# Evaluation of CharacterStrong's Secondary Social-Emotional and Character Education Curriculum: Effects on Student-Reported and Teacher-Reported Outcomes

Author order

James L. Merle[1], Madeline F. Larson[2], Clayton, R. Cook[2]

Completion Date

July 31st, 2023

Affiliations

[1]University of Utah Intermountain Healthcare Department of Population Health Sciences, Division of Health System Innovation and Research, Spencer Fox Eccles School of Medicine, University of Utah, Salt Lake City, UT, USA

[2]CharacterStrong, LLC, Pallayup, WA, USA

Abstract

Social-emotional learning (SEL) and character education are important components of adolescent development. In this study, we evaluated a randomized controlled trial of *CharacterStrong*, a curriculum that combines SEL and character education that included 1609 students and 242 teachers across 14 schools. This study applied baseline target moderation analysis to examine the differential effect of CharacterStrong given students ’baseline (pre-test) scores on self-reported measures of social-emotional competence and personal attributes (i.e., Self-efficacy, Grit, Learning strategies, Growth mindset, Self-management, and Social-awareness). Results indicated that students with the most room to improve at baseline (pre-test) who received CharacterStrong improved significantly more than those who also had the most room to improve in the control group for three out of the six measured outcomes: Self-efficacy, Learning strategies, and Self-management. Teachers who received training for CharacterStrong reported improvements in their own Self-efficacy and students ’SEL competencies compared to teachers in the control condition. This study provides evidence supporting the effects of *CharacterStrong* as well as adds to the school-based prevention research on performing moderation analyses of baseline status to reveal the sub-group of youth who respond most favorably to universal prevention programs.

*Keywords*: CharacterStrong, SEL, Character Education, baseline target moderation, schools, preventive interventions

## Introduction

Social-emotional learning (SEL) and character education (i.e., educational approaches that focus on developing the social and emotional skills of students) are considered vital aspects of students ’school experience and healthy development, particularly when delivered during times of acute stress and uncertainty such as during the COVID-19 pandemic, to prevent deleterious outcomes from occurring. Since SEL competencies develop and mature around early adolescence (Jones & Doolittle, 2017), there is a need for SEL programs that serve secondary students. Moreover, programs that specifically integrate character education as a complementary component to SEL have been tied to student achievement and behavioral outcomes (Jeynes, 2019). The purpose of this study was to conduct a cluster randomized controlled trial of the effects of *CharacterStrong*’s secondary curriculum.

### Importance of SEL for Adolescent Development

SEL has been recognized as influential in shaping the well-being of teenagers and fostering their healthy development, enabling individuals to cultivate valuable skills, virtues, and ethical behavior necessary for personal and social growth (Durlak et al., 2011). To date, there has been a proliferation of research demonstrating the effectiveness of SEL programs, although the majority of studies are conducted in the elementary grades (Durlak et al., 2022). Research consistently shows that when schools intentionally implement quality SEL curriculum and instruction, positive effects on a wide range of student outcomes follow, including academic engagement and performance, positive relationships, well-being, and longer-term follow-up effects (Corcoran et al., 2018). Because of its effectiveness and potential for widespread reach, schools nationwide implement SEL programs, and state educational agencies have developed policies calling for widespread SEL program use (Eklund et al., 2018; Taylor et al., 2017).

Research on SEL effectiveness at the middle and high school level is promising but limited to date, with evidence to suggest that SEL can have significant positive impacts on middle and high school students' social-emotional skills, academic outcomes, behavior, and school climate (Espelage et al., 2015; Taylor et al., 2017). By promoting self-control, empathy, and responsible decision-making, SEL programs contribute to a decrease in risky behaviors, substance abuse, aggression, and delinquency in adolescents, as well as higher grades, increased motivation, and better classroom behavior (Sklad et al., 2012). In addition to their effect on student-level outcomes, SEL interventions in secondary schools contribute to the development of a positive school climate characterized by respectful and caring relationships. These programs foster a sense of belonging, inclusivity, and connectedness among students and between students and staff (Payton et al., 2008).

**Integrating Character Education into the SEL Framework**

Although SEL is important for adolescent development, the integration of SEL and character education offers an important opportunity for even greater tailoring to the unique developmental tasks of adolescence, an especially critical stage in moral, social, and personal development linked to successful engagement in the social, professional, and personal tasks of adulthood (Zhang et al., 2023). The challenges and opportunities specific to adolescence (e.g., identity formation, moral development, peer influence, increased autonomy, and the transition to adulthood) make integrating SEL and character education particularly important for adolescents (Beets et al., 2009; Jeynes, 2019). By addressing these specific needs, expanding SEL to include character education could help adolescents further prepare for the complexities and responsibilities of adulthood.

Character education and SEL overlap in their goals of fostering positive personal and social development. SEL focuses on teaching students skills to understand and manage their emotions, develop positive relationships, make responsible decisions, and effectively handle challenges and conflicts to ultimately enhance students' overall social and emotional well-being and improve their academic performance (Durlak et al., 2011). Character education emphasizes students ’development of virtues (e.g., honesty, respect, responsibility, and empathy), ethical behavior, and moral reasoning to cultivate moral character, personal qualities, and responsible citizenship (Berkowitz & Bier, 2007). Given their complementary differences, character education integrated into an SEL approach holds particular promise for adolescents.

Research indicates that delivery of character education in secondary school is associated with positive outcomes, including: increased emotional wellbeing, positive interpersonal relationships, and better self-control, empathy, and conflict resolution skills (Jeynes, 2019); improved academic achievement, including higher grades and standardized test scores (Katilmis et al., 2011); and reduced aggressive behavior, improved social skills, and increased prosocial behavior(Berkowitz & Bier, 2005; Thapa et al., 2012). Moreover, the delivery of character education has been associated with long-term benefits for students, including higher educational attainment, career success, and overall well-being in adulthood when compared to students who do not receive character education (Berkowitz et al., 2017).

**CharacterStrong Middle School & High School Curricula**

*CharacterStrong Secondary* curriculum (hereafter referred to as *CharacterStrong)* is a universal program designed for all students in a given secondary school (i.e., 6th through 12th grades). The curriculum includes vertically-aligned content across each grade that integrates traditional SEL skills-focused instruction (e.g., emotion regulation and interpersonal skills) with

character education for secondary students. *CharacterStrong* is designed to build social and emotional competence, develop character, and cultivate strong educator-student relationships. *CharacterStrong* Middle School curriculum consists of 35 lessons per grade level for 6th through 8th grades. *CharacterStrong* High School curriculum consists of 25 lessons per grade level for 9th through 12th grades.

Each lesson is designed to be delivered with minimal preparation needed and few supplies to minimize implementation barriers, as well as with a strong focus on student voice and choice to maximize student responsiveness. Lessons are delivered in approximately 30 minutes with the flexibility to adapt to a variety of timing needs. Each lesson features an evidence-based approach to structure and delivery grounded in community-building to develop psychological safety and trust among peers and teacher, explicit instruction, practice and application. Lessons are delivered via five core segments: *Warm Welcome*, a *Community-Building activity*, interactive *Content*, a *Character Challenge* to apply learning in their real lives, and an *Optimistic Closure*. These segments allow for flexible delivery and consistent structure to promote routine implementation across a range of schedules and contextual needs.

Grounded in evidence on healthy youth development (Bernat & Resnick, 2006), *CharacterStrong* was constructed around five primary principles: Emotion Understanding & Regulation, Empathy & Compassion, Values & Purpose, Goals & Habits, and Leadership & Teamwork. Each of these focuses on building specific SEL and character competencies that support three proximal outcomes of well-being, belonging, and engagement. These outcomes serve as enablers for success in both school and life.

**Gaps in the Literature**

It is well-documented that not all participants benefit equally from evidence-based prevention programs (Gardner, 2023). A common problem is a 'rich-get-richer 'effect, where participants with fewer risks improve more than those with greater risks. Prevention researchers suggest that the effectiveness of a universal program likely varies according to children's baseline status (Greenberg & Abenavoli, 2017; Howe, 2019). Children with higher needs at baseline are likely to be more responsive to universal programs than peers who are higher in strengths and low on difficulties at baseline (i.e., demonstrating a larger or steeper change over time)(Calhoun et al., 2020). Thus, it is important to extend main effect analyses by exploring if the treatment effect varies across subgroups based on baseline status on outcomes as well as child- and teacher-level moderators (e.g., demographics, fidelity; Spybrook et al., 2020). Baseline target moderation (BTM) analysis, examining whether intervention effects are moderated by starting levels, can reveal for whom interventions do and do not work (Gardner, 2023; Howe, 2019).

**Purpose of this Study**

The overarching purpose of this study was to add to the growing literature on SEL and Character Education in secondary schools. This study provided a unique opportunity to examine the effectiveness of *CharacterStrong,* which is an integrated SEL and Character Education program. More specifically, the aim of this study was to conduct a cluster randomized controlled trial of the differential effects of the *CharacterStrong* on student- and teacher-reported outcomes using baseline target moderation analysis. Specific research questions included:

- Were there significant differences between treatment and control groups in student-level outcomes (i.e., self-efficacy, grit, learning strategies, growth mindset, self-management, and social awareness) between pre-post- assessment?

- Did students who received *CharacterStrong* and started lower on outcomes at pre-assessment significantly improve at post-assessment more so than students in the Control condition who also started in the bottom third of the sample on outcomes?
- Were there significant differences between treatment and control groups in teacher-level outcomes (i.e., social-emotional learning competencies, self-efficacy, and enthusiasm,) between pre- post-assessment?

## Method

### Study Context

While research on integrated SEL and character education effectiveness at the secondary level is important in its own rite, this study is particularly unique as it was carried out during school closures caused by the onset of the COVID-19 pandemic in Spring 2020. Study recruitment, *CharacterStrong* professional development, and lesson implementation began prior to the onset of the pandemic. Following initial implementation, *CharacterStrong* facilitated a massive change of teaching modality in the participating schools to promote continuity of care within the study samples and implementation continued in a virtual environment into the following year when students returned to in-person instruction at school. This study is also unique in that it evaluated the longer-term reach of the impact of *CharacterStrong* across two full academic years and in the context of the pandemic.

### Participants

As described in the Design section below, we initially recruited a total of 14 secondary schools ($n_{student}$=2097, $n_{teacher}$=282) from a large and diverse urban school district in the Pacific Northwest. In an effort to optimize the accuracy and transparency in reporting, we present the descriptive statistics for both the initial sample and the final analytic sample used to evaluate the

effectiveness of the *CharacterStrong* (i.e., remaining cases after list-wise deletion of all the cases missing on pretest outcomes).

**Procedures**

Schools were randomly assigned to early start and delayed start conditions. Schools in both conditions eventually received training to implement *CharacterStrong*, with the early start schools starting delivery on *CharacterStrong* during the 2019-2020 and the delayed start condition starting delivery during the 2020-2021 school year. This design allowed for an intervention control comparison using data from the 2019-2020 school year and then the comparison of dosage/exposure using the 2020-2021 with both conditions implementing *CharacterStrong* but students in the early start condition receiving more exposure.

We followed the CONSORT reporting guideline for cluster trials (Figure 1).

**Measures**

***Student & Teacher Demographics***

A set of student and teacher demographic variables were added to the survey data (Table 1), which are theoretically relevant to the research questions and commonly used in educational or implementation research. The demographic variables were obtained via either administrative data reported by the participating schools or the self-report data from the first section of the web-based survey. To eliminate potential confounding effects and optimize statistical power, the demographic covariates were included in the analytic models if statistical equivalence between intervention and control groups was not established at baseline (pre-test).

***Panorama Social-Emotional Learning Surveys (P-SELS)***

The P-SELS is a suite of validated scales developed by Panorama Education that assesses the outcomes associated with implementation of Social-Emotional Learning curricula. The suite

captures the change in student self-reports on personal attributes, student perceptions of the school environment and supports, and teachers ’perspectives and skills, as well as three domains of student SEL outcomes: social relationships (Gehlbach & Robinson, 2016; Panorama Education, 2020) motivation (Hulleman & Harackiewicz, 2009), and self-regulation (Duckworth et al., 2011). In a large-scale validation study produced acceptable reliability and validity evidence of the P-SELS scales used in this study (Panorama Education, 2020).

As part of their routine screening procedures within the school or district context, participating school systems assessed students' reports of their personal attributes before and after the implementation (pre-/post-test) of the CharacterStrong using seven scales from the P-SELS: Grit, Learning Strategies, Self-Efficacy, Self-Management, Social Awareness, and Growth Mindset (Panorama Education, 2020). Teacher-report of outcomes was collected with P-SELS subscales, including their own enthusiasm and self-efficacy, and student SEL competency. The pre-test was completed between October 1 and November 8, 2019, while the post-test was completed between May 25 and Jun 19, 2020. In this study sample, the reliability estimates for each scale ranged from Acceptable to Good ($\alpha$=.72–.85)(Cronbach, 1951).

**Analysis Plan**

Prior to analysis, baseline equivalence tests (paired-sample *t*-tests) were conducted to examine differences in student and teacher demographics and baseline levels of response variables across the treatment and control groups. The student-level control group had a significantly higher proportion of students eligible for free or reduced-price lunch [FRPL; $t(1163)$=–2.03, $p$=.04]. Thus, FRPL status was entered as a student-level covariate (Level 1) in all inferential models. The two groups were not significantly different on any other variables at

*a*=.05 level. Missing data accounted for <1% of all observed variables, and data were assumed to be missing at random.

Intraclass correlation coefficients (ICCs) were calculated to assess design effects (i.e., whether student or teacher scores were nested within schools; (Raudenbush & Bryk, 2002). Student score ICCs ranged from .002–.02 and Teacher score ICCs ranged from .02–.15. Although some of the ICCs suggest clustering may be ignorable, some exceeded 0.05, suggesting a need for clustering. Given these findings and that the study employed a cluster randomization at the school level, we used mixed-linear modeling, including school-level clustering for all analyses, testing the intervention at the unit of randomization. Reduced maximum likelihood (REML) was used in reported results, and full maximum likelihood (ML) was used to compare the fit of nested models. All analyses were run in *R* (R Core Team, 2021).

To answer Research Question 1, intervention effects were tested using treatment group as the independent variable predicting the change in outcome from pre-post, while controlling for FRPL. The dependent variable was modeled as the difference in pre-post test score (i.e., post *minus* pre) (Model 1). Next, the following predictors were included to determine whether changes in outcomes were moderated by the following level-1 covariates: student grade, race (dummy coded across all categories with white as the reference), gender, and special education status (Model 2).

To answer Research Question 2, a cross-level interaction was conducted between Treatment group and pre-assessment score, predicting the difference in pre-post assessment. Any significant predictors from model 2 were also included, and if still significant, were also included as an interaction term. Model fit indices (AIC, deviance test, $X^2$ tests) were used to identify the

best-fitting models. Grand mean centering was used for main effects variables involved in the interaction.

To answer Research Question 3, the predictor for treatment group was included as a dependent variable predicting the change in outcome variable from pre-post. Next, the following predictors were included to determine whether changes in teacher outcomes were moderated by the following level-1 covariates: teacher years of experience and gender.

Effect sizes appropriate for mixed-linear modeling were calculated using a delta score, which accounts for relative change over time between the intervention and control group *delta* = $(M_{\text{change–T}} / SD_{\text{pre–T}}) - (M_{\text{change–C}} / SD_{\text{pre–C}})$(Feingold, 2009). This effect size can be interpreted similar to Cohen's *d* (Cohen, 1988).

## Results

For all pre-outcome data, n=1780, and for post-outcome data, n=1884. The full sample with no NAs removed was N=2097, and n=1609 after list-wise deletion. A total of 1609 students and 242 teachers were included for modeling at level 1 across 14 schools at level 2. Student and teacher demographics are displayed in Table 1.

### Research Question 1.

The main effects, baseline model for student outcomes included FRPL as a predictor, given that baseline un-equivalency was found between treatment and control groups. There were no significant treatment effects for any measured outcome. FRPL (students who paid for lunch) was a significant predictor of improvements in Social-awareness ($\beta$=-0.09, $SD$=.03, $p$<.001) and Learning strategies ($\beta$=-0.11, $SD$=.04, $p$=.002) from Pre-Post assessment (Table 2). A treatment by FRPL interaction was run for each model, though none were significant at α=.05.

Table 3 includes conditional models for each student-level outcome, where grade, race, sex, and whether the student was receiving special education services were included as covariates. Most models were associated with worse or non-significant change in fit except for Social-awareness and Self-management. The conditional model for Social-awareness improved fit and indicated that students who paid for lunch ($\beta$=-0.10, $SD$=.03, $p$<.001), students in higher grades ($\beta$=0.04, $SD$=.01, $p$<.001), and students *not* receiving Special education services ($\beta$=-0.14, $SD$=.05, $p$=.01) were associated with statistically significant positive change from pre-post-assessment. The conditional model for Self-management had improved fit compared to the unconditional model and indicated that students in higher grades ($\beta$=0.03, $SD$=.01, $p$=.011) were associated with statistically significant improvement from pre-post-assessment.

**Research Question 2.**

Table 4 includes the baseline targeted moderation interaction models between treatment group pre-score for each outcome (Model 3). The deviance test ($X^2$) comparison was conducted between the better-fitting prior model (Model 1 or Model 2). If significant predictors were found in Model 2, they were included in the interaction, which was the case for Learning strategies, Self-management, and Social-awareness. For Learning strategies, FRPL was no longer significant when accounting for pre-assessment score, so it was dropped from the final model. For Self-management, student grade was also no longer significant when accounting for pre-assessment score, so it was also dropped. The best-fitting model for Social-awareness involved no interactions, with pre-post score change moderated by *higher* student grade ($\beta$=0.04, $SD$=.01, $p$=.002), students *not* receiving SPED services ($\beta$=-0.16, $SD$=.04, $p$<.001), and students who scored lower on Social awareness at pre-assessment ($\beta$=-0.45, $SD$=.02, $p$<.001).

In the case of non-significant deviance tests between models, the most parsimonious model was kept (Table 4).

Results indicate that students who were in the treatment group and had the largest room for improvement at baseline improved significantly compared to students who were in the control group and had the largest room for improvement for Self-efficacy ($\beta$=-0.11, *SD*=.04, *p*=.01), Learning Strategies. ($\beta$=-0.10, *SD*=.04, *p*=.02), and Self-management ($\beta$=-0.08, *SD*=.04, *p*=.03). To visualize these interactions, the student sample was split into tertiles (low, middle, high) according to each students 'baseline score across outcomes. Students with the most room for improvement at baseline for each outcome (bottom third of each pre-score) are compared between treatment groups and displayed visually in Figure 2.

**Research Question 3**

Table 5 displays the teacher-level models. Teachers in the treatment group significantly improved their reported self-efficacy ($\beta$=0.15, *SD*=.07, *p*=.04). When years of experience and gender were added to the model, the effect of treatment fell out of significance, and years of experience was significant ($\beta$=0.01, *SD*=.004, *p*=.03), indicating that teachers with more years of experience were associated with positive changes in self-efficacy from pre-post-test regardless of treatment condition. However, this model did not significantly improve fit $X^2$=5.19, *p*=.07. Teachers in the treatment group also indicated that their students had significant improvements in their SEL competencies ($\beta$ =0.24, *SD*=.09, *p*=.03). In the best fitting model, gender ($\beta$ =0.10, *SD*=.05, *p*=.04) and years of experience ($\beta$ =0.01, *SD*=.00, *p*=.03) were also significant. Positive $\beta$ values indicate that teachers in the treatment group who were male and had more years of experience were associated with increases in social-emotional learning competencies from pre-post assessment. An interaction model was run, though no terms were significant.

### Effect Sizes

Effect size estimates were calculated for each student- and teacher-level outcome. All but one student-level effect size were in the direction favoring the treatment group. The highest effect size was associated with Learning strategies *(d*=0.10) followed by Self-management (*d*=0.06), Growth mindset *(d*=0.04), Grit (*d*=0.03), and Self-efficacy (*d*=.001), while Social awareness *(d*=-0.07) favored the control condition. Teacher-level effect sizes were all relatively larger than student effect sizes and favoring the treatment group. The highest teacher-level effect size was for Enthusiasm (*d*=0.51), followed by SEL Competencies (*d*=0.40), and Self-efficacy (*d*=0.18).

## Discussion

This study experimentally examined the effects of the *CharacterStrong* on student- and teacher-level outcomes and investigated its differential effect on students with high needs at baseline (pre-test) using baseline target moderation analysis. In general, this study provided evidence supporting the effectiveness of *CharacterStrong*, in particular for the sub-group of students with the greatest need and room for support at baseline. Several findings emerged from this study that are worthy of discussion, and that add to the growing body of experimental research on SEL and Character Education at the secondary grades.

### Study Findings

Consistent with prior research evaluating the effects of universal prevention programs, no significant main effects were found for *CharacterStrong* on student self-reported outcomes (Smith et al., 2023). This finding was anticipated because main effect analyses assume that all students have room for improvement at baseline. The most sensitive analyses of the effects of universal prevention programs, like *CharacterStrong*, are for those that disaggregate data to

identify effectiveness for those students who have the greatest need and room for growth at baseline (Howe & Leijten, 2023). In line with this literature, significant effects were found for students who had the greatest need and room for growth at baseline on three out of the six student-self-reported outcomes (Self-efficacy, Learning strategies, and Self-management) relative to comparable students in the control group. These findings were notable considering the importance of these constructs as academic enablers to students ’school success, especially in the secondary grades where greater independence and self-sufficiency are needed (Strunk, 2014). Results extend school-based prevention research for adolescents by highlighting the sub-group of students who may be in most need of universal prevention programming .

Significant effects were also found on two of the teacher-reported outcomes. Specifically, teachers who received training for *CharacterStrong* reported higher levels of self-efficacy, suggesting that as a result of delivering *CharacterStrong* teachers were more confident in their ability to meet the needs of the students they serve. This is a noteworthy finding considering the literature on correlates of teacher self-efficacy, such as job satisfaction (Ortan et al., 2021) and retention (Ahn et al., 2021) as well as student outcomes such as academic engagement and performance (Sökmen, 2021). Findings also revealed a significant effect favoring *CharacterStrong* on teachers ’ratings of student SEL competencies. The SEL competencies captured in the survey include growth mindset, self-efficacy, and social awareness, which have been linked to short-term educational outcomes and long-term outcomes into adulthood (Ross & Tolan, 2018).

### *Effect size interpretations for Universal Prevention Programs*

In a review of meta-analyses of universal prevention programs for school-aged youth (ages 5–18), median effect sizes tended to be lower and more variable than typical benchmark

effect size interpretations (Tanner-Smith et al., 2018). Authors recommend that previously identified benchmarks for effect sizes (0.20, 0.50, and 0.80 being indicative of small, medium, and large effects) are not appropriate for interpreting the magnitude of effect for universal prevention studies for school-age youth, finding that median effect sizes tended to fall between 0.07–0.16. Instead, effect sizes should be contextualized to aid more accurate interpretation. In this study, effect sizes at the student-level fell between –0.07 and 0.10, and teacher-level effect sizes fell between 0.18 and 0.51. These findings align with previous evaluations of preventive interventions and indicate that *CharacterStrong* had moderate practical effects on student outcomes and moderate-to-large effects on teacher outcomes.

**Limitations and Future Directions**

As with all studies, this study has limitations that are important to note when interpreting findings and pinpoint directions for future research. The first limitation is that of the measures used in this study as all findings are related to the methods used and outcomes assessed. While this study included both student self-report and teacher report to provide a multi-informant approach. Findings would be strengthened if other methods were used such as direct observation of student behavior or school administrative records of outcomes like behavior, attendance, and grades. Future research on the effects of *CharacterStrong* or other secondary SEL programs should include a multi-method, multi-informant approach. There are also limitations with regard to the sample of schools, students, and teachers included in this study, which impacts the generalizability of this study's findings. In this vein, two other RCTs have been completed on *CharacterStrong* evaluating it in schools operating in other geographic contexts serving students with different demographic characteristics. Other research groups should seek to replicate the findings to evaluate whether findings generalize beyond the sample in this study.

Another key limitation was associated with COVID-19. Originally, the post assessment was scheduled for April 2020, but with the school closure and the school system's focus on transitioning students to remote learning, the data collection effort was not sufficiently rolled out and accessed by all students. For this reason, the post assessment was moved to the Fall of the next academic year prior to a time when control schools initiated the implementation of *CharacterStrong*. While this is a limitation, this study still provided a test of the effects of *CharacterStrong* relative to schools, teachers and students in a control group that had not yet initiated the implementation of *CharacterStrong*.

Another limitation is the lack of focus on implementation. Future research is warranted that examines the implementation process more closely and examines the degree to which implementation outcomes such as fidelity and reach are associated with student outcomes. There are also opportunities in future research for hybrid effectiveness-implementation trials that continue to build on the evidence of *CharacterStrong* that simultaneously assess the effectiveness of *CharacterStrong* as well as the effects of implementation strategies that improve the uptake and use of *CharacterStrong* (Landes et al., 2019).

**Implications for Prevention Science**

There are several implications for Prevention Science stemming from this study. First, this study suggests the need for subgroup analyses of the effects of universal prevention programs consistent with recommendations from Prevention Science (Howe & Leijten, 2023), especially according to the baseline status of students. Second, there are implications for prevention science focused on evaluating the unique and combined effects of SEL and Character Education for secondary-aged students. To date, a treatment component analysis study does not exist, and there is potential that combining approaches may offer additive value beyond either

approach alone. Last, there are implications for prevention researchers to continue to focus on SEL in the secondary grades as it is a unique developmental context in which prevention efforts are timely and needed.

**Conclusion**

*CharacterStrong* offers promise as a universal prevention program in secondary schools. While more research is warranted, the findings from this study provide evidentiary support for *CharacterStrong* to improve students ’perceptions of their self-efficacy, learning strategies, and self-management as well as teacher reports of their own self-efficacy and students ’SEL competencies. This research is important in light of the literature indicating few programs are currently well suited for the unique implementation and developmental needs of secondary students and schools

**Compliance with Ethical Standards**

Funding: This work was supported in part by The National Institutes of Health, NLM 2 T15 LM 007124-26 postdoctoral training slot to JLM.

Ethics approval: This study was completed in compliance with district Institutional Review Boards as well as district, state, and federal data privacy policies.

Conflicts of interest/Competing interests: CRC and MFL are employed by CharacterStrong, LLC. The study was completed in collaboration with an external research entity to ensure unbiased results and reporting.

Consent to participate: This study was completed as a secondary data analysis project in collaboration with school districts. Thus, participating schools and districts adhered to their own consent/assent policies for routine screening for student social-emotional needs that may require intervention. Researchers were not responsible for assent/consent in this study.

## References


Ahn, I., Chiu, M. M., & Patrick, H. (2021). Connecting teacher and student motivation: Student-perceived teacher need-supportive practices and student need satisfaction. *Contemporary Educational Psychology*, *64*, 101950.

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K. K., Burns, K., Washburn, I. J., & Durlak, J. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American journal of public health*, *99*(8), 1438-1445. https://doi.org/10.2105/ajph.2008.142919

Berkowitz, M. W., & Bier, M. C. (2005). What works in character education: A research-driven guide for educators. In: Washington, DC: Character Education Partnership.

Berkowitz, M. W., & Bier, M. C. (2007). What works in character education. *Journal of Research in Character Education*, *5*(1), 29-48.

Berkowitz, M. W., Bier, M. C., & McCauley, B. (2017). Toward a science of character education. *Journal of Character Education*, *13*(1), 33-51.

Bernat, D. H., & Resnick, M. D. (2006). Healthy Youth Development: Science and Strategies. *Journal of Public Health Management and Practice*, *12*. https://journals.lww.com/jphmp/Fulltext/2006/11001/Healthy_Youth_Development__Science_and_Strategies.4.aspx

Calhoun, B., Williams, J., Greenberg, M., Domitrovich, C., Russell, M. A., & Fishbein, D. H. (2020). Social Emotional Learning Program Boosts Early Social and Behavioral Skills in Low-Income Urban Children. *Front Psychol*, *11*, 561196. https://doi.org/10.3389/fpsyg.2020.561196

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/https://doi.org/10.1016/j.edurev.2017.12.001

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social psychological and personality science*, *2*(2), 174-181.

Durlak, J. A., Mahoney, J. L., & Boyle, A. E. (2022). 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*(11-12), 765-782. https://doi.org/10.1037/bul0000383

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 Dev*, *82*(1), 405-432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Eklund, K., Kilpatrick, K. D., Kilgus, S. P., & Haider, A. (2018). A Systematic Review of State-Level Social–Emotional Learning Standards: Implications for Practice and Research. *School Psychology Review*, *47*(3), 316-326. https://doi.org/10.17105/SPR-2017.0116.V47-3

Espelage, D. L., Rose, C. A., & Polanin, J. R. (2015). Social-emotional learning program to reduce bullying, fighting, and victimization among middle school students with disabilities. *Remedial and Special Education*, *36*(5), 299-311. https://doi.org/10.1177/0741932514564564

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychol Methods*, *14*(1), 43-53. https://doi.org/10.1037/a0014699

Gardner, F. (2023). Commentary for Special Issue on Using Baseline Target Moderation to Assess Variation in Prevention Impact: When (and How) to Revise Our Programs. *Prevention Science*, *24*(2), 299-303. https://doi.org/10.1007/s11121-022-01458-1

Gehlbach, H., & Robinson, C. D. (2016). Commentary: The Foundational Role of Teacher–Student Relationships. In *Handbook of Social Influences in School Contexts* (pp. 230-238). Routledge.

Greenberg, M. T., & Abenavoli, R. (2017). Universal Interventions: Fully Exploring Their Impacts and Potential to Produce Population-Level Impacts. *Journal of Research on Educational Effectiveness*, *10*(1), 40-67. https://doi.org/10.1080/19345747.2016.1246632

Howe, G., & Leijten, P. (2023). When Is It Time to Revise or Adapt Our Prevention Programs? Introduction to Special Issue on Using Baseline Target Moderation to Assess Variation in Prevention Impact. *Prev Sci*, *24*(2), 199-203. https://doi.org/10.1007/s11121-022-01456-3

Howe, G. W. (2019). Using baseline target moderation to guide decisions on adapting prevention programs. *Dev Psychopathol*, *31*(5), 1777-1788. https://doi.org/10.1017/s0954579419001044

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410-1412.

Jeynes, W. H. (2019). A Meta-Analysis on the Relationship Between Character Education and Student Achievement and Behavioral Outcomes. *Education and Urban Society*, *51*(1), 33-71. https://doi.org/10.1177/0013124517747681

Jones, S. M., & Doolittle, E. J. (2017). Social and emotional learning: Introducing the issue. *The Future of Children*, 3-11.

Katilmis, A., Eksi, H., & Ozturk, C. (2011). Efficiency of Social Studies Integrated Character Education Program. *Educational Sciences: Theory and Practice*, *11*(2), 854-859.

Landes, S. J., McBain, S. A., & Curran, G. M. (2019). An introduction to effectiveness-implementation hybrid designs. *Psychiatry Research*, *280*, 112513. https://doi.org/https://doi.org/10.1016/j.psychres.2019.112513

Ortan, F., Simut, C., & Simut, R. (2021). Self-Efficacy, Job Satisfaction and Teacher Well-Being in the K-12 Educational System. *Int J Environ Res Public Health*, *18*(23). https://doi.org/10.3390/ijerph182312763

Panorama Education. (2020). Validity brief: Panorama student survey. *Panorama Education*.

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). The Positive Impact of Social and Emotional Learning for Kindergarten to Eighth-Grade Students: Findings from Three Scientific Reviews. Technical Report. *Collaborative for Academic, Social, and Emotional Learning (NJ1)*.

R Core Team, R. (2021). R: A language and environment for statistical computing. In (Vol. Version 4.1.1): R foundation for statistical computing Vienna, Austria.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.

Ross, K. M., & Tolan, P. (2018). Social and emotional learning in adolescence: Testing the CASEL model in a normative sample. *The Journal of Early Adolescence*, *38*(8), 1170-1199.

Sklad, M., Diekstra, R., Ritter, M. d., Ben, J., & Gravesteijn, 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.

Smith, J. D., Carroll, A. J., Fu, E., & Berkel, C. (2023). Baseline Targeted Moderation in a Trial of the Family Check-Up 4 Health: Potential Explanations for Finding Few Practical Effects. *Prev Sci*, *24*(2), 226-236. https://doi.org/10.1007/s11121-021-01266-z

Sökmen, Y. (2021). The role of self-efficacy in the relationship between the learning environment and student engagement. *Educational Studies*, *47*(1), 19-37.

Spybrook, J., Zhang, Q., Kelcey, B., & Dong, N. (2020). Learning From Cluster Randomized Trials in Education: An Assessment of the Capacity of Studies to Determine What Works, For Whom, and Under What Conditions. *Educational Evaluation and Policy Analysis*, *42*(3), 354-374. https://doi.org/10.3102/0162373720929018

Strunk, T. A. (2014). An exploration of the relationships between academic enablers and middle school achievement.

Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically Based Mean Effect Size Distributions for Universal Prevention Programs Targeting School-Aged Youth: A Review of Meta-Analyses. *Prev Sci*, *19*(8), 1091-1101. https://doi.org/10.1007/s11121-018-0942-1

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 Dev*, *88*(4), 1156-1171. https://doi.org/10.1111/cdev.12864

Thapa, A., Cohen, J., Higgins-D'Alessandro, A., & Guffey, S. (2012). School Climate Research Summary: August 2012. School Climate Brief, Number 3. *National School Climate Center*.

Zhang, Y., Cook, C. R., & Smith, B. (2023). PurposeFull People SEL and Character Education Program: A Cluster Randomized Trial in Schools Implementing Tier 1 PBIS with Fidelity. *School Mental Health*. https://doi.org/10.1007/s12310-023-09600-2

Table 1.

*Participant Demographic Characteristics*

| | **Students** | | | **Teachers** | | |
|---|---|---|---|---|---|---|
| **Demographic** | **Combined** *N (%)* | **Treatment** *n (%)* | **Control** *n (%)* | **Combined** *N (%)* | **Treatment** *n (%)* | **Control** *n (%)* |
| Total | 2097 (100%) | 771 (37%) | 1326 (63%) | 282 (100%) | 120 (43%) | 162 (57%) |
| Sex | | | | | | |
| Female | 1141 (54%) | 421 (55%) | 720 (54%) | 159 (56%) | 69 (58%) | 90 (56%) |
| Male | 956 (46%) | 350 (455) | 606 (46%) | 123 (44%) | 51 (42%) | 72 (44%) |
| Race | | | | | | |
| Asian | 61 (3%) | 10 (1%) | 51 (4%) | – | – | – |
| Black | 56 (3%) | 17 (2%) | 39 (3%) | – | – | – |
| Hispanic | 218 (10%) | 90 (12%) | 128 (10%) | – | – | – |
| Native American | 18 (1%) | 7 (1%) | 11 (1%) | – | – | – |
| Pacific Islander | 31 (2%) | 11 (1%) | 20 (2%) | – | – | – |
| White | 1447 (69%) | 540 (70%) | 907 (68%) | – | – | – |
| Multi-racial | 266 (13%) | 96 (12%) | 170 (13%) | – | – | – |
| Grade Level | | | | | | |
| 6 | 84 (4%) | 49 (6%) | 35 (3%) | – | – | – |
| 7 | 1119 (53%) | 462 (60%) | 657 (50%) | – | – | – |
| 9 | 403 (19%) | 89 (12%) | 314 (24%) | – | – | – |
| 10 | 287 (14%) | 105 (14%) | 182 (14%) | – | – | – |
| 11 | 204 (10%) | 66 (9%) | 138 (10%) | – | – | – |
| Free or Reduced Priced Lunch status | | | | | | |
| Paid | 1039 (50%) | 355 (46%) | 684 (52%) | – | – | – |
| Reduced | 239 (11%) | 101 (13%) | 138 (10%) | – | – | – |
| Free | 819 (40%) | 315 (41%) | 504 (38%) | – | – | – |
| Special Education | | | | | | |
| Yes | 174 (8%) | 61 (8%) | 113 (9%) | – | – | – |
| No | 1932 (92%) | 710 (92%) | 1213 (91%) | – | – | – |
| ELD status | | | | | | |
| Yes | 90 (4%) | 24 (3%) | 66 (5%) | – | – | – |
| No | 2007 (96%) | 747 (97%) | 1260 (95%) | – | – | – |
| Grade Taught | | | | | | |
| Elementary | – | – | – | 0 (0%) | 0 (0%) | 0 (0%) |
| Middle School | – | – | – | 104 (37%) | 50 (42%) | 54 (33%) |
| High School | – | – | – | 127 (45%) | 64 (53%) | 63 (39%) |
| EL or MS | – | – | – | 4 (1%) | 4 (1%) | 0 (0%) |
| MS or HS | – | – | – | 46 (16%) | 1 (<1%) | 45 (28%) |
| Option School | – | – | – | 1 (<1%) | 1 (<1%) | 0 (0%) |
| | | | | *M (SD)* | *M (SD)* | *M (SD)* |
| Years of Experience | – | – | – | 12.4 (8.6) | 13.4 (8.8) | 11.6 (8.4) |

Table 2.

*Student-level Unconditional Intervention Effects Model for Each Measured Outcome*

| **Parameter** | **Outcome** | | | | | |
|---|---|---|---|---|---|---|
| | *Self-Efficacy* | *Learning Strategies* | *Growth Mindset* | *Self-management* | *Social-awareness* | *Grit* |
| | Fixed Effects *β (SD)* | | | | | |
| **Intercept (Post-Pre)** | -0.16 (.03)*** | -0.05 (.04) | -0.11 (.05) | -0.07 (.06) | 0.06 (.03) | -0.03 (.03) |
| **Level 1 (student)** | | | | | | |
| FRL (1=Eligible) | -0.01 (.04) | -0.11 (.04)** | -0.00 (.05) | -0.03 (.03) | -0.09 (.03)*** | -0.01 (.03) |
| **Level 2 (school)** | | | | | | |
| Treatment Group (1=CharacterStrong) | -0.00 (.04) | 0.10 (.05) | 0.07 (.07) | 0.06 (.05) | 0.01 (.05) | -0.01 (.04) |
| **Effect size** | | | | | | |
| *Delta* | 0.01 | 0.10 | 0.04 | 0.06 | -0.07 | 0.03 |
| | Random Effects $\sigma^2$ *(SD)* | | | | | |
| School Intercept ($U_{0j}$) | .00 (.00) | .00 (.06) | .01 (.10) | .01 (.08) | .00 (.06) | .00 (.03) |
| Residuals ($r_{ij}$) | .60 (.77) | .49 (.70) | .77 (.88) | .26 (.51) | .29 (.53) | .42 (.64) |
| # Parameters | 5 | 5 | 5 | 5 | 5 | 5 |
| AIC | 3741.5 | 3426.2 | 4167.5 | 2421.7 | 2570.4 | 3172.2 |
| Deviance | 3731.5 | 3416.1 | 4157.7 | 2414.6 | 2560.4 | 3162.2 |
| $X^2$ (*df*) | .08 (2) | 13.43 (2)* | .77 (2) | 2.94 (2) | 12.5 (2) | .20 (2) |

*Note:* * $p<.05$. ** $p<.01$. *** $p<.001$. $X^2$ comparison is to fully unconditional multilevel model. Level 1 (student $n$=1609*),* Level 2 (school $n$=14)

Table 3.

*Student-level Conditional Intervention Effects Model for Each Measured Outcome*

| **Parameter** | **Outcome** | | | | | |
|---|---|---|---|---|---|---|
| | *Self-Efficacy* | *Learning Strategies* | *Growth Mindset* | *Self-management* | *Social-awareness* | *Grit* |
| | Fixed Effects *β (SD)* | | | | | |
| **Intercept (Post-Pre)** | -0.35 (.12)* | -0.40 (.17)* | -0.21 (.19) | -0.34 (.11)** | -0.03 (.08)** | -0.17 (.11) |
| **Level 1 (student)** | | | | | | |
| FRL (1=Eligible) | -0.02 (.04) | -0.12 (.04)** | 0.00 (.08) | -0.03 (.03) | -0.09 (.03)*** | -0.02 (.03) |
| Grade | 0.02 (.01) | 0.04 (.02) | 0.01 (.02) | 0.03 (.01)* | 0.04 (.01)*** | 0.01 (.01) |
| Race (1=white) | -0.02 (.04) | 0.02 (.04) | -0.04 (.02) | -0.01 (.03) | 0.01 (.03) | 0.05 (.04) |
| SPED (1=receives SPED) | 0.06 (.07) | -0.02 (.07) | (-0.04) (.05) | -0.01 (.05) | -0.14 (.05)** | -0.03 (.06) |
| **Level 2 (school)** | | | | | | |
| Treatment Group (1=CharacterStrong) | -0.01 (.04) | 0.14 (.07) | 0.08 (.08) | 0.06 (.04) | 0.01 (.03) | -0.01 (.04) |
| | Random Effects $\sigma^2$ *(SD)* | | | | | |
| School Intercept ($U_{0j}$) | .00 (.00) | .00 (.06) | .01 (.11) | .0 (.05) | .00 (.00) | .00 (.03) |
| Residuals ($r_{ij}$) | .60 (.77) | .49 (.70) | .77 (.88) | .26 (.51) | .29 (.53) | .42 (.65) |
| # Parameters | 8 | 8 | 8 | 8 | 8 | 8 |
| AIC | 3743.8 | 3429.2 | 4172.7 | 2419.2 | 2555.2 | 3174.4 |
| Deviance | 3786.8 | 3413.2 | 4156.7 | 2403.2 | 2539.2 | 3158.4 |
| $X^2$ (*df*) | 3.74 (3) | 3.08 (3) | .81 (3) | 8.42 (3)* | 21.2 (3)*** | 3.79 (3) |

*Note:* * $p<.05$. ** $p<.01$. *** $p<.001$. $X^2$ comparison is to unconditional intervention effects model (Table 2). Level 1 (student *n*=1609*),* Level 2 (school *n*=14)

Table 4

*Relationship between students with most room for improvement and treatment group*

| **Parameter** | **Outcome** | | | | | |
|---|---|---|---|---|---|---|
| | *Self-Efficacy* | *Learning Strategies* | *Growth Mindset* | *Self-management* | *Social-awareness* | *Grit* |
| | Fixed Effects *β (SD)* | | | | | |
| **Intercept (Post-Pre)** | -0.17 (.02)*** | -0.10 (.03)* | -0.09 (.04)* | -0.08 (.03)* | -0.14 (.04)** | -0.05 (.02)* |
| **Level 1 (student)** | | | | | | |
| Pre-Score | -0.44 (.04)*** | -0.44 (.03)*** | -0.65 (.02)*** | -0.42 (.02)*** | -0.45 (.02)*** | -0.49 (.02)*** |
| Grade | | | | | 0.04 (.01)** | |
| SPED | | | | | -0.16 (.04)*** | |
| **Level 2 (school)** | | | | | | |
| Treatment Group (1=CharacterStrong) | 0.01 (.04) | 0.7 (.05) | 0.04 (.08) | 0.04 (.04) | 0.00 (.03) | -0.02 (.04) |
| **Interaction** | | | | | | |
| Treatment x Pre-Score | -0.11 (.05)* | -0.10 (.04)* | -0.07 (.04)+ | -0.08 (.04)* | -- | -0.04 (.04) |
| | Random Effects $\sigma^2$ *(SD)* | | | | | |
| School Intercept ($U_{0j}$) | .00 (.03) | .00 (.04) | .01 (.11) | .00 (.06) | .00 (.03) | .00 (.05) |
| Residuals ($r_{ij}$) | .47 (.69) | .37 (.61) | .46 (.68) | .19 (.44) | .24 (.49) | .30 (.55) |
| # Parameters | 6 | 6 | 6 | 6 | 7 | 6 |
| AIC | 3375.4 | 2979.6 | 3357.8 | 1952.6 | 2139.8 | 2667.1 |
| Deviance | 3363.4 | 2967.6 | 3345.8 | 1940.6 | 2125.8 | 2655.1 |
| $X^2$ (*df*) | 368.1 (1)*** | 448.7 (1)*** | 811.7 (1)*** | 0 (2)◊ | .58 (1)◊ | 507.2 (1)*** |

*Note:* + $p<.10$ * $p<.05$. ** $p<.01$. *** $p<.001$. $X^2$ deviance test comparison is to better fitting model between unconditional intervention effects model (Table 2) or (◊) conditional model (Table 3). Level 1 (student *n*=1609*),* Level 2 (school *n*=14)

Table 5

*Teacher-level Intervention Effects Models for Each Measured Outcome*

| **Parameter** | **Outcome** | | | | | |
|---|---|---|---|---|---|---|
| | *Self-Efficacy* | *SEL Competencies* | *Enthusiasm* | *Self-Efficacy* | *SEL Competencies* | *Enthusiasm* |
| | Fixed Effects *β (SD)* | | | | | |
| **Intercept (Post-Pre)** | -0.03 (.05) | -0.08 (.06) | 0.38 (.16) | -0.13 (.08) | -0.20 (.07)* | -0.42 (.19) |
| **Level 1 (teacher)** | | | | | | |
| Sex (1=M) | -- | -- | -- | -0.03 (.07) | 0.10 (.05)* | -0.01 (.12) |
| Years Experience | -- | -- | -- | 0.01 (.00)* | 0.01 (.00)* | 0.00 (.01) |
| **Level 2 (school)** | | | | | | |
| Treatment Group (1=CharacterStrong) | 0.15 (.07)* | 0.23 (.09)* | 0.25 (.24) | 0.13 (.07) | 0.24 (.09)* | 0.25 (.24) |
| **Effect Size** | | | | | | |
| *delta* | 0.18 | 0.40 | 0.51 | -- | -- | -- |
| | Random Effects $\sigma^2$ *(SD)* | | | | | |
| School Intercept ($U_{0j}$) | .00 (.00) | .02 (.12) | .11 (.34) | .00 (.07) | .01 (.08) | .11 (.33) |
| Residuals ($r_{ij}$) | .31 (.56) | .13 (.37) | .94 (.97) | .31 (.56) | .13 (.36) | .95 (.98) |
| # Parameters | 4 | 4 | 4 | 6 | 6 | 6 |
| AIC | 412.2 | 217.6 | 789.2 | 411.0 | 212.2 | 801.8 |
| Deviance | 404.2 | 209.6 | 790.2 | 399.0 | 200.2 | 789.8 |
| $X^2$ (*df*) | -- | -- | -- | 5.19 (2) | 9.46 (2)** | .43 (2) |
| Level 1 (teacher) *n* | 242 | 242 | 280 | 242 | 242 | 280 |
| Level 2 (school) *n* | 15 | 15 | 16 | 15 | 15 | 16 |

*Note:* * $p<.05$. ** $p<.01$. *** $p<.001$.

Figure 1.

CONSORT

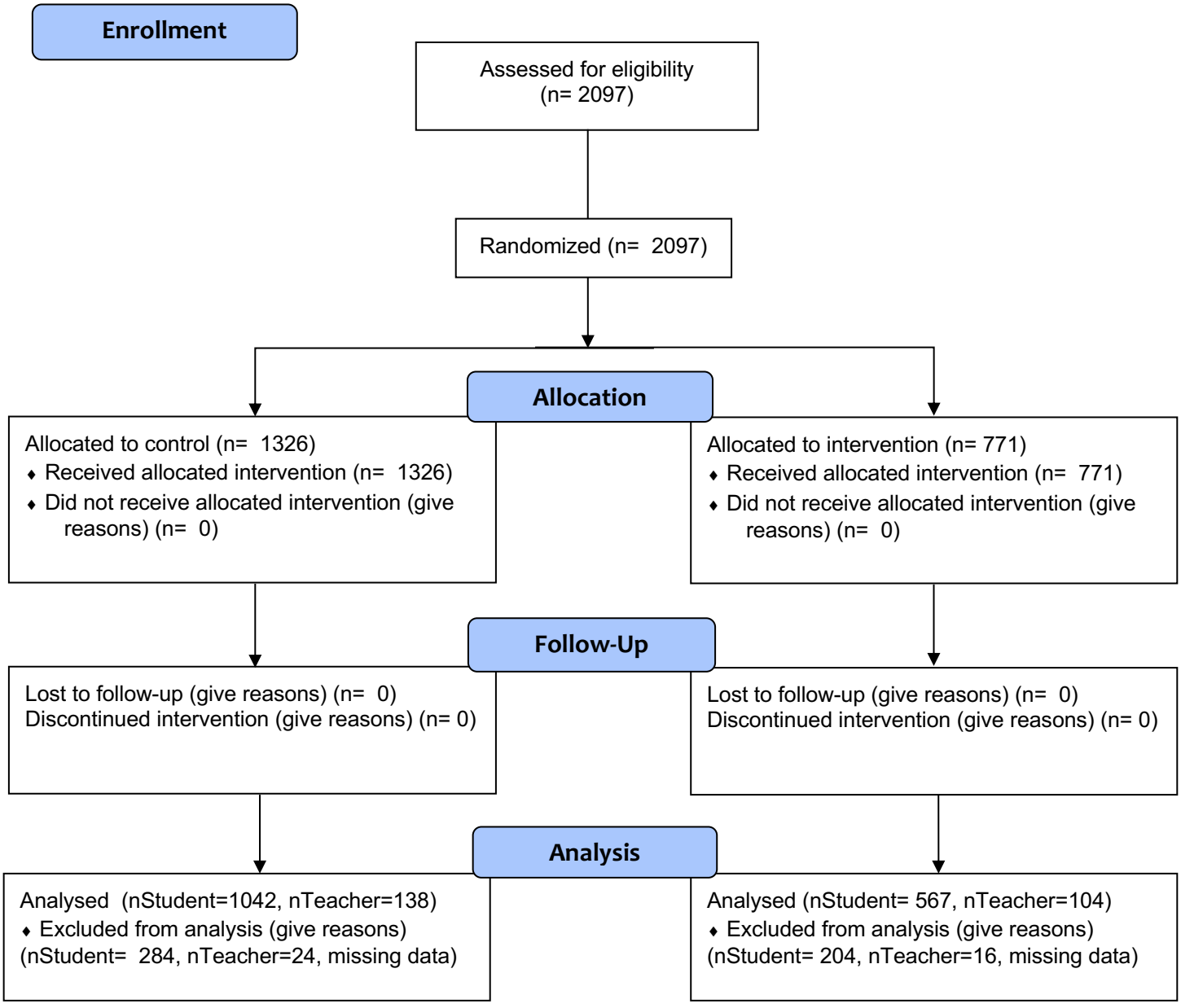
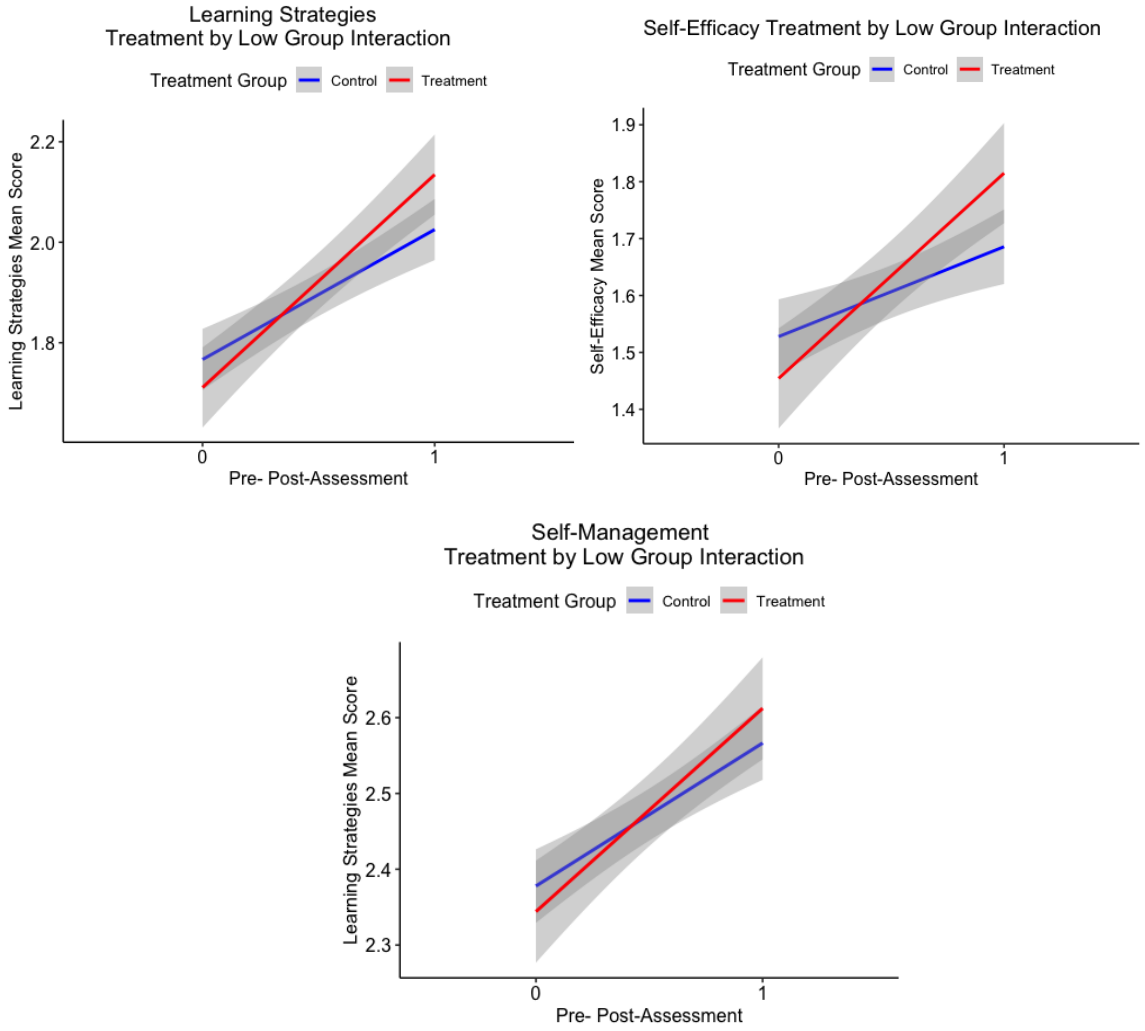


Figure 2

*Treatment by Low Group Interactions*




*Note*. This sample only includes students who scored in the bottom-third of the outcome variable (Learning Strategies: *N*=574, *n*=202 treatment, *n*=327 control; Self-efficacy: *N*=584, *n*=199 treatment, *n*=385 control; Self-management: *N*=556, *n*=204 treatment, *n*=352 control)