The Effects of Second Step on Middle School Students' Academic, Behavioral, and Social–Emotional Well-being

Nicholas A. Gage, PhD

April 2025

WestEd.org

Gage, N. A. (2025). *The effect of Second Step on middle school students' academic, behavioral, and social–emotional well-being.* 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. 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.

Executive Summary

This report describes a study evaluating the impact of the Second Step Digital Middle School Program, a universal social-emotional learning (SEL) curriculum, on academic achievement, behavior, attendance, and social–emotional well-being among 4,903 middle school students in a large school district in the Southeastern U.S. over two years. Using quasi-experimental methods with propensity score weighting and difference-in-differences models, the study found that faithful implementation of Second Step led to small but significant improvements in English language arts (ELA) performance, substantial reductions in behavioral infractions (office referrals, in-school and out-of-school suspensions), and increased school attendance (about 2.5 more days per year). Additionally, treated students reported better teacher-student relationships, greater school belonging, and improved perceptions of school climate, though no significant changes were observed in self-management or peer supportive relationships. The findings highlight the value of sustained SEL programming during the challenging middle school years for enhancing multiple dimensions of student development, while noting limitations such as lack of individual baseline academic data and generalizability concerns.

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

LIST OF TABLES

The Effects of Second Step on Middle School Students' Academic, Behavioral, and Social–Emotional Well-being

Middle school is a critical developmental time for students, marked by profound changes in identity formation, peer dynamics, and academic demands. During this period, students often experience declining motivation, increases in behavioral challenges, and dips in academic performance and school connectedness. National data reflect troubling trends: disciplinary incidents increase in early adolescence, school attendance declines, and students report lower levels of emotional well-being as they transition through middle grades. These challenges underscore the need for school-based strategies that can simultaneously address academic engagement, behavioral outcomes, and social–emotional health. In this study, we explore the impact of a universal social–emotional Learning (SEL) program, Second Step, on students' academic achievement, behavior, school attendance, and social–emotional well-being.

Challenges in Middle School

Developmentally, middle school is characterized by significant transformations in identity formation, peer dynamics, and academic expectations. Adolescents during this period begin to re-examine and reconstruct their self-concepts as they interact with new social environments and evolving peer groups (Onetti et al., 2019). Prior research has demonstrated that active engagement in school not only facilitates academic progress but also supports the development of a cohesive identity, whereas school burnout may impede this process (Erentaitė et al., 2018). Furthermore, as young adolescents increasingly seek to define themselves, factors such as ethnic and social identity become more pronounced, further highlighting the complex interplay between individual development and the school environment (Erentaitė et al., 2018).

The transition into middle school often introduces heightened academic demands and new organizational structures that can clash with adolescents' growing need for autonomy. This mismatch can lead to a decline in intrinsic motivation as students struggle to meet increased academic expectations without corresponding support from educators. Empirical evidence indicates that when teachers are viewed as supportive—by maintaining high expectations and personalizing instruction—students tend to exhibit higher levels of motivation (Hornstra et al., 2018). Conversely, the absence of such supportive practices can exacerbate motivational decline and contribute to adverse behavioral outcomes during this transitional period. Moreover, the shift from elementary to middle school has been associated with a range of negative outcomes, including reduced academic performance and increased behavioral challenges, largely due to the adjustment stress imposed by the new educational environment (Akos, 2006).

Another critical aspect during the middle school years is the erosion of school connectedness. Empirical studies have found that disruptions in students' feelings of belonging and support within the school community are closely linked to declines in academic achievement and increases in disciplinary issues (Niehaus et al., 2012). As academic pressures mount and peer dynamics evolve, students who experience lower levels of connectedness may become more susceptible to behavioral problems and disengagement from school activities—factors that further compound academic difficulties (Niehaus et al., 2012). Together, these findings underscore that middle school is a complex phase during which declines in motivation, emergent behavioral challenges, and reductions in school connectedness collectively contribute to challenges in academic performance and personal development.

Universal SEL curricula have gained attention as a promising approach to fostering students' interpersonal skills, self-regulation, and resilience. By proactively embedding SEL instruction into the fabric of the school day, these programs aim to build a foundation for improved student behavior, greater school belonging, and ultimately, stronger academic outcomes. Importantly, universal SEL approaches are designed to reach all students—not only those identified as at risk—thereby promoting equitable access to skill-building opportunities and preventive supports.

A growing body of research has demonstrated the positive effects of SEL programs on a range of student outcomes. Meta-analyses suggest that SEL interventions are associated with improvements in students' emotional regulation, social competence, and attitudes toward school, as well as reductions in conduct problems and disciplinary referrals. Furthermore, SEL participation has been linked to small but meaningful gains in academic achievement. However, the magnitude and consistency of these effects vary, particularly across developmental stages and implementation contexts. Middle school, in particular, poses unique challenges for SEL implementation given students' increasing autonomy and the complexity of peer relationships during early adolescence.

The **Second Step Middle School Program** is one of the most widely used universal SEL curricula in United States schools. Grounded in cognitive–behavioral and social–emotional development theory, Second Step focuses on teaching emotion regulation, empathy, problem-solving, and responsible decision-making. While prior evaluations have shown that Second Step can reduce bullying and aggression and increase prosocial behavior, relatively few studies have examined its effects on broader school functioning indicators, such as attendance and academic performance, especially within rigorous analytic frameworks and real-world settings.

Moreover, few evaluations of universal SEL programs in middle school have taken a comprehensive approach to assessing multiple domains of student functioning simultaneously. Much of the existing literature tends to focus on either behavioral or socioemotional outcomes in isolation, without considering how these dimensions intersect with academic engagement and school climate. There is a growing need for research that integrates multiple student-level outcomes, leverages longitudinal designs, and assesses program effects within authentic school contexts.

This study addresses these gaps by evaluating the effects of the Second Step Middle School Program on student academic achievement, behavior, attendance, and well-being. Using a difference-in-differences design with school and student-level controls, we examine whether exposure to Second Step is associated with changes in course grades, disciplinary outcomes, school attendance, and student-reported measures of school climate and social–emotional skills. Our goal is to provide a rigorous and comprehensive assessment of Second Step's impact in middle school settings, contributing to the evidence base for universal SEL programming during a critical period of student development.

Method

Sample

We collected de-identified student-level data from 4,903 middle school students in four middle schools from a large school district in South Carolina from the 2022–2023 and the 2023–2024 school year. Two middle schools were defined by the Committee for Children and the school district as high implementers of the digital Second Step program, defined as completing more than 60 percent of the lessons. The two middle schools served as the treatment group. Two different middle schools had access to the Second Step program, but did not complete any lessons or only a few lessons sporadically. These two middle schools served as the comparison group. Table 1 provides descriptive statistics for the students. Over half of the students in both conditions were White. There were more Hispanic students in the comparison condition and more Black students and multiracial students in the treatment condition. There were also more students with disabilities in the treatment schools.

Measures

Academic Performance

Student achievement in English language arts (ELA) and mathematics was measured using final course grades reported by the district. Grades were scaled from 0 to 100 and reflect cumulative performance across the academic year. These grades were treated as continuous outcomes, with higher values indicating stronger academic performance, and were available for a wide range of course enrollments and were standardized across schools and grade levels. These course grades were

For ELA, students were enrolled in a variety of grade-level and accelerated courses, including standard and honors designations. Common courses included ELA 6 (e.g., 10010600), ELA 7 (20010700), ELA 8 (20010800), and their honors equivalents (e.g., 1001H600, 2001H700, 2001H800). Additional ELA courses included English 1 (302400CW) and English 1 Honors (302400HW), as well as specialized offerings such as Reading Assistance and Language Arts Extensions (e.g., 10240600, 2916X800, 2910X700).

The same was true for mathematics, with data available from enrollment in grade-level and accelerated courses, including standard and honors designations achievement. These included general education mathematics courses (e.g., 11100600 for Math 6, 21100700 for Math 7), honors courses (e.g., 1110H600, 2110H700, 2110H800), and advanced mathematics offerings such as Algebra 1 (411400CW), Algebra 1 Honors (411400HW), Geometry Honors (412200HW), and various accelerated tracks (e.g., 2916X801, 1916X601).

Behavioral Outcomes

Behavioral outcomes were captured using administrative records from the district's student information system. Three cumulative indicators were used to quantify student behavior across the school year:

- **Office Disciplinary Referrals (ODRs)**: The total number of documented behavioral referrals submitted by school staff for rule violations or conduct issues.
- **In-school Suspensions (ISS)**: The cumulative number of times a student was assigned to in-school suspension as a disciplinary response.
- **Out-of-school Suspensions (OSS)**: The cumulative number of times a student was suspended from school and not allowed on campus.

Each of these variables was treated as a count outcome, with higher values indicating more frequent behavioral infractions.

Attendance

Student attendance was measured using the total number of **days attended** during the school year, as reported by the district's attendance tracking system. This measure reflects the number of days a student was marked present for instruction and was used as a continuous indicator of school engagement. Days attended were treated as a count variable, with higher values reflecting greater instructional access and participation.

Social–Emotional Well-being

We used the *Panorama Student Survey*, a validated instrument designed to capture student perceptions on various aspects of their educational experience, including teaching, learning, and school climate. The survey incorporates multiple topics, such as pedagogical effectiveness, school climate, and student engagement. We used the Panorama subtests used by middle schools in the school district, which included the following:

- *Teacher-Student Relationships*: Evaluates the strength of the social connection between teachers and students within and beyond the classroom. An example item is: "If you walked into class upset, how concerned would your teacher be?". All items used a 5-point Likert scale.
- *School Belonging*: Assesses the extent to which students feel valued and included in their school community. An example item is: "Overall, how much do you feel like you belong at your school?" All items used a 7-point Likert scale.
- *School Climate:* Measures students' perceptions of the overall social and learning environment of the school. Only one item was used by the district: " How often do your teachers seem excited to be teaching your class?" The one item used a 5-point Likert scale.
- *Self-Management:* Assesses students' ability to manage their emotions, thoughts, and behaviors in different situations. An example item is: "During the past two weeks, how often did you come to class prepared?" All items used a 5-point Likert scale.
- *Supportive Relationships*: Examines the quality of students' relationships that provide support within the school context. All items were dichotomous.

Table 2 provides descriptive statistics and reliability (i.e., internal consistency) using data-in-hand for each fall and spring administration of the Panorama survey. A few important patterns emerged from these data. First, for both school years, the fall scores are slightly higher than the spring scores. Second, all measures demonstrate adequate reliability (> .70) except for the school climate measure because there is only one item, therefore composite reliability isn't relevant in the Supportive Relationships domain. The low reliability score for the Supportive Relationships domain suggests that results for this composite score should be interpreted with caution. Table 3 provides the descriptive statistics by Panorama domain and measurement period and Table 4 provides the raw standardized mean differences by domain and by measurement period.

Data Analysis

Propensity Score Weighting

We used inverse probability of treatment weighting (IPTW) using propensity scores to address potential selection bias and establish baseline equivalence between the treatment and comparison groups. The propensity score, representing the probability of a student being assigned to the treatment group given their observed characteristics, was estimated using logistic regression. Predictor variables included gender, race/ethnicity, English language learner status, students with disabilities status, grade level, and fall 2022 scores from the Panorama domains. Students with missing data on these covariates were excluded from the propensity score estimation.

Weights were calculated as follows:

- Treatment Group: Weight = 1 / Propensity Score
- Control Group: Weight = 1 / (1 - Propensity Score)

This approach creates a weighted sample wherein the distribution of observed covariates is independent of treatment assignment, approximating the conditions of a randomized controlled trial. This approach ensures that the student meets the What Works Clearinghouse (WWC) evidence standards with reservations. 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

Covariate Balance Assessment

We conducted a covariate balance check using standardized mean differences (SMDs) to evaluate the success of the weighting strategy in balancing baseline characteristics between the groups. SMDs were computed using the *bal.tab()* function from the *cobalt* package, and inverse probability weights were applied to adjust for covariate differences. The results are presented in Table 5. All baseline covariates, including fall 2022 Panorama composite scores, were equivalent, defined as < 0.25 standard deviation units. The balance analysis included 1,585 students in the control group and 750 students in the treatment group.

Modeling Approach

Academic Achievement and Panorama Composite Scores

We used a Difference-in-differences (DiD) design and fixed-effects linear regression models to evaluate the impact of the intervention on academic achievement and Panorama composite scores. This approach accounts for unobserved, time-invariant individual characteristics by including student-level and school-level fixed effects, isolating within-student variation over time. The primary independent variables were the treatment indicator, time-period indicators, and their interactions, allowing assessment of differential changes between the treatment and control groups across specified time frames. Propensity score weights were incorporated to adjust for baseline differences between groups, and standard errors were clustered at the school level to account for intra-school correlations.

The fixed-effects linear regression model is specified as follows:

We estimate the treatment effect using the following difference-in-differences model:

$$Yit_{it} = \beta_0 + \beta_1 Treatmenti_i + \beta_2 Postt_t + \beta_3(Treatmenti_i \times Postt_t) + \alpha i_i + \epsilon it_{it}$$

where

- *Yit* represents the outcome measure for student *i* at time *t*.
- $Treatment_i$ is a binary indicator for assignment to the treatment group.
- *Postt* is a binary indicator for the post-treatment period.
- $Treatment_i \times Postt$ is the interaction term capturing the difference-in-differences (DiD) estimate of the treatment effect.
- *αi* denotes individual fixed effects to control for time-invariant characteristics.
- *ϵit* is the error term.

This model leverages the DiD approach to compare changes over time between the treatment and control groups, effectively controlling for unobserved, time-invariant factors that could confound the estimated treatment effects.

This model satisfies the baseline equivalence requirement of to meet the What Works Clearinghouse (WWC) standards with reservations for the Panorama composite score models because the fall 2022 (baseline) Panorama composite scores are in included in the models in the propensity score weights. Unfortunately, we do not have a baseline student achievement score. Therefore, to meet the WWC standards with reservations, we added the school-level percentage of students performing at the meets or exceeds level on the SC READY in English language arts and mathematics from the 2021–2022 (year prior to Second Step implementation) to the achievement models. Per the WWC, to meet standards with reservations in the absence of individual-level pretest data, the analysis is required to meet baseline equivalence on (a) a broad, continuous, and standardized measure of student academic achievement, and (b) at least two demographic characteristics, such as grade level and race/ethnicity (WWC, 2024, pp. 54–55). Table 6 provides the school-level mean and standard deviations for each content area. The pretreatment differences between groups were $d = -0.06$ for ELA and $d = -0.15$ for mathematics, both below the 0.25 equivalence threshold. Therefore, adding the prior-year school-level achievement satisfies the baseline requirement.

Behavioral and Attendance Outcomes

We estimated a series of negative binomial regression models to examine the effects of treatment exposure and time on student behavioral and attendance outcomes, including ODRs, ISS, OSS, and days attended. Given the count nature of these dependent variables and evidence of overdispersion, we used a negative binomial model specification (glm.nb from the *MASS* package in R) with analytic weights and school fixed effects.

Each model included a binary indicator for treatment status (treat), a categorical variable representing school year (year), and their interaction. The models also controlled for prior academic achievement using the school-level 2021–2022 ELA performance. The fixed-effects specification in the models controlled for unobserved, time-invariant individual characteristics, isolating the within-student variations over time. Propensity score weights were applied to adjust for baseline differences between the treatment and control groups, ensuring comparability. Standard errors were clustered at the school level to account for intraschool correlations, acknowledging that students within the same school may exhibit correlated behaviors.

The model took the following form:

$$\log(\mu_i) = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{Year}_i + \beta_3(\text{Treatment}_i \times \text{Year}_i) + \beta_4 \text{ELA22}_i + \gamma_s(_i) + \log(w_i)$$

where

- μ_i is the expected count of the outcome for student i,
- Treatment_i is a binary indicator of treatment assignment,
- Year_i captures post-treatment years,
- $\text{Treatment}_i \times \text{Year}_i$ is the interaction term representing the DiD estimate,
- ELA22_i is the prior ELA score,
- $\gamma_s(_i)$ are school fixed effects, and
- - w_i is the analytic weight (included via the weights argument).

To aid interpretation, we exponentiated the model coefficients to produce incidence rate ratios (IRRs). In the context of a negative binomial model, the IRR expresses the proportional change in the expected count of the outcome associated with a one-unit increase in the predictor variable, holding other variables constant. For example, an IRR of 1.20 for the treatment-by-year interaction term would indicate that students in treatment schools had 20% more of the outcome (e.g., days attended or ODRs) relative to comparison students in that year. Conversely, an IRR of 0.80 would indicate a 20% reduction in the expected count. We report both the IRRs and their 95% confidence intervals to reflect the uncertainty of the estimates. Separate models were estimated for each outcome variable (ODRs, ISS, OSS, and attendance), allowing us to examine whether the treatment was associated with improvements in behavior and attendance over time.

Results

Academic Achievement

We estimated weighted Ordinary Least Squares (OLS) models predicting grade 1 ELA and mathematics scores for the 2022–2023 and 2023–2024 school years to evaluate the effect of treatment on academic outcomes. Each model controlled for school-level baseline (2021–2022 school year) performance in the relevant subject, included fixed effects for School ID, and applied analytic weights (wt). Models were restricted to complete cases on all included variables. The results are presented in Table 7.

Treatment (i.e., Second Step implementation with fidelity) was significantly associated with higher ELA scores, even after adjusting for students' baseline achievement. In 2022–2023, students in the treatment group scored an average of 2.16 points higher on ELA than those in the control group ($SE = 0.51$, $p < .001$), corresponding to $d = 0.14$. This advantage slightly increased in 2023–2024, where the treatment effect was 2.34 points ($SE = 0.49$, $p < .001$), equivalent to $d = 0.17$. These results reflect small but consistent gains in ELA outcomes for treatment students over two years.

In contrast, treatment effects on mathematics scores were not statistically significant in either year. In 2022–2023, the treatment effect was close to zero ($b = 0.38$, $p = .45$; $d = 0.02$), and in 2023–2024, while slightly larger, it remained nonsignificant ($b = 0.64$, $p = .17$; $d = 0.05$). These effect sizes suggest that the treatment had little to no measurable impact on mathematics achievement in either year.

Behavioral Outcomes

We used weighted negative binomial regression to model student behavior outcomes (referrals, ISS, and OSS) as count variables (see Table 8). Each model included a treatment-by-year interaction, baseline ELA scores (ELA22), and fixed effects for school. Students in the treatment group received significantly fewer ODR overall ($b = –0.44$, $SE = 0.07$, $p < .001$). The interaction between treatment and spring 2024 was not statistically significant ($b = –0.17$, $p = .10$), suggesting that the treatment effect on referrals was relatively consistent across both treatment years. Treatment students had significantly fewer ISS overall ($b = –0.37$, $SE = 0.10$, $p < .001$), with a significantly larger reduction in spring 2024 ($b = –0.45$, $p = .003$). This suggests the treatment had a **stronger impact on reducing ISS after two years** of implementation. Treatment was also associated with fewer OSS incidents overall ($b = –0.22$, $SE = 0.11$, $p = .036$), with a significant interaction in spring 2024 (b = –0.40, p = .012), suggesting an **increased treatment effect over time.**

To aid in the interpretation of the treatment effects, we exponentiated the negative binomial regression coefficients to obtain **incidence rate ratios** (IRRs), which reflect the **multiplicative change in the expected count** of behavioral outcomes associated with treatment exposure.

For **ODR**, the main effect of treatment indicated that students in the treatment group received approximately **36% fewer referrals** than students in the control group (*IRR* = 0.64, standardized mean difference [g] = -0.81). While the interaction with year was not statistically significant, the effect was consistent across time points, suggesting a robust pattern of reduced referrals. For **ISS,** the treatment effect in spring 2024 was particularly pronounced. The interaction term indicated that treatment students had **36% fewer ISS incidents** in spring 2024 relative to control students during the same period
(*IRR* = 0.64/g = -0.81).

Similarly, for **OSS**, the interaction between treatment and spring 2024 was significant, with treated students experiencing **33% fewer OSS incidents** than control students (*IRR* = 0.67, g = -0.73) in that year. This pattern suggests that the intervention became **increasingly effective in reducing more severe behavioral infractions** in its second year of implementation.

Attendance

We used a weighted negative binomial regression to predict the number of days attended, adjusting for treatment status, year, prior year ELA performance, and school fixed effects.

Results suggest that students in the treatment group attended significantly more days overall than their peers in the control group (b = 0.014, *SE* = 0.002, p < .001). This main effect corresponds to an **IRR of 1.014,** indicating that, on average, treatment students attended **1.4% more school days** than control students. Assuming a 180-day school year, this translates to approximately **2.5 additional days of attendance** for treatment students compared to their peers.

The interaction between treatment and spring 2024 was not statistically significant (b = 0.002, p = .575), suggesting that the positive effect of the intervention on attendance was stable across both years of implementation. Similarly, there was no significant main effect of spring 2024 (p = .666).

Baseline ELA performance was positively associated with attendance (b = 0.00052, p < .001), indicating that schools that have higher-performing students tended to attend school more consistently.

Panorama Domains

We estimated DiD models to examine the effects of Second Step on perceived school climate and well-being outcomes, focusing on changes from spring 2023 to spring 2024. The results are presented in Table 10. All models included student-level fixed effects and school fixed effects and were weighted using the propensity score inverse probability weights. Interaction terms between treatment and spring 2024 were used to isolate the differential impact of the intervention over time.

Statistically significant and positive DiD effects were found for several school climate constructs. Students in the treatment group experienced a statistically significant increase in **teacher–student relationships** ($b = 0.22$, $p = .001$), **school belonging** ($b = 0.21$, $p = .003$), and **school climate** ($b = 0.20$, $p = .031$) compared to control students. These changes corresponded to d **effect sizes of 0.40, 0.37, and 0.27**, respectively, indicating **moderate and meaningful gains** in students' perceptions of their school environments over time. No significant DiD effects were found for **self-management** ($b = 0.04$, $p = .26$; $d = 0.12$) or for **supportive relationships** ($b = –0.01$, $p = .60$; $d = –0.06$), suggesting that the intervention had less impact on students' intrapersonal regulation or peer-based support networks.

Discussion

This study evaluated the effects of the Second Step Middle School Program on a comprehensive set of student outcomes: academic achievement, behavioral incidents, school attendance, and social–emotional well-being. Findings contribute to the growing evidence base for universal SEL programs in secondary schools and underscore the importance of attending to the multiple dimensions of students' school experience during early adolescence.

Consistent with prior research (Durlak et al., 2011; Corcoran et al., 2018), our results demonstrate that implementation of Second Step with fidelity was associated with significant improvements in student outcomes. Most notably, students in treatment schools demonstrated small but statistically significant gains in ELA performance across two academic years. While effects on mathematics achievement were not significant, the ELA results align with broader literature suggesting SEL programs may be more tightly linked to literacy domains (Taylor et al., 2017). These findings support the theory that SEL competencies—such as emotion regulation, self-efficacy, and interpersonal communication—may facilitate students' ability to engage with academic content, particularly in literacy-rich disciplines.

In addition to academic gains, the intervention yielded robust effects on behavioral outcomes. Students in Second Step schools experienced significantly fewer ODRs, ISS, and OSS, with the magnitude of effects increasing during the second year of implementation. These findings highlight the cumulative and compounding benefits of sustained SEL exposure and suggest that the program may support more positive school climates through reductions in behavioral infractions. The behavioral findings are particularly compelling in the context of middle school, a developmental period marked by heightened behavioral risk (Akos, 2006; Onetti et al., 2019). That the program reduced not only minor referrals but also more serious infractions, such as OSS, further underscores its promise as a Tier 1 behavioral support.

Similarly, attendance improved among students in the treatment group, with gains equivalent to approximately 2.5 additional days per school year. These findings add to a limited but growing body of literature linking SEL interventions to improved school engagement (Jones et al., 2021). As chronic absenteeism continues to be a national concern—particularly in the wake of the COVID-19 pandemic—SEL programs like Second Step may serve as a viable pathway to improving instructional access and student connectedness.

Analyses of student-reported outcomes revealed significant improvements in perceived teacher–student relationships, school belonging, and overall school climate. These effects were moderate in magnitude and suggest that students in Second Step schools experienced a more affirming and supportive educational environment over time. Notably, no effects were found for self-management or supportive relationships, possibly reflecting either a ceiling effect or the relatively short window of program implementation. That said, the significant growth in relational constructs speaks to one of the central aims of SEL: fostering warm, inclusive, and respectful school communities.

Together, these findings offer several implications. First, universal SEL programs can positively impact not only students' emotional development but also concrete academic and behavioral outcomes when implemented with fidelity. Second, SEL appears to be particularly effective when delivered consistently over time, with stronger effects emerging in the second year of implementation. This suggests the importance of long-term investment and sustained programming rather than short-term or one-off efforts. Third, while the program improved ELA performance, it did not affect mathematics achievement. Future research should explore whether different academic domains respond differently to SEL exposure and whether content-specific adaptations could enhance academic impacts.

Limitations and Future Directions

Despite the study's strengths—including a large sample, rigorous quasi-experimental design with propensity score weighting, and the integration of multiple data sources—it is not without limitations. First, although weighting methods improved group comparability, unmeasured confounding variables cannot be fully ruled out. Future research should use randomized controlled trial designs to increase internal validity. Second, the lack of baseline individual academic performance data required reliance on school-level achievement proxies for equivalence testing in academic models. Third, generalizability is limited by the geographic and demographic characteristics of the sample, which may differ from other districts implementing Second Step.

Future research should explore the mediating mechanisms through which SEL programs influence academic and behavioral outcomes, such as changes in school climate, student motivation, or executive function. Moreover, studies should consider whether certain student subgroups benefit more from SEL interventions, particularly students with disabilities, multilingual learners, or those experiencing chronic absenteeism. Given the growing policy emphasis on whole-child education, further research is also needed to assess the cost-effectiveness, scalability, and sustainability of SEL programs in secondary settings.

Conclusion

The results of this study support the efficacy of the Second Step Middle School Program in promoting academic engagement, improving behavioral outcomes, and fostering more positive school experiences among adolescents. While middle school is often characterized by heightened risk and declining motivation, this work adds to a growing literature base indicating that universal SEL programs can serve as a protective factor—buffering students from developmental risks and creating conditions that support learning, belonging, and growth.

References

Akos, P. (2006). Middle school transition: Navigating change, keeping connected. *Middle School Journal*, *38*(1), 5–12. https://doi.org/10.1080/00940771.2006.11461527

Corcoran, R. P., Cheung, A. C. K., Kim, E., & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review*, *25*, 56–72. https://doi.org/10.1016/j.edurev.2017.12.001

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

Erentaitė, R., Vosylis, R., Gabrialavičiūtė, I., & Raižienė, S. (2018). How does school experience relate to adolescent identity formation over time? Cross-lagged associations between school engagement, school burnout and identity processing styles. *Journal of Youth and Adolescence*, *47*(4), 760-774.

Hornstra, L., Stroet, K., van Eijden, E., Goudsblom, J., & Roskamp, C. (2018). Teacher expectation effects on need-supportive teaching, student motivation, and engagement: A self-determination perspective. *Educational Research and Evaluation, 24*(3-5), 324-345. https://doi.org/10.1080/13803611.2018.1550841

Jones, S. M., Bailey, R., Brush, K., & Kahn, J. (2021). Navigating SEL from the inside out: Looking inside and across 33 leading SEL programs. Harvard Graduate School of Education, EASEL Lab. https://www.wallacefoundation.org/knowledge-center/pages/navigating-social-and-emotional-learning-from-the-inside-out.aspx

Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology, 50*(4), 443-460.

Onetti, W., Fernández-García, J. C., & Castillo-Rodríguez, A. (2019). Transition to middle school: Self-concept changes. *PloS One*, *14*(2), https://doi.org/10.1371/journal.pone.0212640

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

What Works Clearinghouse. (2024). *WWC standards handbook version 5.0*. U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v5.pdf

Table 1. Descriptive Statistics for Categorical Variables

		Comparison		Treatment	
Variable		***Frequency***	***Percentage***	***Frequency***	***Percentage***
Gender					
	Female	3,146	49.09	2,602	47.06
	Male	3,182	49.65	2,896	52.38
	Non-Binary	1	0.02	1	0.02
	Missing	80	1.25	30	0.54
Race					
	White	3,527	55.03	3,306	59.79
	Black	1,232	19.22	1,476	26.7
	Hispanic	1,227	19.14	339	6.13
	Asian	134	2.09	90	1.63
	Native American	9	0.14	8	0.14
	Native Hawaiian or Other Pacific Islander	14	0.22	24	0.43
	Multiracial	186	2.90	256	4.63
	Missing	80	1.25	30	0.54
Disability Status					
	Not SWD	5,529	86.27	4,632	83.78
	SWD	880	13.73	897	16.22

Table 2. Means, Standard Deviations, and Internal Consistency Estimates (Cronbach's α) for Composite Variables by Year.

Year	Scale	α	M	SD
Fall 2022				
	Teacher–Student Relationship	.868	3.60	0.95
	School Belonging	.875	3.83	0.99
	School Climate	—	3.44	1.17
	Self-Management	.842	4.03	0.62
	Supportive Relationships	.634	1.84	0.21
Fall 2023				
	Teacher–Student Relationship	.856	3.70	0.91
	School Belonging	.875	3.93	0.96
	School Climate	—	3.52	1.13
	Self-Management	.840	4.06	0.62
	Supportive Relationships	.635	1.86	0.20
Spring 2023				
	Teacher–Student Relationship	.873	3.49	0.97
	School Belonging	.887	3.68	1.04
	School Climate	—	3.26	1.19
	Self-Management	.858	3.99	0.65
	Supportive Relationships	.647	1.85	0.21
Spring 2024				
	Teacher–Student Relationship	.878	3.62	0.95
	School Belonging	.891	3.92	1.01
	School Climate	—	3.42	1.15
	Self-Management	.869	4.04	0.66
	Supportive Relationships	.687	1.86	0.21

Table 3. Means and Standard Deviations for Composite Variables by Year and Treatment Group

Year	Group	*TSR*	*SB*	*SC*	*SM*	*SR*
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Fall 2022	Control	3.59 (0.96)	3.80 (1.00)	3.45 (1.17)	4.07 (0.63)	1.84 (0.21)
	Treatment	3.62 (0.94)	3.88 (0.97)	3.42 (1.16)	3.99 (0.61)	1.85 (0.21)
Spring 2023	Control	3.52 (0.96)	3.73 (1.03)	3.33 (1.18)	4.01 (0.65)	1.86 (0.20)
	Treatment	3.44 (0.99)	3.57 (1.04)	3.12 (1.21)	3.97 (0.66)	1.84 (0.22)
Fall 2023	Control	3.68 (0.92)	3.91 (0.98)	3.52 (1.14)	4.06 (0.62)	1.86 (0.20)
	Treatment	3.75 (0.90)	3.97 (0.92)	3.54 (1.11)	4.07 (0.62)	1.86 (0.20)
Spring 2024	Control	3.60 (0.97)	3.91 (1.04)	3.46 (1.18)	4.03 (0.66)	1.86 (0.21)
	Treatment	3.66 (0.91)	3.93 (0.94)	3.34 (1.08)	4.05 (0.65)	1.87 (0.21)

Note. TSR = Teacher–Student Relationships, SB = School Belonging, SC = School Climate, SM = Self-Management, SR = Supportive Relationships.

Table 4. Establishing Equivalence on Panorama

Year	TSR	SB	SC	SM	SR
Fall 2022	0.03	0.08	-0.03	-0.13	0.05
Spring 2023	-0.08	-0.15	-0.18	-0.06	-0.10
Fall 2023	0.08	0.06	0.02	0.02	0.00
Spring 2024	0.06	0.02	-0.11	0.03	0.05

Note. All statistics are *d* effect sizes. TSR = Teacher–Student Relationships, SB = School Belonging, SC = School Climate, SM = Self-Management, SR = Supportive Relationships.

Table 5. Covariate Balance Table

Covariate	Type	Std. Mean Diff (SMD)
Grade 6	Binary	0.03
Grade 7	Binary	0.02
Grade 8	Binary	–0.04
Gender: Male	Binary	–0.04
Gender: Female	Binary	0.04
Gender: Other/Unknown	Binary	–0.00
Race: White	Binary	0.21
Race: Black	Binary	–0.05
Race: Hispanic/Latino	Binary	–0.16
Race: Asian	Binary	–0.01
Race: American Indian	Binary	–0.00
Race: Pacific Islander	Binary	0.00
Race: Multiracial	Binary	0.01
English Learner (ELL)	Binary	0.04
Student with Disability (SWD)	Binary	0.02
Teacher-Student Relationships	Continuous	0.03
School Belonging	Continuous	0.07
School Climate	Continuous	-0.02
Self-Management	Continuous	-0.08
Supportive Relationships	Continuous	0.05

Note. **Standardized mean differences (SMDs) above 0.10 in absolute value are typically interpreted as indicating meaningful imbalance. All values reported here are post-weighting.**

Table 6. School-Level English Language Arts and Mathematics Performance from the 2021-2022 School Year

Condition	School	ELA M (SD)	Math M (SD)
Treatment		47.3 (11.9)	37.1 (18.5)
	Middle School #1	55.7	50.1
	Middle School #2	38.9	24.0
Comparison		49.1 (40.0)	42.1 (43.6)
	Middle School #3	77.3	72.9
	Middle School #4	20.8	11.2

Note. Values are the percentage of students performing at the meets or exceeds level on the SC READY summative assessment.

Table 7. Weighted OLS Regression Results Predicting English Language Arts and Mathematics Scores, Adjusting for Prior Achievement and School Fixed Effects

Outcome	Predictor	Estimate	SE	t	p	Adj. R^2	RMSE	N
22–23 ELA	Intercept	77.07	0.71	108.88	< .001	.142	15.8	2,951
	Treat	2.16	0.51	4.27	< .001			
	Baseline ELA	0.14	0.01	12.25	< .001			
23–24 ELA	Intercept	76.54	0.73	104.29	< .001	.124	13.7	2,801
	Treat	2.34	0.49	4.78	< .001			
	Baseline ELA	0.14	0.01	12.15	< .001			
22–23 Math	Intercept	73.28	0.60	122.45	< .001	.110	15.6	2,937
	Treat	0.38	0.50	0.76	.447			
	Baseline Math	0.18	0.01	17.54	< .001			
23–24 Math	Intercept	69.44	0.60	114.88	< .001	.185	13.1	2,795
	Treat	0.64	0.47	1.36	.174			
	Baseline Math	0.25	0.01	24.61	< .001			

Note. **All models include school fixed effects (with two reference categories removed for collinearity) and apply analytic weights using the wt1 variable. RMSE = Root Mean Squared Error. Adj. R^2 = adjusted coefficient of determination.**

Table 8. Negative Binomial Regression Results Predicting Student Behavior Outcomes

Outcome	Predictor	Estimate	SE	*z*	*p*
Referrals	Intercept	1.25	0.08	15.69	< .001
	Treat	–0.44	0.07	–6.66	< .001
	Spring 2024	–0.12	0.07	–1.70	.089
	Treat × Spring 2024	–0.17	0.10	–1.64	.100
ISS	Intercept	–0.18	0.11	–1.66	.098
	Treat	–0.37	0.10	–3.73	< .001
	Spring 2024	–0.06	0.10	–0.55	.582
	Treat × Spring 2024	–0.45	0.15	–3.00	.003
OSS	Intercept	0.65	0.10	6.58	< .001
	Treat	–0.22	0.11	–2.10	.036
	Spring 2024	–0.19	0.11	–1.73	.084
	Treat × Spring 2024	–0.40	0.16	–2.51	.012

Note. ODR is office discipline referrals, ISS is in-school suspension, and OSS is out-of-school suspension. All models adjust for baseline English language arts performance and include fixed effects for school (reference categories omitted due to collinearity). Negative binomial regressions were weighted using propensity score inverse probability weighting.

Table 9. Negative Binomial Regression Results Predicting Number of Days Attended

Predictor	Estimate	SE	z	p
Intercept	5.091	0.00314	1622.18	< .001
Treat	0.01438	0.00243	5.93	< .001
Spring 2024	–0.00112	0.00259	–0.43	.666
ELA22 (baseline ELA)	0.00052	0.00005	11.14	< .001
Treat × Spring 2024	0.00208	0.00370	0.56	.575

Note. Model includes fixed effects for School ID (two dropped due to collinearity) and is weighted using propensity score inverse probability weighting.

Table 10. Difference-in-differences (DiD) Results Predicting School Climate and Well-Being Outcomes

Outcome	Treat Effect	Spring 2024	Treat × Spring 2024	RMSE	d
Teacher-Student Relationships (TSR)	1.12	–0.08	**0.22**	0.54	**0.40**
School Belonging (SB)	–0.66	0.01	**0.21**	0.55	**0.37**
School Climate (SC)	–0.98	0.02	*0.20*	0.73	**0.27**
Self-Management (SM)	0.08	0.01	0.04	0.31	0.12
Supportive Relationships	0.19	0.02	–0.01	0.20	–0.06

Note. **"Treat × Spring 2024" represents the DiD estimate—the additional change for the treatment group in Spring 2024 compared to the control group.** ***d*** **= (Interaction ÷ RMSE**

WestEd.org