). The Panorama survey was only administered to students in grades three through five. Therefore, we aggregated within-school mean scores for the Panorama domains. Second, we excluded the social and emotional skills because those are indicators of individual student skills and not reflective of a schoolwide status. We then reviewed the three school environment domains and found that they were highly correlated with each other ($r > .80$). Therefore, to reduce multicollinearity, we identified the domain that was most strongly predictive of the outcome and least correlated with treatment (Steiner & Cook, 2013). Among the three school environment domains, sense of belonging had the highest correlation with the outcomes and smallest correlations with the treatment indicator based on point-biserial correlations.

All models included a random intercept for school to account for the nested structure of students within schools. Models were estimated using the *glmer* function from the *lme4* package with Laplace approximation. To facilitate interpretation of treatment effects, we calculated approximate odds ratio-based standardized effect sizes (i.e., Cohen's d) using the approximation from Chinn (2000):

$$d = \frac{\ln(OR) \times \sqrt{3}}{\pi}$$

This approach assumes a logistic model with a binary outcome and equal variance between groups.

Finally, we estimated a two-level random intercept model to explore the effect of SS and SS3 on student attendance. The outcome variable, a percentage, was skewed (2.7). Therefore, we used a square root transformation to more closely approximate normality (square root transformed skew = 0.7). We used the same approach as the discipline outcomes and included the school-level average of students' sense of belonging in fall 2023, and student grade level modeled as a factor, with kindergarten as the reference group in the model.

Results

School Environment and Social and Emotional Skills

Primary Treatment Effect Models

Table 3 provides the empty model intraclass correlation coefficients (ICCs). We observed small ICCs across all models, ranging from 0.075 to 0.153, suggesting modest between-school variation in students' SEL growth. School-level random intercepts were included to account for clustering. Across the six SEL outcome domains, one statistically significant effect was observed. Students in schools with adequate fidelity of implementation (>60% Second Step fidelity) demonstrated a greater increase in self-management compared to students in lower-implementation schools ($B = 0.0415$, $SE = 0.0191$, $p = .037$). The standardized mean difference for this effect was 0.08, indicating a small effect size. No significant differences were found for other SEL domains (see Table 4).

Different Levels of Fidelity

In our dose-response analysis using the SS3 variable (Table 5), only one statistically significant effect was again detected. Students in schools with a high level of Second Step implementation (>80% of lessons completed) demonstrated a significantly greater increase in self-management compared to students in schools with no or low usage ($B = 0.059$, $SE = 0.024$, $p = .020$). While other contrasts did not reach statistical significance, positive point estimates for high-dosage students were observed in five out of six outcome domains. The standardized mean difference for the significant self-management effect (high vs. none) was approximately 0.12, indicating a small effect size according to Cohen's conventions. These findings provide some evidence that higher exposure to the Second Step program may be associated with improvements in students' self-management skills over the academic year, though effects on other domains were small and nonsignificant.

Discipline Outcomes

Primary Treatment Effect Models

The ICC for the ODR, ISS and OSS models were 0.12, 0.36, and 0.27, respectively. These ICCs suggest little between-school variance for ODR, but moderate between-school variance for ISS and OSS. The GLMM models revealed that Second Step implementation fidelity was not significantly associated with the likelihood of ODR ($OR = 0.75$, $p = .183$), ISS ($OR = 0.61$, $p = .250$), or OSS ($OR = 0.50$, $p = .050$), although the effect for OSS was marginally significant. Across models, higher grade levels were consistently associated with increased odds of disciplinary outcomes. Mean sense of belonging in fall 2023 was not significantly associated with any outcome.

Different Levels of Fidelity

The model for ODR revealed no statistically significant effects of varying fidelity implementation of Second Step. Compared to students in low-fidelity schools, the odds of an ODR were not significantly lower in moderate-fidelity schools (SS31: OR = 0.92, $p = .738$) or high-fidelity schools (SS32: OR = 0.60, $p = .057$). Sense of belonging and grade level were positively associated with increased odds of incidents, with higher grade levels predicting more frequent behavioral incidents.

The model for ISS also showed no significant associations between Second Step fidelity and ISS. Odds ratios for moderate (OR = 0.67, $p = .455$) and high (OR = 0.55, $p = .271$) fidelity schools were not statistically different from low-fidelity schools. However, higher grade levels were strongly predictive of increased ISS likelihood.

In contrast, the model for OSS found a significant effect for high-fidelity implementation. Students in high-fidelity schools (SS32) had significantly reduced odds of out-of-school suspension compared to students in low-fidelity schools (OR = 0.37, $p = .028$). The approximate standardized effect size was $d = -0.37$, indicating a moderate effect size. The coefficient for moderate fidelity was not significant (SS31: OR = 0.65, $p = .302$). Sense of belonging trended toward a protective effect but was not statistically significant ($p = .071$). Again, higher grade levels were strongly predictive of increased OSS.

Attendance

Primary Treatment Effect Model

The ICC was estimated as 0.072, indicating that 7.2% of the variance in absenteeism was attributable to differences between schools. The effect of Second Step implementation on student absenteeism was not statistically significant, suggesting no meaningful difference in absenteeism rates between students in schools implementing the program with high versus low fidelity. School-level sense of belonging was significantly associated with lower absenteeism ($b = -0.0744$, $SE = 0.0287$, $p = .015$), indicating that higher average levels of belonging in schools were associated with lower average student absence rates. There were also statistically significant effects for student grade level.

Different Levels of Fidelity

The model revealed a significant negative association between mean sense of belonging and student absences ($b = -0.077$, $SE = 0.028$, $p = .011$), indicating that higher belonging was associated with lower student absenteeism. Students in schools with the highest fidelity of Second Step implementation (SS3 = 2) had marginally lower absenteeism compared to those in low fidelity schools ($b = -0.025$, $SE = 0.012$, $p = .059$). The effect size (d) is -0.22. No significant difference was observed between moderate fidelity (SS3 = 1) and low fidelity (SS3 = 0) schools ($p = .87$). Grade level was significantly associated with absenteeism, with older students exhibiting slightly lower rates.

Discussion

This study examined the effects of the Second Step Program on elementary students' social-emotional competencies, behavior, and attendance in a large urban district using rigorous propensity score weighting and multilevel modeling approaches. Overall, we observed small but meaningful benefits associated with the high-fidelity implementation of Second Step.

In the school environment and social and emotional skills domains, students in schools with adequate program fidelity (>60%) demonstrated statistically significant improvements in self-management ($B = 0.042$, $p = .037$; $d = 0.08$), a key social and emotional skill related to emotion regulation and attention control. Moreover, students in schools with high fidelity (>80%) showed even greater gains in self-management ($B = 0.059$, $p = .020$; $d = 0.12$). These findings suggest a potential dose-response relationship between Second Step exposure and improvement in students' regulation and goal-setting skills, consistent with prior research (e.g., Espelage et al., 2013). No statistically significant effects were observed for the other related outcomes, although point estimates generally favored students in high-fidelity schools.

In terms of disciplinary outcomes, we found mixed evidence. While Second Step implementation was not significantly associated with reduced office discipline referrals or in-school suspensions, students in high-fidelity schools had significantly lower odds of receiving out-of-school suspensions compared to students in low-fidelity schools ($OR = 0.37$, $p = .028$; $d = 0.37$). This represents a meaningful reduction in the likelihood of exclusionary discipline for students in schools with strong Second Step implementation.

Attendance outcomes showed no significant main effects for treatment. However, the model revealed a marginal effect for high-fidelity implementation ($B = -0.025$, $p = .059$; $d = -0.22$), suggesting that students in these schools may experience slightly lower rates of absenteeism. Additionally, a significant association was found between school-level sense of belonging and attendance, indicating that schools with higher perceived belonging had lower rates of student absence.

Implications for Research and Practice

These findings offer nuanced support for the effectiveness of Second Step in elementary settings. While effects were generally small, the consistent improvement in self-management and reduction in OSS among students in high-fidelity schools suggests that when implemented well, universal SEL programs can positively influence important student outcomes. These results underscore the importance of program fidelity as a critical moderator of SEL effectiveness.

For practitioners, the findings suggest that simply adopting an SEL curriculum may be insufficient. Instead, school leaders should invest in implementation support, teacher training, and monitoring to ensure delivery with fidelity. Furthermore, the apparent association between school belonging and positive outcomes in both attendance and behavior underscores the importance of fostering a strong school climate alongside SEL instruction.

For researchers, this study highlights the utility of combining propensity score methods with multilevel modeling to estimate treatment effects in real-world school contexts. Future work should continue to explore mediating mechanisms (e.g., how teacher implementation quality influences outcomes), longitudinal effects, and potential equity implications by student subgroups.

Limitations

This study has several limitations. First, the lack of student-level demographic data limited our ability to examine differential effects by race/ethnicity, socioeconomic status, or other factors. Second, although we used rigorous weighting procedures for some models, residual confounding may remain due to unmeasured variables. Third, the use of change scores assumes linear change and may be sensitive to regression to the mean. Fourth, self-reported school environment and social and emotional skills outcomes are subject to social desirability bias, and some subscales had lower reliability estimates (e.g., school safety, climate). Finally, while implementation fidelity was measured through lesson completion rates, more granular fidelity data (e.g., quality of delivery) could provide a clearer picture of dosage effects.

Conclusion

In sum, this study provides evidence that high-fidelity implementation of the Second Step program can support modest improvements in students' self-management skills and reduce out-of-school suspensions. The findings suggest that implementation quality matters and that supportive school climates, particularly feelings of belonging, may be an important contextual factor. Continued research is needed to identify the conditions under which SEL programs yield the greatest benefits and how they can be sustained in diverse school environments.

References

- Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, 19(22), 3127–3131. [https://doi.org/10.1002/1097-0258\(20001130\)19:22%3C3127::AID-SIM784%3E3.0.CO;2-M](https://doi.org/10.1002/1097-0258(20001130)19:22%3C3127::AID-SIM784%3E3.0.CO;2-M)
- Collaborative for Academic, Social, and Emotional Learning. (2023). *What is SEL?* <https://casel.org/what-is-sel/>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Espelage, D. L., Low, S., Polanin, J. R., & Brown, E. C. (2013). The impact of a middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health*, 53(2), 180–186. <https://doi.org/10.1016/j.jadohealth.2013.02.021>
- Jones, S. M., & Kahn, J. (2017). *The evidence base for how we learn: Supporting students' social, emotional, and academic development*. The Aspen Institute National Commission on Social, Emotional, and Academic Development. <https://nationathope.org/>
- Panorama Education. (2020). *Reliability and validity of Panorama's survey topics for students: 2020 update*. Panorama Education.
- Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892–909. <https://doi.org/10.1002/pits.21641>
- Steiner, P. M., & Cook, T. D. (2013). Matching and propensity scores: The state of the art. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods: Vol. 1. Foundations* (pp. 259–284). Oxford University Press.
- Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-

analysis of follow-up effects. *Child Development*, 88(4), 1156–1171.
<https://doi.org/10.1111/cdev.12864>

Table 1. Composite Scores for School Environment and Social and Emotional Skills in the Fall and Spring of the 2023-2024 School Year

Composite	Fall 2023 M (SD) α	Spring 2024 M (SD) α
School Environment		
Classroom Teacher–Student Relationships	4.08 (0.82), $\alpha = .70$	4.01 (0.87), $\alpha = .74$
School Safety	3.53 (0.86), $\alpha = .63$	2.48 (0.87), $\alpha = .66$
Sense of Belonging	3.80 (0.85), $\alpha = .71$	3.71 (0.85), $\alpha = .73$
School Climate	4.67 (1.13), $\alpha = .62$	4.54 (1.15), $\alpha = .65$
Social and Emotional Skills		
Self-Efficacy	3.47 (0.78), $\alpha = .66$	3.50 (0.78), $\alpha = .70$
Self-Management	3.92 (0.68), $\alpha = .81$	3.92 (0.69), $\alpha = .83$

Table 2. Unweighted and Weighted Covariate Balance

Covariate	Type	Unweighted SMD	Weighted SMD
Classroom Teacher-Student Relationships (Fall 2023)	Continuous	0.128	-0.0005
School Safety (Fall 2023)	Continuous	0.103	-0.0033
Sense of Belonging (Fall 2023)	Continuous	0.089	-0.0007
School Climate (Fall 2023)	Continuous	0.112	-0.0017
Self-Efficacy (Fall 2023)	Continuous	0.095	0.0009
Self-Management (Fall 2023)	Continuous	0.076	-0.0040
Grade 3	Binary	0.030	-0.0001
Grade 4	Binary	0.044	0.0005
Grade 5	Binary	0.027	-0.0004

Table 3. School Environment and Social and Emotional Skills Models Intra-Class Correlation Coefficient

Outcome	Random Intercept Variance (School)	Residual Variance	ICC
Teacher-Student Relationships	0.021	0.260	0.075
School Safety	0.085	0.470	0.153
Sense of Belonging	0.034	0.250	0.120
School Climate	0.041	0.290	0.124
Self-Efficacy	0.027	0.240	0.101
Self-Management	0.032	0.220	0.127

Table 4. Multilevel Model Estimates of Second Step Implementation School Environment and Social and Emotional Skill Change Scores

Outcome	Estimate	SE	CI Lower	CI Upper	p
Teacher-Student Relationships	-0.025	0.036	-0.098	0.048	0.490
School Safety	-0.048	0.209	-0.470	0.375	0.821
Sense of Belonging	-0.020	0.026	-0.073	0.033	0.444
School Climate	0.023	0.054	-0.087	0.133	0.673
Self-Efficacy	0.009	0.025	-0.043	0.060	0.736
Self-Management	0.042	0.019	0.003	0.080	0.037

Table 5. Multilevel Model Estimates of SS3 Dosage on SEL Change Scores

Outcome	SS3 Level	B	SE	95% CI Lower	95% CI Upper	p
Teacher-Student Relationships	Low vs None	-0.053	0.044	-0.142	0.035	0.228
Teacher-Student Relationships	High vs None	0.006	0.045	-0.086	0.098	0.890
School Safety	Low vs None	0.201	0.251	-0.306	0.707	0.428
School Safety	High vs None	-0.324	0.260	-0.850	0.202	0.220
Sense of Belonging	Low vs None	-0.034	0.032	-0.098	0.031	0.295
Sense of Belonging	High vs None	-0.005	0.033	-0.072	0.063	0.890
School Climate	Low vs None	0.016	0.067	-0.120	0.152	0.809
School Climate	High vs None	0.031	0.070	-0.110	0.172	0.660
Self-Efficacy	Low vs None	-0.020	0.030	-0.081	0.041	0.514
Self-Efficacy	High vs None	0.041	0.031	-0.023	0.105	0.202
Self-Management	Low vs None	0.026	0.023	-0.021	0.073	0.267
Self-Management	High vs None	0.059	0.024	0.010	0.108	0.020

Table 6. Fixed Effects from Logistic Mixed-Effects Models Predicting Discipline Outcomes with Dichotomous Treatment Indicator

Outcome	Predictor	Estimate	SE	OR	95% CI	p
Incident	SS	-0.284	0.214	0.75	[0.50, 1.05]	.183
ISS	SS	-0.493	0.429	0.61	[0.26, 1.29]	.250
OSS	SS	-0.694	0.354	0.50	[0.25, 1.00]	.050

Note. All models included grade-level as a fixed-effect factor covariate. ISS is in-school suspensions, OSS is out-of-school suspensions

Table 7. Fixed Effects from Logistic Mixed-Effect Models Predicting Discipline Outcomes with Varying Fidelity Levels for the Treatment Indicator

Outcome	Predictor	Estimate	Std. Error	OR	95% CI (OR)	p-value
Incident	SS31 (mod fidelity)	-0.08	0.25	0.92	[0.56, 1.29]	.738
	SS32 (high fidelity)	-0.51	0.27	0.60	[0.36, 1.03]	.057
ISS	SS31	-0.39	0.53	0.68	[0.24, 1.75]	.455
	SS32	-0.60	0.55	0.55	[0.20, 1.33]	.271
OSS	SS31	-0.44	0.42	0.65	[0.30, 1.34]	.302
	SS32	-0.99	0.45	0.37	[0.16, 0.97]	.028 *

Note: All models included grade-level as a fixed-effect factor covariate. ISS is in-school suspensions, OSS is out-of-school suspensions

