# EXAMINING THE EFFECTS BETWEEN TEACHER QUALIFICATIONS AND PROGRAM IMPLEMENTATION ON IMPROVING SOCIAL SKILLS AMONG STUDENTS IN K-3RD GRADE

by

MARK A. HAMMOND

A DISSERTATION

Presented to the Department of Counseling Psychology and Human Services
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2021

DISSERTATION APPROVAL PAGE

Student: Mark A. Hammond

Title: Examining the Effects Between Teacher Qualifications and Program Implementation on Improving Social Skills Among Students in K-3rd Grade

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Counseling Psychology and Human Services by:

| | |
|---|---|
| Atika Khurana | Chairperson and Advisor |
| Leslie Leve | Core Member |
| Dave DeGarmo | Core Member |
| John Seeley | Institutional Representative |

and

| | |
|---|---|
| Krista Chronister | Vice Provost for Graduate Studies |

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded December 2021

DISSERTATION ABSTRACT

Mark A. Hammond

Doctor of Philosophy

Department of Counseling Psychology and Human Services

December 2021

Title: Examining the Effects Between Teacher Qualifications and Program Implementation on Improving Social Skills Among Students in K-3rd Grade

School-based social-emotional learning (SEL) programs for children at risk for behavioral problems can be effective, affordable, and scalable. Despite evidence of efficaciousness, few SEL programs progress to later implementation stages due to challenges with delivery in real-world settings. Program implementation in schools and subsequent student outcomes can be impacted by the qualifications of teachers implementing the program and teachers' experiences during program delivery and interactions with support staff. I conducted a sub-group analysis using longitudinal data from two treatment arms ($N$ = 188 teacher-student dyads in K-3$^{rd}$ grade) of a large RCT ($N$ = 379 teacher-parent-student triads), of First Step Next (FSN), a Tier-2 SEL program over 30 consecutive school days. I examined the teacher-student dyads in the treatment arms to understand the influence of teachers' qualifications (i.e., education and years of experience at their school) on dosage (i.e., sessions delivered) and students' social skills improvement. I tested the effects of teachers' implementation experiences (i.e., program satisfaction, working alliance) and interaction effects with dosage on social skills improvement. Most students in my sample ($N$ = 188; age range = 5-10 years) were eligible for reduced-price school meals (72%), male (72%), and Black (55%). Teachers

(*N* = 188) were predominantly female (92%), White (87%), and averaged 11 years of teaching. Regression analyses revealed that teachers' education was associated with dosage but not students' social skills improvement. Though dosage was associated with students' social skills improvement, there was no indirect effect of teachers' education on students' social skills through dosage. Teachers' program satisfaction and working alliance had direct effects on students' social skills improvement, with no interaction effects with dosage. The teacher's years of experience at a school did not influence dosage or students' social skills improvement. Higher dosage, better working alliance, and greater program satisfaction were positively associated with social skills improvement. Results underscore the impact of teacher education on delivering the FSN program and implementation experiences on improving social skills. Future FSN trials should monitor teachers' knowledge and implementation experiences during training, support, and delivery, to examine the dynamic effects of process measures on program delivery and student outcomes.

CURRICULUM VITAE

NAME OF AUTHOR: Mark A. Hammond

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Michigan, Ann Arbor
St. Olaf College, Northfield

DEGREES AWARDED:

Doctor of Philosophy, Prevention Science, 2021, University of Oregon
Master of Public Health, Health Behavior and Health Education, 2015, University of Michigan School of Public Health
Bachelor of Arts, Psychology, 2010, St. Olaf College

AREAS OF SPECIAL INTEREST:

Quantitative Methods, Program Evaluation, Health Education, Health Communication, and Implementation Science

PROFESSIONAL EXPERIENCE:

Project Coordinator and Lab Manager, University of Michigan, Ann Arbor, MI, 2015-2017

Program Coordinator and Data Manager, University of Michigan School of Public Health, Ann Arbor, MI, 2015

Junior Scientist, University of Minnesota, Minneapolis, MN, 2012-2013

Research Technician, University of Minnesota, Minneapolis, MN, 2012-2013

Mental Health Counselor, South Metro Human Services, St. Paul, MN, 2010-2013

GRANTS, AWARDS, AND HONORS:

First Year Fellowship, University of Oregon, 2017-2018

Dean's Scholarship, University of Michigan School of Public Health, 2015

PUBLICATIONS:

Hammond, M. A., Khurana, A. & Stormshak, E. A. (2021) Adolescent Measures of Family Socioeconomic Status: Reliability, Validity, and Effects on Substance Use Behaviors in Adolescence and Young Adulthood. *Preventive Medicine Reviews.*

Leonard, H., Khurana, A. & Hammond, M. A. (2021) Bedtime Media Use and Sleep: Evidence for Bidirectional Effects and Associations with Attention Control in Adolescents. *Sleep Health*.

ACKNOWLEDGMENTS

I offer my heartfelt appreciation to Dr. Atika Khurana for her unconditional support, kindness, and guidance throughout my doctoral training. I am also indebted to Dr. John Seeley for his mentorship, cultivating my research interests, and providing the resources and the data used in the present dissertation. I thank Dr. Leslie Leve for her acute feedback in shaping my dissertation and leadership within the Prevention Science program. I appreciate Dr. Dave DeGarmo for his enthusiastic input and advice on the analytical framing of the present dissertation.

I would also like to thank Jason Smalls for his assistance with accessing the source materials and data used for the dissertation. I send my appreciation to Jeff Gau for his patient technical support with running my analyses.

Finally, I extend my deepest gratitude to my brilliant wife, Katie Hammond, whose humor, kindness, and counsel were instrumental in helping me complete this undertaking.

To my family.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

# CHAPTER I

## INTRODUCTION

Children who struggle with behavioral problems during elementary school are at an elevated risk for emotional and behavioral disorders in later years (Beyer et al., 2012; Fanti & Henrich, 2010). They are also more likely to demonstrate antisocial behavior and difficulties with maintaining healthy peer relationships (Campbell et al., 2006). Given the long-term negative implications of early disruptive behaviors, early detection and remediation have been a public health and educational priority for serving K-12 students (Lloyd et al., 2019). Social-emotional learning (SEL) programs have emerged as a popular and broadly endorsed type of evidence-based program (EBP) for addressing student behavioral problems (Greenberg et al., 2003; Wigelsworth et al., 2016). SEL programs include trainings and curriculums that aim to improve children's ability to comprehend and regulate emotions and develop social skills that promote positive social behaviors and relationships (CASEL, 2019). SEL programs have been widely adopted by schools given their utility (Foster & Jones, 2005) and potential cost-effectiveness (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004), with flexible pricing depending on school needs and availability of personnel for program implementation (Frey et al., 2019).

Accompanying the wide-scale adoption of SEL programs in schools (Foster & Jones, 2005), there has been increasing policy and funding directed towards delivering, evaluating, and sustaining SEL programs in schools, intending to close the gap between efficacy trials and real-world application (Weisz et al., 2005; Wigelsworth et al., 2016). Most SEL programs rely on teachers as the primary implementers, incorporating SEL curricula into classroom contexts (Durlak et al., 2011). The significant variability in

teachers' background, experience, and preparedness to deliver such programs (Lendrum et al., 2013) is an important reason why, despite their effectiveness, SEL programs are often not as effective in school settings as would be expected (de Leeuw et al., 2020; Durlak, 2015).

To promote effective delivery of SEL programs in real-world settings, researchers need to understand how school-level factors, like teacher characteristics and experiences delivering the program, can influence program implementation and student outcomes. This information needs to be collected through early phases of program development (e.g., efficacy trials) and guide pre-implementation planning in later stages (e.g., effectiveness trials). This dissertation aims to contribute to this goal by examining how teacher qualifications and experiences with program implementation can impact program delivery and student outcomes associated with an evidence-based Tier 2 school-based SEL program. The findings will inform future program implementation by identifying areas where further support is needed for teachers to optimize program delivery and student outcomes.

## Background and Significance

Given the negative consequences associated with unaddressed problem behaviors in children, such as externalizing disorders (Bradshaw et al., 2010; Burke et al., 2010), it is critical to understand how to best support school-based SEL program implementation. As specialized curriculums developed to promote social skills, SEL programs have met widespread approval and empirical support for remediating various problem behaviors among preschool and elementary-age children. To ensure that SEL program efficacy translates into school contexts, researchers need to demonstrate that teachers of varying

backgrounds can effectively deliver the program and that the implementation process is agreeable. The program results must also provide clear direction to help decision-makers decide whether the SEL program is a good match for schools and participating students.

**Multitiered Approach to Delivering Evidence-Based Programs in Schools**

In response to the extending range of services that school systems potentially need to deliver, from reading to behavioral support, they have adopted multitiered approaches to better match school services with their students' diverse and unique needs. Within these tiers, EBPs are delivered according to students' needs and response to treatment: Tier 1 includes universal programs where all students receive the same small group programming aimed at preventing initial problem behaviors before they emerge; Tier 2 provides selective programs for students who are at risk for problem behaviors, where focus students receive another level of specialized programming in addition to the Tier 1 programming; finally, Tier 3 involves indicated programs for students who display ongoing problem behaviors and require higher levels of support to reduce the severity of negative behaviors (Mrazek & Haggerty, 1994).

Multitiered systems involve matching students to the appropriate level of support required and assessing program efficacy based on whether adequate change has been achieved because of a program (Gresham, 2007). Making those decisions requires metrics that assess students' response to treatment and can guide clinical decision-making regarding whether to modify, maintain, or increase the intensity of the program to meet students' needs (Gresham, 2004). As schools increasingly adopt multitiered treatment approaches, there is a growing need for research to accurately measure and predict students' response to treatment (August et al., 2018).

### *Measuring Response to Treatment Using Responder Analysis*

In appraising students' response to treatment, testing significance and size of effects for student outcomes may be insufficient for schools to accurately identify which EBPs to implement. Precise evaluation of a program's efficacy and effectiveness requires that the operationalization of a program's impact reflects statistically significant and clinically meaningful change for students (Gresham, 2005; O'Connor & Klingner, 2010). To address this need, education researchers have used responder analysis – a method commonly used in the fields of medicine and public health (Kraemer et al., 2006; Sun et al., 2010) – to assess clinically meaningful change for students in response to a program using a response threshold based on a normative sample (Jacobson et al., 1999). Interpreting a responder analysis assumes that if a student does not show an adequate response to the most appropriate program for their needs, they need additional support, including more intense programs or specialized assistance (Gresham, 2007). For that assumption to be viable, the allotted EBP must also have been effectively implemented (Dane & Schneider, 1998; Sanetti & Kratochwill, 2009).

As teachers are the primary implementers for SEL programs in schools, understanding how their characteristics may influence program implementation and the subsequent student outcomes is critical to support effective program implementation. Substantial work has documented the effects that teachers can have on program implementation (Beets et al., 2008; Johnson et al., 2018; Sy & Glanz, 2008). Teachers' perceptions of the program, their working relationships with program technical support, and their knowledge regarding teaching practices and experience in the classroom are all implicated in effective EBP implementation. In this dissertation, I examine the effects of

teacher qualifications on program dosage and students' social skills improvement, and the effects of teacher experiences delivering an EBP and dosage on students' social skills improvement. Below I summarize the theoretical frameworks and extant findings related to teacher effects on EBP implementation.

**Effects of Teacher Qualifications on Program Implementation**

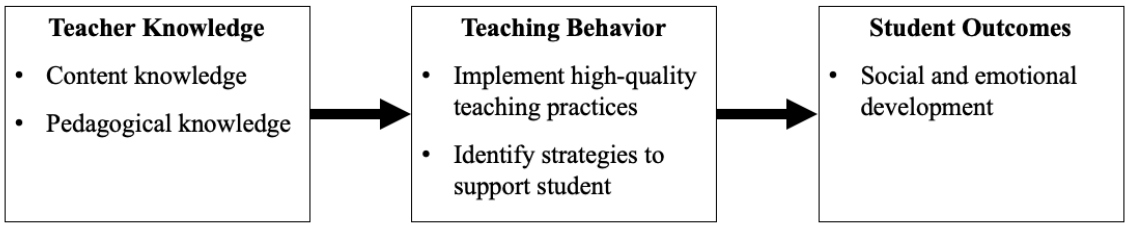
Part of a program's efficacy and implementation may be attributed to teacher characteristics related to their teaching knowledge and behaviors (Domitrovich et al., 2008; Durlak et al., 2011; Williford et al., 2015). Determining teacher qualifications is easier to assess in the early planning stages of program implementation and bears weight for designing appropriate training workshops, implementation tools, and accessible protocols. This information can be benefitial to program coaches to identify areas that teachers may need more support.

Kraft et al. (2018) describe a theory of action for teacher coaching based on the evidence base for teacher-delivered interventions with coaching support. This theory posits that teacher-led program activities and lessons (i.e., teaching practices) mediate the link between teacher knowledge and student outcomes (Kraft et al., 2018). As shown in Figure 1, a teacher's behaviors, like successfully implement high-quality teaching practices and identify appropriate teaching strategies, leads to improved academic achievement and social-emotional development among students. Teachers' program delivery can be influenced by two types of teacher knowledge: a) content knowledge and b) pedagogical knowledge. Content knowledge can be operationalized as a teacher's experience in the classroom context, integrating practices into a network of procedures, expectations, and routines that teachers use in the classroom to manage students.

Pedagogical knowledge indicates the level of education, reflecting knowledge and awareness of frameworks for effective classroom management or SEL, for example (Emmer & Stough, 2001). The two types of teacher knowledge are reflected in the commonly examined metrics of a teacher's professional qualifications of educational attainment (pedagogical knowledge) and years of experience (content knowledge). This model was developed according to a meta-analytic review of coaching interventions for supporting effective teachers and did not directly test the proposed causal links. There seems to be mixed evidence for these proposed relationships in the empirical literature depending on how researchers operationalized teacher education and teaching experience and the implementation outcome.

**Figure 1**

*Theory of Action for Teacher Coaching*




*Note.* Adapted from Kraft et al. (2018)

Studies that have tested the effects of teacher education on program implementation outcomes generally report that more education is associated with better implementation, with some exceptions. In one study examining implementing a school-based program to improve academic performance among the 43 Head Start teachers, advanced educational degree was negatively associated with fidelity. In contrast,

specialty certification was positively associated with program dosage (Vartuli & Rohs, 2009). More often, however, higher educational attainment has been linked to better implementation outcomes when measured according to more specialized education as a degree major (Williford et al., 2015) and highest degree earned (Sutherland et al., 2018). Researchers speculate that higher education may promote EBP implementation due to teachers having had extended practice in structured and group settings in graduate training environments, for example, that mirror many programs' training and professional development curricula (Wanless et al., 2015). More experience learning in advanced academic settings may improve flexibility and knowledge of teaching skills (Forman et al., 2009). In the context of EBP training, more supervised practice with knowledge acquisition skills could help teachers more rapidly integrate new information from the initial training into managing the classroom and student behaviors (Sutherland et al., 2018).

Research on associations between teachers' level of teaching experience and implementation outcomes has also yielded varied results. Some studies suggest that a general measure of years of teaching may not be associated with SEL program dosage (Marti et al., 2018; Williford et al., 2015). However, one study found that more experienced teachers had better fidelity to the program curriculum than their peers (Downer et al., 2009). Another study noted that the number of years a teacher spends working within a particular educational setting was positively associated with dosage (Marti et al., 2018). Given these differences based on teacher characteristics, it is essential to identify and examine these effects so that future school-based SEL

implementation can consider these variations and provide necessary supports to teachers to implement program activities and curriculum successfully.

**Effects of Teacher Qualifications on Student Outcomes**

Although research examining the direct effects of teacher qualifications on student outcomes is relatively scant compared with those testing effects on implementation outcomes, the available research indicates that more professional training and practice, rather than a teacher's general education level or years of experience, may positively influence student outcomes in response to a program (Huang & Moon, 2009). Nevertheless, some studies that have found no effects of teacher qualifications on student outcomes (Marti et al., 2018). Null results may be due to low variance in the measurement of teacher qualifications or because teachers with varying qualifications having equal abilities in effectively delivering a program to relay improvements among students. Testing the possible mediating role of dosage between teacher qualifications and student outcomes would contribute to unpacking the mechanisms of implementation processes (Marti et al., 2018) and guide potential coaching and pre-training for onboarding teachers.

Among the metrics for teachers' qualifications, graduate training and years of teaching experience at their school remain of interest to researchers and policymakers as indicators of effective teaching. While it is reasonable to assume that teachers with advanced training and teaching experience would have better teaching skills, testing the extent to which those qualifications translate into effective implementation of EBPs would be helpful to program developers and schools when planning targeted training and support for teachers. Therefore, examining the extent to which teachers' qualifications

can directly or indirectly impact students' response to treatment is necessary for accurately interpreting program effectiveness and planning for appropriate implementation supports for teachers.

**Program Implementation and Students Response to Treatment**

Program implementation is a dynamic and multi-staged process that defines the process by which a program is administered (Dane & Schneider, 1998; Sanetti & Kratochwill, 2009). Among the different implementation outcomes that are commonly assessed, dosage, defined as the total amount of treatment delivered over time, is a central construct for understanding the efficacy of an EBP (Durlak et al., 2011; Rowbotham et al., 2019). Dosage can reflect both total dose-received and dose-delivered, though alone is insufficient for assessing how well a program was implemented. Other less frequently studied dimensions, such as teacher's program satisfaction, may be equally or more important than dosage in determining program outcomes and should be considered even at the early stages of program development (Durlak, 2015). As all students receiving a program do not respond similarly to the recommended dosage of an EBP, there is also the question of whether teachers' experiences with delivering the program (e.g., satisfaction with the program and working relationship with the training coach) may inhibit the expected change in student outcomes despite the program dosage.

Given the number of school-based EBPs that rely on teachers as key personnel to facilitating program implementation, understanding their experiences with the implementation process, herein referred to as process measures, is essential. Process measures, like teacher satisfaction with the program and working relationship with the coaching staff, can help detect potential problems with implementation, which may mask

program effects. Though infrequently and inconsistently measured, the variance among these process measures can significantly impact programs' outcomes (Durlak, 2016). Theoretical models have identified numerous moderators of the effect of program dosage on participant outcomes, including program complexity, facilitation strategies, quality of delivery, program adaptations, and participant engagement (Berkel et al., 2011; Carroll et al., 2007; Pérez et al., 2015). Given that measuring every proposed moderator is not realistic or appropriate, determining which factors to examine largely depends on the implementation stage. During the early stages of development (e.g., efficacy trial), some of the most critical process measures reflect how teachers receive the program components (e.g., curriculum, coaching). Factors like teachers' satisfaction with the program and working relationship with the program coach (Johnson et al., 2018) are central to determining what sort of implementation supports may be needed for successful implementation.

### *Teacher Satisfaction with the Program*

Teachers' satisfaction with a program can also moderate the effect of dosage on student outcomes. Some research indicates that teachers who are more engaged with a program (i.e., more satisfied) tend to have better behavioral outcomes with the same level of dosage (Lippke et al., 2016). Teachers are also more likely to implement programs with fidelity when they feel they can do so easily and successfully (Ennett et al., 2003) and when the program is perceived to be feasible (Chafouleas et al., 2009). While program satisfaction is measured in various ways, it holds an important role in evaluating implementation quality. Program satisfaction can be viewed as a proxy for how well the

program fits within the delivery context and can help determine whether any program component needs to be changed to facilitate the program better.

### *Working Alliance Between the Teacher and Coach*

A well-tested evidence-based strategy to support the successful implementation of school programs is using coaches to train and support teachers in delivering EBPs (Pas et al., 2015; Sanetti et al., 2014; Stormont et al., 2015). Program coaches can serve numerous functions related to training, engagement, and problem solving, to name a few (Nadeem et al., 2013). Providing coaching with feedback for teachers is one effective approach to support implementation (Kraft et al., 2018; Stormont et al., 2015). Coaching feedback and the established working relationship may improve teachers' skills and knowledge of a new program. Coaching has been associated with improvements in teachers' application of program skills and student behavioral outcomes (Cappella et al., 2012). Working relationship with the coach (i.e., working alliance) has also been found to promote dosage delivered (Johnson et al., 2018). There is also evidence of interaction effects for teacher training with dosage to impact student socio-emotional outcomes (Reyes et al., 2012). The quality of the relationship between the teacher and coach can affect how well the program is implemented and hence may be a moderator of the dose-response relationship.

In this dissertation, I will examine the effect of teacher qualifications and experiences (process measures) delivering an evidence-based Tier 2 SEL program known as the First Step Next (FSN) on program dosage and students' social skills improvement. Below I describe the FSN intervention, how it is implemented in school contexts, and the evidence regarding its effectiveness.

## FSN: An Evidence-Based Tier 2 SEL Program

The FSN program was designed to improve social-behavioral adjustment for young students (PreK-3rd grade) who display early signs of problem behaviors through small group lessons and activities in the classroom (Tier 2). FSN is a manualized social and behavioral program delivered in the classroom by their teacher and a behavioral coach who also works with parents on home-based supports. The program begins with a systematic screening of young students to identify a focus student who demonstrates or is at risk for problem behaviors. Once identified, the focus student participates in classroom-based behavior-management activities that incentivize them to follow their teacher's classroom rules and work toward mastering social skills.

FSN is the latest iteration of a long-standing EBP series that has demonstrated efficacy over several large RCTs for students K-3rd grade (Sumi et al., 2013; Walker et al., 1998; Walker et al., 2009). FSN was recently upgraded to include pre-implementation procedures and resources and improved program structure, content, and delivery. These updates were provided to improve teachers' facility implementing the program, based on feedback from previous teachers. See Walker et al. (2018) for a full explanation of the justification and procedures used to inform developing the FSN program. Though FSN has been successfully implemented, there has been no examination of the within-group efficacy of the program or close inspection of the implementation process for teachers.

## Summary

Given the need for teachers to effectively implement evidence-based SEL programs in their classrooms (Durlak et al., 2011), this dissertation examines the influence of teacher qualifications and experiences on program dosage and students'

social skills improvement. Research examining teachers' impact on implementation outcomes for school-based interventions is growing but remains largely theoretical in the literature base. This dissertation helps to fill this gap by empirically evaluating teacher qualifications as predictors of dosage, and testing potential interaction effects of process measures (program satisfaction and working alliance) and dosage on students' social skills improvement. Examining how process measures can influence implementation outcomes will provide information on how FSN coaches can better support teachers in future implementation trials, and inform models of implementation processes.

# CHAPTER II

## RESEARCH QUESTIONS AND HYPOTHESES

1. Is FSN dosage (proportion of sessions delivered by teacher) associated with student social skills improvement, based on the reliable change index?

*Hypothesis:* Teachers who delivered a higher proportion of sessions will have a higher probability of their focus student demonstrating social skills improvement.

2. Are teacher qualifications (years at the school and graduate education) and process measures (program satisfaction and working relationship with coach) associated with dosage?

*Hypotheses:*

   a) Teachers with more years at the school will have a higher dosage than teachers with fewer years of experience at the school.
   b) Teachers with a graduate education are expected to have a higher dosage than teachers with fewer years of education.
   c) Teachers who reported more satisfaction with the program are expected to have higher dosage than those with lower levels of satisfaction.
   d) Teachers who reported a positive working relationship with the program coach are expected to have higher program dosage than teachers who had less positive working relationship with the program coach.

3. Are teacher qualifications and process measures associated with student social skills improvement?

*Hypotheses:*

a) Teachers with more years at the school will have a higher probability of their focus student showing social skills improvement.

b) Teachers with more education will have a higher probability of their focus student showing improved social skills.

c) Teachers who reported more satisfaction with the program are expected to have a higher probability of their focus student showing improved social skills. Shown in Figure 3, represented as pathway $b_2$.

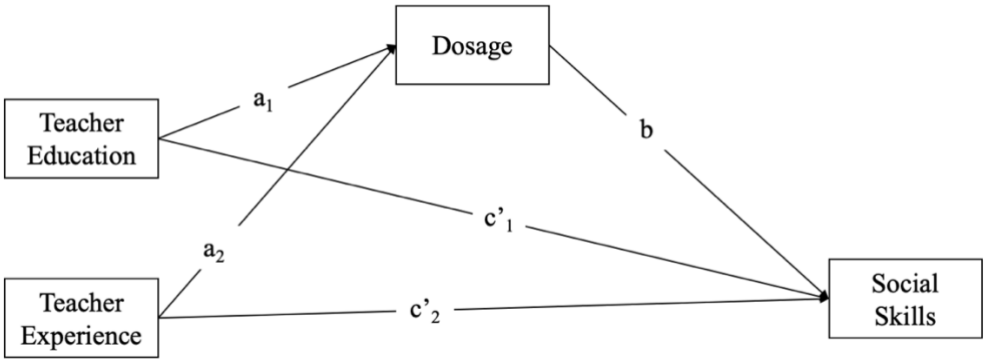
d) Teachers who reported a more positive working alliance with the program coach will have a higher probability of their student showing improved social skills. Shown in Figure 4, represented as pathway $b_2$.

4. Does program dosage mediate the potential effect of teacher qualifications on student social skills improvement?

*Hypothesis:* Dosage will account for part of the effect that teacher qualifications have on student social skills improvement. See Figure 2 for illustration of pathway.

**Figure 2**

*Hypothesized Mediation Model*



5. Does the potential effect of program dosage on student response to treatment vary based on process measures?

*Hypothesis:* The effect of dosage on student social skills improvement would be stronger when teachers report more satisfaction with the program (see Figure 3) and better working alliance (see Figure 4).

**Figure 3**

*Hypothesized Moderated Mediation Process Model with Program Satisfaction*

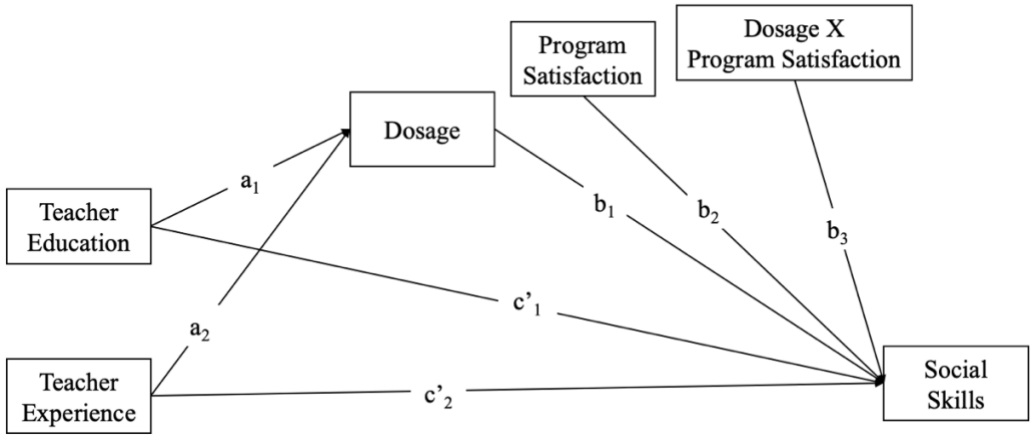
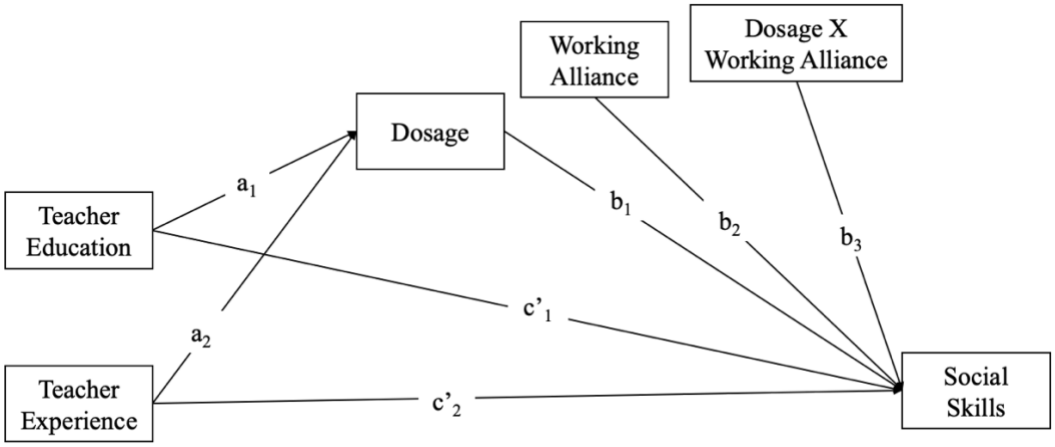



**Figure 4**

*Hypothesized Moderated Mediation Process Model with Working Alliance*



# CHAPTER III

# METHOD

## Study Setting and Participants

The present study uses data from 188 teacher-student dyads randomly assigned to one of the two treatments arms of an FSN efficacy trial. These data were drawn from a large RCT trial with 379 teacher-parent-student triads from 100 schools in five districts in Kentucky and Indiana. They were randomly assigned to the treatment arms ($n$ = 188) or business-as-usual group ($n$ = 191). The current study sample of 188 teacher-student dyads was used to examine the effects of teacher qualifications on program dosage and teacher qualifications and process measures on student response to treatment. Most participating students were male (72%), Black (55%), and eligible for reduced-price meals at school (72%). At baseline, students were aged 5-10 years old, and they were enrolled in K-3rd grade (see Table 1 for student demographics at baseline).

Few of the students' parents held a bachelor's degree or higher (14%), while nearly 70% of parents were currently working, and approximately one-third (36%) were living below the poverty level. The teachers in the sub-sample had worked in the teaching profession for an average of 11 years ($SD$ = 8.3), were predominantly female (92%) and White (87%). See Table 2 for teacher demographics at baseline.

**Table 1**

*Student Baseline Demographic Characteristics*

| Demographic Characteristic | Total (*N* = 188) |
|---|---|
| Age *M (SD)* | 6.82 (1.24) |
| Female | 28% |
| Grade | |
| Kindergarten | 25% |
| 1st grade | 23% |
| 2nd grade | 28% |
| 3rd grade | 24% |
| Sped. Services (IEP) | 23% |
| Hispanic | 5% |
| Race | |
| African American | 55% |
| Caucasian | 36% |
| More than one race | 7% |
| Unknown or unreported | 2% |
| Free or reduced lunch eligible | 70% |
| Parent college education | 14% |
| Parent employment | 70% |

**Table 2**

*Teacher Baseline Demographic Characteristics*

| Measures | Total (*N* = 188) |
|---|---|
| Years of experience at school *M (SD)* | 6.50 (6.47) |
| Education | |
| BA/BS | 17% |
| BA/BS with 1+ year coursework | 13% |
| MA/MS | 52% |
| MA/MS with 1+ year coursework | 17.5% |
| PhD | 0.5% |
| Female | 93% |
| Non-Hispanic | 97% |
| Race | |
| Hawaiian/Pacific Islander | 9% |
| African American | 1% |
| Caucasian | 87% |
| More than one race | 1% |
| Unknown or unreported | 2% |

## Data Collection Procedures

Before randomizing students to treatment and control groups, all participating teachers completed a pre-test questionnaire answering questions related to their demographic background, years of teaching experience, education level, years at the school, and questions regarding the focus student's social skills. At post-test, teachers completed a survey containing questions about the implementation process, including program satisfaction and working alliance (i.e., teachers' reports of how well they worked with their FSN coach), and the focus student's social skills. Parents received $50 to complete the screening process, while teachers received $75 for completing the survey at each timepoint.

## Recruitment and Screening

Following university and school district institutional review boards' approval, project staff recruited teachers across five cohorts from 2015-2020. Eligible students were identified for inclusion in the project using teacher and parent-reported screening data across a two-step process. First, teachers completed Stages 1 and 2 from the Systematic Screening for Behavior Disorders (SSBD; Walker et al., 2014) to identify five students in their classrooms who were at elevated risk for externalizing behaviors. Project staff rank-ordered students who met SSBD Stage 2 cutoff criteria based on severity and focused on the highest-ranked student in each classroom for inclusion. Second, project staff collected the externalizing scale of the Child Behavior Checklist (CBCL; Achenbach et al., 2001) from the parents of the target student to verify the student was struggling with behavior across the school and home settings.

Project staff recruited a family to participate if the student was in the borderline or clinical range on the CBCL externalizing scale. If the highest-ranked student on the SSBD did not meet CBCL criteria or if staff were unable to screen or recruit the highest-ranked student, researchers repeated the process with the next highest-ranked student in the classroom. In this study, data were from student-teacher dyads who either received FSN only treatment or FSN with a parent component, totaling 188 participants between both treatment groups.

## First Step Next (FSN) Description

The FSN program uses multiple influences in the student's life (e.g., parents, teachers, peers). The program includes three major tasks: social skills instruction, the green-card game, and home-school connections. During social skills instruction, a behavioral coach delivers school skills lessons (i.e., *Super Student Skills lessons*) to help the focus student build positive interpersonal relationships, develop problem-solving skills, and improve self-regulation. During the green card game, the teacher uses a color-coded card to provide subtle but direct, non-verbal feedback to encourage the student to either continue using the Super Student Skills or stop, think, and change their behavior. The green card game is integrated into daily academic, social skills lessons, and other classroom activities. Throughout FSN implementation, the game is played regularly during the school day with increasing duration. For the parent component, caregivers receive daily feedback via a note or phone call from the FSN coach and materials to encourage positive parenting strategies at home.

### *FSN Program Protocol*

Before the FSN program begins, the coach leads a series of pre-implementation activities that include the coach, teacher, parent, focus student, and the other students in the classroom. The coach meets with the teacher to explain the program, obtain the teacher's buy-in, and gather more information about the focus student and their classroom in general. Coaches also evaluate students on social skills and problem behaviors by observing the focus student in the classroom and on the playground while interacting with peers. The coach also meets with the student and their classroom peers to introduce the program and solicit buy-in.

FSN implementation takes place in three phases over a total of 30 consecutive days of delivery. First, during the coach phase (program days 1-7), a trained coach delivers the skills lessons directly to the focus student, implementing the green-card game, and communicating daily with the student's parents or caregiver via notes or phone calls (for the parent component). Second, during the transition phase (program days 8-10), the coach transitions control and management to the teacher. During this phase, the teacher delivers the skill lessons directly to the focus student, implements the green-card game, and communicates daily with the student's parent/caregiver. The coach continues to support the focus student, delivering skill lessons while the teacher provides support for the coach. Third, during the teacher phase (program days 11-30), the teacher supervises playing the green-card game and reviews the skill lessons with the focus child as needed. If the student does not successfully complete a day's activities during days 1-18, depending on the FSN phase (i.e., coach, transition, or teacher), the coach (program days 1-10) or teacher (program days 8-18) repeats the last day of programming when the student was successful. During days 19-30, if the student is not successful for 2-3

consecutive days, the teacher returns to the beginning of successful days and repeats the designated activities. Therefore, a lower dosage will not necessarily reflect how many days teachers delivered the program, and only how many days the program was delivered, and students completed the daily point requirements when teachers provided the curriculum.

All the FSN coaches were university employees who were trained by the project research staff. Across the five cohorts, 33 coaches were trained and participated in FSN implementation. During training, coaches practiced introducing and implementing FSN to student-teacher dyads one at a time and role-played program activities with the focus student. During FSN implementation, coaches met weekly with research staff to troubleshoot implementation issues.

The parent component of FSN, called homeBase (hB), includes three to six 60-minute home visits throughout the intervention, from pre-randomization through post-intervention follow-up assessment. During these sessions, coaches use motivational interviewing techniques to help parents adjust their parenting practices to align with the five universal principles of positive behavior support (Sprague & Golly, 2013). hB sessions are delivered using a multi-step process to increase intrinsic motivation for adopting and using FSN with integrity. Research staff certified by the Motivational Interviewing Network of Trainers trained participating coaches using the Motivational Interviewing Training and Assessment System (Frey et al., 2017). For more information on this home component, see Frey et al. (2015).

***Business-As-Usual Group***

Teachers who were randomized to the business-as-usual (BAU) group were offered a 4-hr training session in classroom management and principles of positive behavior support. During these sessions, teachers discussed their experiences with positive behavior support and learned strategies for promoting a positive classroom environment (Sprague & Golly, 2013). Training in the BAU group was more generic than the FSN program (e.g., specific program strategies were not provided), but teachers in the BAU group received some positive behavior support training.

## Measures

### *Outcome Variable*

**Student Social Skills.** Focus student social skills were measured using the 46 social skills items (e.g., "The focus student makes friends easily") from the teacher-version of the Social Skills Improvement System Rating Scales (SSiS; Gresham & Elliott, 2018), collected at pre and post-test. Items were rated on a 4-point frequency scale (*Never, Seldom, Often, and Almost Always*) with a total possible score of 184 ($M$ = 83.70, $SD$ = 12.46, $\alpha$ = .90). The correlation between pre and post-test social skills scores was a medium effect size ($r$ = 0.39, $p < .001$). For the analyses, the SSiS scores were revised into a dichotomous 'response to treatment' variable, assigning participants to one of two groups: improved (1) and not improved (0), based on the procedure described below.

***Measuring response to treatment.*** To measure response to FSN, I calculated a clinically significant post-test response measure using a reliable change index (Jacobson & Truax, 1991). Calculating the RCI is a two-step process that accounts for the magnitude of change in scores and the change in functioning, reflected by student scores

moving across a specified cutoff (Jacobson et al., 1999). To assess change in functioning for students' social skills, I calculated the RCI by dividing the difference between T2 and T1 social skills scores, where T1 = baseline, T2 = post-test, and $SE_{diff}$ = standard error of difference between T2 and T1 scores:

$$RCI = \frac{Social\ Skills_{T2} - Social\ Skills_{T1}}{SE_{diff}}$$

Thus, students who demonstrated reliable change in socials skills will have a normalized score of 1.96 or greater. Given the distribution of the social skills, I classified students into two groups based on the following criteria: (a) *Improved* group if at post-test they had an RCI equal to or greater than 1.96 (*n* = 72); (b) *Not improved* group if they had an RCI less than 1.96 (*n* = 102).

***Predictor Variables***

Teacher qualifications were measured using two common metrics, of teaching experience and educational background, which were assessed as follows:

**Teaching Experience (at current school).** The amount of teaching experience was measured using teachers' reports of the number of years they have been teaching at their current school (*M* = 6.50 years, *SD* = 6.47 years; Range = 0-25 years).

**Teacher Education.** Teachers self-reported their level of education based on the following categories: BA/BS (*n* = 32, 17%); BA/BS with 1+ year coursework (*n* = 25, 13%); MA/MS (*n* = 97, 52%); MA/MS with 1+ year coursework (*n* = 33, 17.5%); PhD (*n* = 1, 0.5%). Because more than half of the teachers attained a master's degrees or more, the teacher education variable was recoded into two categories: bachelor's degree = 0 (30%) and graduate degree = 1 (70%).

***Proposed Mediator***

**Program Dosage.** During the teacher-phase of FSN implementation, days 11-30, teachers filled out classroom monitoring forms daily to record their delivery of program activities and the number of points that the focus student earned by behaving according to the program directions. At the end of implementation, classroom monitoring forms documented the total number of program days the teacher had completed. For a teacher to record a program day to be complete, the student needs to earn a predetermined number of points for the day's activities. Following previous First Step studies (Sumi et al., 2013; Walker et al., 2009), I calculated program dosage as the proportion of completed program days out of the possible 20 days of the teacher phase ($M$ = 0.47, $SD$ = 0.26; Range = 0-1). Therefore, program dosage reflects both the teacher's delivery of program instruction and the focus student's compliance with the program directions.

***Proposed Moderators***

Teachers' perceived experiences with implementing the FSN program were measured according to their self-reported satisfaction with the program and their working relationship with their assigned FSN coach. Both measures are described in further detail below.

**Program Satisfaction.** Teachers completed a 13-item satisfaction measure that assessed their satisfaction with the FSN program's usability, level of support, and perceived program effectiveness ($\alpha$ = .91). Satisfaction items were scored on a five-point Likert scale from *strongly disagree* to *strongly agree*, and the variable used for analysis was a composite score using the average of the 13 items ($M$ = 3.92, $SD$ = 0.65; Range = 1-5)

**Working Alliance.** The Therapeutic Alliance Scale is an 8-item scale that was used to measure the quality of teachers' working relationships with their FSN coach. This scale has been used in other large-scale efficacy trials (Sumi et al., 2013; Walker et al., 2009). Teachers rated items on a five-point frequency scale ranging from *never* (1) to *always* (5) to assess their coach's approachability, communication skills, follow-through, shared goals, willingness to collaborate, and overall effectiveness ($\alpha$ = .96). These items were combined to create a mean composite score for working alliance ($M$ = 4.73, $SD$ = 0.57; Range = 1-5). Due to the skewness of distribution towards higher scores (i.e., *always*), this variable was dichotomized into two groups: '*always allied'* = 1 ($n$ = 104, 55.32%), and '*not always allied*' = 0 ($n$ = 64, 34.04%).

***Covariates***

Due to the program design and individual differences between students that could have confounding effects on students' social skills in response to participating in the FSN program, I statistically controlled for the following factors.

**Treatment Group.** There may be differences in exposure to treatment between the two treatment arms included in the current analyses. The homeBase program included activities for parents to reinforce the focus student's social skills development at home. So students randomly assigned to only receive the FSN program at school ($n$ = 94) may have less support with building social skills than those who received the FSN and homeBase programs ($n$ = 94). The treatment group variable was coded as follows: 'FSN and homeBase group' = 0, 'FSN group' = 1.

**Student Age.** I included the focus student's age (*M* = 6.82 years, *SD* = 1.24 years) as a covariate as there may be developmental differences in students' social skills acquisitions during childhood (Gil-Madrona et al., 2019).

**Student Biological Sex.** Differences in acquisition of social skills may be influenced in part by sex. There is evidence that during childhood, males and females differ in levels of externalizing behaviors and caring for others (Hastings et al., 2000) which may differentially impact their social skills. Similarly, among children with ADHD, females tend to display lower levels of inattention, internalizing behavior, and peer aggression than boys (Gaub & Carlson, 1997). To account for any differences based on biological sex, I included a dichotomous sex covariate in the analyses: male = 0, female = 1.

**Student Individual Education Plan (IEP).** Students with special needs may have been receiving social skills support regarding behavioral or learning needs if they were enrolled in an Individual Education Plan (IEP) (Kurth et al., 2019). To account for potential differences related to special education services in student social skills improvement, I included a dichotomous covariate coded as: 'no IEP' = 0, 'IEP' = 1.

**Problem Behaviors.** Programs that aim to improve social skills can be impeded by students' problem behaviors, such as bullying and externalizing behaviors (Gresham, 2004). To account for this possibility, I included the baseline scores for problem behaviors as a model covariate using the subscale for problem behaviors from the teacher-version Social Skills Improvement System Rating Scales (SSiS; Gresham & Elliott, 2018). The problem behaviors subscale consists of 33 items (e.g., "The focus student bullies others") that were rated on a 4-point frequency scale (*Never, Seldom,*

*Often, and Almost Always*) with a total possible score of 184 ($M$ = 135.04, $SD$ = 12.76, $\alpha$ = .82).

## Data Analysis Plan

I evaluated whether data appeared to meet the assumptions for discrete dependent variable models, including independence among observations, multicollinearity, and normal distribution of the probability of an event, using bivariate associations (e.g., correlations), histograms, and density plots in R version 4.1.0 (R Core Team, 2021). All models were tested using M*plus* version 8.5 (Muthén & Muthén, 2020).

### Missing Data

I examined patterns and distributions of missingness to evaluate whether the data appear to be missing at random (MAR). Missing data were limited as the data has 96.4% complete observations. Six missing data patterns were identified but, these were not significantly related to any of the study variables or student or teacher demographics. Given the dichotomous nature of the outcome variable, I accounted for missing data using the weighted least-squares with robust standard errors, mean, and variance adjustment (WLSMV; Muthén et al., 1997). This approach applies full information maximum likelihood (FIML) to treat missing data among exogenous variables but does not account for missing data among the endogenous (outcome) variables.

### Main Analyses

Significance tests were examined at $\alpha \leq .05$. I evaluated the effect size of correlation coefficients according to Cohen's (1988) conventions: Small = .10, moderate = .30, large = .50. Model fit was evaluated according to the following global model fit indices and their recommended thresholds (Hu & Bentler, 1999): $\chi^2$ likelihood ratio test,

standardized root-mean squared residual (SRMR ≤ .08), root-mean square error of approximation (RMSEA ≤ .06), and the comparative fit index (CFI ≥ .90). If the $\chi^2$ likelihood ratio test was significant, then Hoelter's N was consulted to test the robustness of findings regarding sample size. As each focus student ($n$ = 188) had one teacher ($n$ = 188), there were no level 2 effects. I specified two-level models to control for the potential clustering effects for the 33 program coaches who were shared among the teachers.

**Bivariate Analyses**

For research questions 1 and 3, I used point-biserial correlation coefficients ($r_{pb}$) to evaluate bivariate associations between continuous (e.g., program dosage, teacher education) and dichotomous variables (e.g., reliable change index). For research question 2, I evaluated bivariate associations among continuous variables with Pearson $r$ as the correlation coefficient. I evaluated all correlation coefficients using $\alpha < .05$ for the significance threshold.

**Mediation Analysis**

To address research question 4 (see Figure 2), I fit a mediation model beginning with testing teacher qualification measures as observed antecedent variables, predicting social skills improvement as the endogenous dichotomous outcome variable. I then included dosage as a mediator and evaluated the statistical significance of indirect effects using 95% confidence intervals derived from bias-corrected bootstrap resampling with 1000 draws. To investigate whether the assumption of homogeneity of regression slopes is met, I also tested individual interaction effects between the predictor variables (i.e., teacher education, teacher experience, and dosage) and covariate measures. These

analyses helped determine if any of the direct effects of predictor variables were contingent on the value of the covariates.

**Moderated Mediation Process Analysis**

I tested research question 5 by creating two models, building on the mediation model from RQ4. First, I created a model with an interaction effect between teacher satisfaction and program dosage on social skills improvement (see Figure 3). This model tested whether the potential effects of program dosage on social skills vary based on levels of teacher satisfaction. In a second model (see Figure 4), I tested an interaction effect between working alliance and program dosage on social skills, testing whether the potential effects of program dosage on social skills differed between levels of working alliance. To ease interpretation of the interaction effects that include continuous variables (i.e., dosage and progam satisfaction), I mean-centered variables before running analyses. I also conducted a post-hoc investigation of interaction effects by calculating the region of significance using the Johnson-Neyman intervals and simple slopes analysis (Preacher et al., 2006).

# CHAPTER IV

# RESULTS

## Bivariate Analyses

See Table 3 for correlation coefficients between teacher qualifications, implementation process measures, social skills, and covariates. All but one predictor (teaching experience) held positive significant associations with student social skills improvement, with effect sizes ranging from small ($r < .30$) to moderate ($r \geq .30$ and $r < .50$). Among teacher qualifications, teacher education had a small but significant correlation with social skills ($r_{pb} = 0.17$, $p = .028$), whereas teacher experience was not significantly associated with social skills ($r_{pb} = 0.04$, $p = .64$). Both implementation process measures were significantly correlated with social skills, with a small effect size for alliance ($r_{pb} = 0.20$, $p = .008$) and a moderate size for satisfaction ($r = 0.38$, $p < .001$). The association between program dosage and social skills was also significant and small ($r_{pb} = 0.27$, $p = .004$). Those patterns indicate that teachers with an advanced degree, with better implementation experiences (good working relationship with the coach and satisfaction with the program), and who delivered more program days, also had a higher likelihood of their student improving in social skills.

**Table 3**

*Correlation Matrix of Study Variables*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Stu. SS | . | | | | | | | | | | |
| 2. Prg. Dosage | 0.22** | . | | | | | | | | | |
| 3. Tch. Exp | 0.04 | 0.09 | . | | | | | | | | |
| 4. Tch. Ed | 0.17* | 0.20** | 0.32** | . | | | | | | | |
| 5. Working Alliance | 0.20** | 0.19* | 0.15† | 0.10 | . | | | | | | |
| 6. Prg. Satisfaction | 0.38** | 0.26** | 0.21** | 0.21** | 0.55** | . | | | | | |
| 7. Treatment Group | -0.02 | -0.06 | -0.07 | -0.01 | -0.03 | 0.03 | . | | | | |
| 8. IEP | 0.08 | 0.05 | 0.00 | 0.01 | 0.00 | 0.02 | -0.03 | . | | | |
| 9. Prb. Bhvr. | 0.12 | -0.13† | -0.18* | -0.09 | -0.07 | -0.04 | 0.03 | 0.03 | . | | |
| 10. Stu. Age | -0.17* | -0.06 | -0.12 | 0.02 | 0.00 | -0.03 | 0.04 | -0.05 | 0.04 | . | |
| 11. Stu. Bio sex | 0.11 | 0.06 | 0.05 | 0.13† | 0.13† | 0.17† | -0.01 | -0.10 | 0.34** | 0.01 | . |
| N (total = 188) | 174 | 173 | 187 | 188 | 168 | 169 | 188 | 188 | 182 | 188 | 188 |

*Note.* Stu = student, SS = social skills, Prg = program, Tch = teacher, Exp = experience, Ed = education, IEP = individualized education plan, Prb Bhvr = problem behavior. Point-biserial correlation coefficients are presented for associations among dichotomous (social skills, teaching education, working alliance, treatment group, IEP, and student biological sex) and continuous variables (dosage, student age, satisfaction).

†$p < .10$, *$p < .05$, **$p < .01$.

I found similar correlation patterns and effect sizes between predictor variables and program dosage. For teacher qualifications, teacher experience was not significantly associated with dosage ($r$ = 0.08, $p$ = .257), whereas the association between teacher education and dosage was significant with a small effect size ($r_{pb}$ = 0.27, $p$ = .007). There were also significant small associations between alliance ($r_{pb}$ = 0.18, $p$ = .015) and program dosage, as well as program satisfaction and program dosage ($r$ = 0.25, $p$ < .001). Those correlations of predictor variables with program dosage reflect that, on average, teachers with an advanced degree, who were '*always allied'* with their coach, and had better satisfaction with the program, also delivered more program sessions.

**Mediation Analysis**

To examine whether teacher qualifications had an indirect effect on social skills, I fit a pathway analysis testing the direct and indirect effects of teacher education and teacher experience on social skills, with dosage as a mediator, statistically controlling for student's IEP, problem behaviors, age, biological sex, and group assignment. Unstandardized estimates for this model are presented in Table 4. The model satisfied all a priori thresholds of standards for goodness of fit [$\chi^2(5) = 3.58$, p = .61; RMSEA = .00 (.00-.09), P-close = .79; CFI = 1.00; SRMR = .02]. Findings from this model indicated that teacher education but not teacher experience was significantly associated with dosage. Neither teacher education nor teacher experience had significant indirect associations with social skills. The only significant effects observed were for dosage ($B$ = 1.05, $SE$ = 2.30, $p$ = .02) and student age ($B$ = -0.19, $SE$ = -2.03, $p$ = .04) on social skills. See Figure 1 for standardized regression estimates from the mediation model.

**Table 4**

*Predictors of Social Skills with Indirect Effects of Teacher Qualifications Through Dosage*

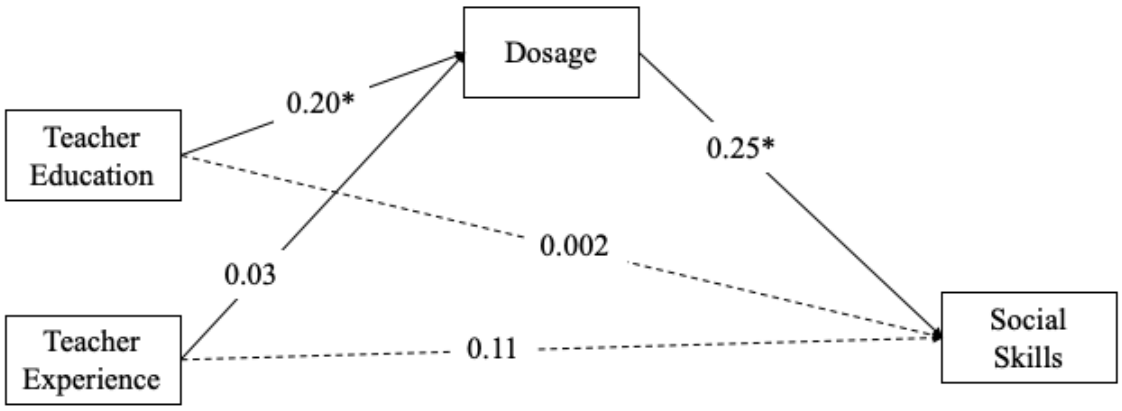
| Model Pathways | *B* | SE | *z* | *p* |
|---|---|---|---|---|
| **Direct Effects on Social Skills** | | | | |
| Dosage | 1.05 | 0.46 | 2.30 | .022 |
| Teacher experience | -0.01 | 0.02 | -0.31 | .760 |
| Teacher education | 0.38 | 0.23 | 1.64 | .101 |
| Treatment group | -0.05 | 0.19 | -0.27 | .786 |
| IEP | 0.24 | 0.22 | 1.10 | .271 |
| Problem behaviors | 0.01 | 0.01 | 1.35 | .178 |
| Student age | -0.19 | 0.09 | -2.03 | .042 |
| Student biological sex | 0.15 | 0.21 | 0.70 | .485 |
| **Direct Effects on Dosage** | | | | |
| Teacher experience | 0.002 | 0.004 | 0.40 | .640 |
| Teacher education | 0.10 | 0.07 | 1.47 | .024 |
| **Indirect Effects on Social Skills through Dosage** | | | | |
| Teacher experience | 0.002 | 0.003 | 0.73 | .685 |
| Teacher education | 0.11 | 0.06 | 1.64 | .142 |

*Note.* Unstandardized estimates are probit coefficients.

None of the interaction effects between the covariate measures and predictor variables (i.e., teacher education, teacher experience, and dosage) were significant, which supports the assumption of homogeneity of regression slopes. The final model, therefore, was defined according to the original tested model as depicted in Figure 5, with standardized estimates for the predictor variables and outcome measures.

**Figure 5**

*Mediation Model with Standardized Estimates*




*Note.* The solid lines reflect direct effects and the dashed lines represent the indirect effects through Dosage.

*$p < .05$.

**Moderated Mediation Process Analysis**

To examine potential interaction effects of the two implementation process measures with program dosage, I tested two more models building on the mediation model previously depicted in Figure 5. First, to examine whether the effect of program dosage on student social skills improvement would be stronger for teachers who reported greater working alliance with the coach, I included working alliance as a predictor to the model (see Table 5 for output). The model met thresholds for good model fit [$\chi^2(6) = 12.09$, p = .06; RMSEA = .08 (.00-.14), P-close = .204; CFI = .92; SRMR = .03] and working alliance was significantly associated with social skills ($B = 0.51$, $SE = 0.25$, $p = .044$). I then added an interaction term between working alliance and dosage to the model, creating a moderated mediation process model. This model did not meet all

thresholds for good fit of the data [$\chi^2(7) = 18.64$, p = .009; RMSEA = .10 (.04-.15), P-close = .06; CFI = .86; SRMR = .03]. The estimates are presented in Table 5.

**Table 5**

*Predictors of Social Skills with Indirect Effects of Teacher Qualifications Through Dosage and Interaction Effect Between Working Alliance and Dosage*

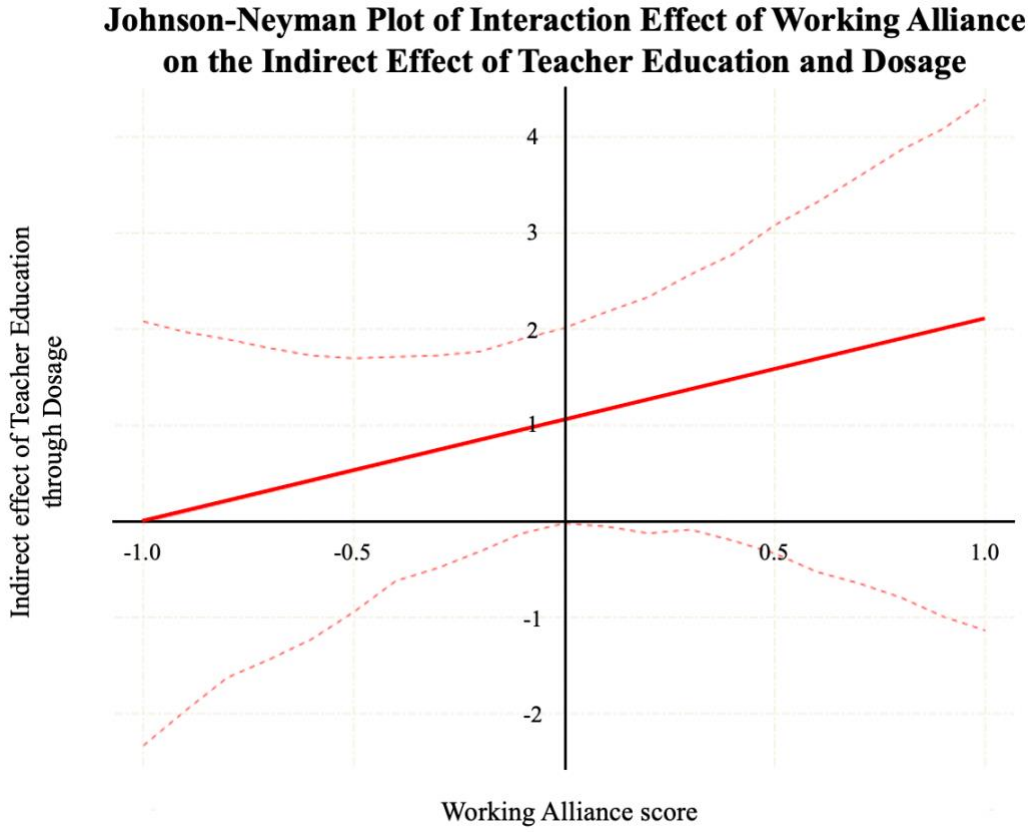
| Model Pathways | Mediation | | | Moderated mediation | | |
|---|---|---|---|---|---|---|
| | *B* | SE | *p* | *B* | SE | *p* |
| **Direct Effects on Social Skills** | | | | | | |
| Dosage | 1.05 | 0.47 | .026 | 1.05 | 0.57 | .066 |
| Working alliance | 0.51 | 0.25 | .044 | 0.52 | 0.24 | .033 |
| Dosage X Working alliance | . | . | . | 1.06 | 1.18 | .371 |
| Teacher experience | -0.01 | 0.05 | .796 | -0.01 | 0.07 | .882 |
| Teacher education | 0.34 | 0.68 | .615 | 0.34 | 0.93 | .717 |
| Treatment group | -0.04 | 0.18 | .823 | 0.002 | 0.19 | .991 |
| IEP | 0.23 | 0.22 | .297 | 0.21 | 0.25 | .390 |
| Problem behaviors | 0.01 | 0.01 | .123 | 0.01 | 0.01 | .099 |
| Student age | -0.20 | 0.09 | .026 | -0.20 | 0.09 | .025 |
| Student biological sex | 0.06 | 0.22 | .775 | 0.04 | 0.20 | .829 |
| **Direct Effects on Dosage** | | | | | | |
| Teacher experience | 0.002 | 0.004 | .602 | 0.002 | 0.01 | .852 |
| Teacher education | 0.11 | 0.06 | .057 | 0.11 | 0.16 | .489 |
| **Indirect Effects on Social Skills through Dosage** | | | | | | |
| Teacher experience | 0.01 | 0.05 | .782 | 0.002 | 0.02 | .923 |
| Teacher education | 0.05 | 0.05 | .339 | 0.11 | 0.32 | .725 |

*Note.* All estimates are unstandardized probit coefficients.

To fully examine the potential interaction effect between dosage and working alliance, I conducted a post-hoc test of the indirect effect on social skills using the Johnson-Neyman technique to calculate the region of significance for the interaction intervals (Preacher et al., 2006). See Figure 6 for a visualization of the indirect effect and the confidence intervals. As the lower confidence interval did not cross above 0, there was no evidence that there were significant differences in the effect of dosage on social skills between levels of working alliance.

**Figure 6**

*Region of Significance for the Interaction Effect of Working Alliance on the Indirect Effect of Teacher Education on Social Skills Improvement as Mediated by Dosage*




*Note.* The solid red line represents values of the adjusted indirect effect of teacher education on social skills improvement through dosage. The dashed red lines reflect 95% confidence intervals.

The moderated mediation process model with the interaction term between working alliance and dosage (Table 5) indicated that neither teacher qualification measure held significant direct or indirect associations with dosage or social skills. The only significant effects were for dosage ($B$ = 1.05, $SE$ = 0.47, $p$ = .026) and student age ($B$ = -0.20, $SE$ = 0.09, $p$ = .026), showing that higher dosage and alliance is more likely to result in improved social skills, however, there was no significant difference in the effect of dosage level on social skills between teachers who were *always allied* and those who were *not always allied* with their coach. Among the covariates, older students were less likely to respond to treatment.

In the second moderated mediation process model, I replaced the working alliance variable with the program satisfaction variable as an exogenous predictor variable for social skills improvement (see Table 6). This model met all prescribed thresholds for goodness of model [$\chi^2(6) = 12.09$, p = .060; RMSEA = .08 (.00-.14), P-close = .20; CFI = .92; SRMR = .03]. I then added an interaction effect between dosage and program satisfaction to the model, fitting a second process moderated mediation model. This model did not meet all standards for good model fit [$\chi^2(7) = 18.64$, p = .009; RMSEA = .10 (.04-.15), P-close = .06; CFI = .86; SRMR = .03] and did not successfully converge when calculating 95% confidence intervals using bias-corrected bootstrap resampling. See Table 4 for the coefficients for this model. Given the model fit statistics of those two models, and null interaction effect, I chose the former, mediation model as the final model (results presented in Table 6 and Figure 7).

**Table 6**

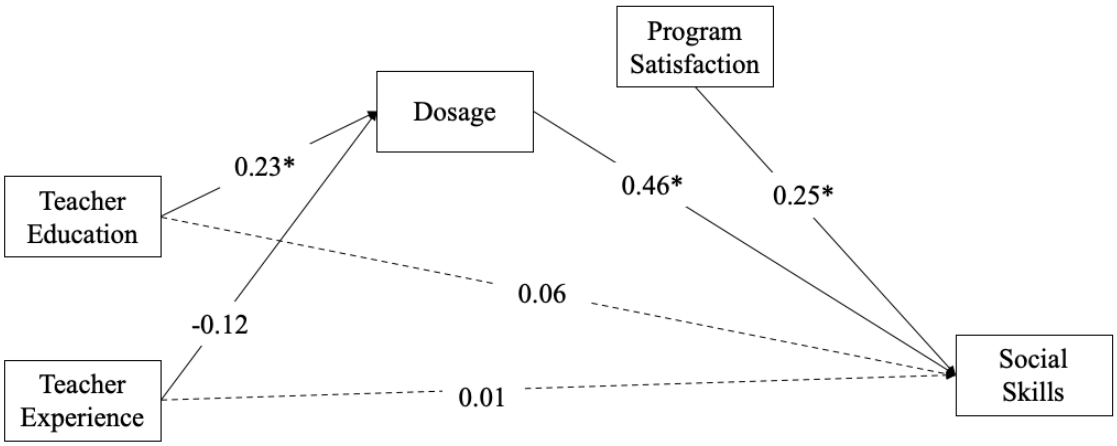
*Program Satisfaction and Other Predictors of Social Skills with Indirect Effects of Teacher Qualifications Through Dosage*

| Model Pathways | *B* | SE | *z* | *p* |
|---|---|---|---|---|
| **Direct Effects on Social Skills** | | | | |
| Dosage | 1.04 | 0.46 | 2.24 | .025 |
| Program satisfaction | 0.71 | 0.20 | 3.58 | .000 |
| Teacher experience | -0.02 | 0.02 | -0.93 | .354 |
| Teacher education | 0.19 | 0.28 | 0.68 | .494 |
| Treatment group | -0.09 | 0.17 | -0.54 | .591 |
| IEP | 0.22 | 0.22 | 1.00 | .319 |
| Problem behaviors | 0.01 | 0.02 | 1.64 | .100 |
| Student age | -0.19 | 0.08 | -2.38 | .017 |
| Student biological sex | -0.02 | 0.18 | -0.12 | .907 |
| **Direct Effects on Dosage** | | | | |
| Teacher experience | 0.002 | 0.004 | 0.53 | .595 |
| Teacher education | 0.12 | 0.06 | 2.14 | .033 |
| **Indirect Effects on Social Skills through Dosage** | | | | |
| Teacher experience | 0.002 | 0.005 | 0.43 | .665 |
| Teacher education | 0.12 | 0.09 | 1.39 | .163 |

*Note.* All estimates are unstandardized probit coefficients.

**Figure 7**

*Mediation Model Including Program Satisfaction with Standardized Estimates*




*Note*. The solid lines reflect direct effects and the dashed lines represent indirect effects through Dosage.

*$p < .05$.

In the final model including program satisfaction, the effect for satisfaction on social skills was significant and positive. The effects of teacher education and dosage were also significantly associated with higher likelihood of improved social skills. No indirect effects were significant for the teacher qualification variables. Student age was the only covariate variable that had a significant effect of social skills improvement, which was negative.

# CHAPTER V

## DISCUSSION

The results from this dissertation contribute to the literature examining the effects of teachers' qualifications and delivery of the FSN program on improving students' social skills in K-3rd grade who were at risk for behavioral problems. I examined the link between program dosage and social skills improvement, tested teacher qualifications as predictors of program dosage, and direct and interaction effects of implementation process measures with dosage on student social skills improvement. Results indicated that teachers with more advanced degrees delivered more dosage than those with bachelor's degrees. Teachers' years of experience at their school were not associated with how much of the program they delivered. After accounting for the effects of covariates, neither teacher qualification measure was related to social skills improvement. Although program dosage was associated with a higher likelihood of social skills improvement, dosage did not mediate the effect of teacher education on social skills. Teacher's program satisfaction and working alliance were directly linked to social skills improvement, but they did not moderate the effect of dosage.

The results guide future implementation projects, specifically regarding how to support teachers with delivering the FSN program. Notably, the findings substantiate the underlying core components of the FSN curriculum, where more progress through the program (i.e., dosage) promotes reliable improvement in social skills.

Teachers, on average, completed approximately half of the scheduled program days during the teacher-stage of implementation. This deviation from planned dosage may have been related to issues with sustaining the program over 20 school days, or teachers

may have had varying levels of ability to apply the program in-class due to knowledge and skills that are related to their level of formal education. More information may be needed to understand how to prepare teachers are for delivering FSN. Attention to teachers' educational background (pedagogical knowledge) may help improve the proportion of the program delivered. Additionally, enhancing teachers' satisfaction with the program and working relationship with the coach, even when reported scores were high, remained relevant to optimizing student outcomes in a randomized controlled trial. It is promising that teachers with different years of experience at their respective schools were able to deliver the same amount of dosage. With both teaching qualification measures, these factors did not differ in the likelihood of their students responding to treatment. Altogether, these findings support the FSN program's theoretical underpinnings and point to the influence of general education teachers as implementers of the program in K-3rd grade to meaningfully improve students' social skills.

Policymakers and educational systems continue to emphasize terminal degrees and prior experience teaching as a priority for addressing shortages among qualified teachers. Simultaneously, school systems are applying multitiered systems for implementing targeted student support, thus requiring extended training and professional development for teachers who act as the primary implementers for many of these interventions. The extent to which traditional teachers' qualification metrics affect implementation outcomes and program effects would be of interest for guiding future applications of a program. The present study contributes to this goal by exploring the effects of teacher qualifications (teacher education and years teaching at their school) and

implementation experiences with dosage and students' social skills improvement for an evidence-based Tier 2 SEL program, First Step Next (FSN).

**Effects of Teacher Qualifications on Dosage and Student Outcomes**

As teacher-led SEL programs continue to build support as effective models for improving students' functioning in the classroom (Durlak, 2016; Eklund et al., 2018), it is important to measure how teacher-level characteristics impact program implementation and what supports are needed to promote program efficacy. While numerous psychological metrics (e.g., self-efficacy, perceptions of intervention, philosophy of mind) and interpersonal factors (e.g., quality of relationship with students) have been examined with program implementation and student outcomes (Beets et al., 2008; Biggs et al., 2008; Rubie-Davies et al., 2012), determinants of teachers effectively implementing a program also deserve attention. Research and theory in implementation science (Durlak, 2015) and professional development (Kraft et al., 2018) point to the importance of teachers' knowledge of content and pedagogical methods as two factors that determine teachers' ability to deliver aspects of SEL programs such as dosage.

My results support the hypothesized positive effect of teacher's education level on dosage but do not find evidence to support my hypothesis for teaching experience at the school. The finding for teacher education generally aligns with prior research that has examined the effects of teacher education, where advanced education, in the form of specialty certifications (Vartuli & Rohs, 2009) or highest degree earned (Sutherland et al., 2018) is positively associated with the dosage of school-based programs. One explanation for this association is that graduate education contributes to better

understanding and exposure to teaching practices and learning frameworks, translating into improved ability to deliver program activities.

Teachers with a graduate degree in social sciences may be more likely to have encountered program designs or advanced models of behavior that are reflected in the FSN curriculum. Though FSN is a new iteration of the First Step to Success program, the foundation for the curriculum is based on social learning theory and has existed for quite some time (Patterson, 1982; Patterson et al., 1992). It may be that teachers with graduate degrees previously studied social learning theory and similar models of behavior change before starting the intervention. So teachers with an advanced degree may better understand the FSN components, which helped them deliver more activities than teachers who were new to programs based on social learning principles.

It is also possible that teachers with a graduate degree were more likely to possess skills that enabled them to more readily integrate new information from the initial training into classroom management that supported their ability to implement high-quality programs. For instance, previous work has found that teachers with a graduate education delivered a program curriculum with higher levels of adherence and more competence (Sutherland et al., 2018). As a graduate degree is not required for elementary school teachers, and I was unable to determine whether teachers' advanced education was in teaching or human development, further inquiry into this finding is needed.

Regarding teaching experience and dosage, the null findings suggest that the years spent teaching at a school did not impact the teacher’s ability to deliver the intervention. Given the low average proportion of dosage, it would be of value to test potential predictors of variability in implementation among teacher characteristics. Where previous

researchers have found evidence that years teaching at a specific teaching center is associated with higher levels of dosage (Marti et al., 2018) and fidelity (Downer et al., 2009), the results of my analyses indicate otherwise, as has been found in research where years of teaching experience was not associated with SEL program dosage (Marti et al., 2018; Williford et al., 2015). Teachers with more years of experience at a school may be more familiar with the school curriculum, schedule, and classroom organization. Programs that provide sufficient training and coaching support enable teachers with different teaching experiences to deliver it with similar levels of effectiveness. Thus, results support the assumption that FSN training and coaching prepared teachers equally well for delivering the intervention in their classroom, despite the teachers' range of experiential knowledge.

Concerning the lack of association between teachers' experience and student outcomes, this may also speak to the general effectiveness of FSN for teachers with a range of experience. While research examining general teaching effectiveness on student performance (e.g., math and reading), has presented evidence that years of teaching at a specific grade level are associated with academic achievement gains (Huang & Moon, 2009), this may not be the case for teacher experience on improving SEL-related skills. Instead, my findings concur with research that has found no effect of teacher experience level on student outcomes (Cappella et al., 2012; Domitrovich et al., 2019). My results could reflect that teacher experience has little impact on impacting the change in social skills. While social skills are considered as supporting academic gains, they may not share the impact that teaching experience has on math and reading skills.

Another interpretation for the mixed evidence may be because of differences in the program supports provided to implement the program. As teaching experience is a metric to gauge a teacher's ability to integrate practices into their classroom's set of procedures, expectations, and routines to manage students (Emmer & Stough, 2001), FSN training may have equalized differences in teaching experience, resulting in no substantial differences in associations with improved social skills.

A program that trains teachers as the primary implementers, ideally, should work equally well regardless of a teacher's qualifications. Considering the results of my analysis, this may be the case for FSN concerning years of teaching experience, but not teacher education.

**Dosage Effects on Social Skills Improvement**

With increased attention to implementation outcomes, most research has attended to quantifying the extent to which a program was delivered, as the central component to successful implementation (Durlak et al., 2011). Program dosage, as measured in this study, accounted for both delivery of the program by the teacher and the compliance of the student. Thus, testing the association of dosage with student social skills improvement also tests the theoretical structure of the curriculum, wherein the program delivery leads to focus students following the teachers' rules or instruction, then affect students' meaningful improvement of social skills.

Consistent with my hypothesis, teachers who completed more than the average proportion of the program (dosage) were more likely to see a meaningful improvement in their student's social skills. There are a few ways to interpret this association. First, it may be that more dosage of the program contributes to social skills because students get

exposed to more days of the structured program. Another consideration may be that the latter half of the program's curriculum is especially effective yet difficult for teachers and students to reach. The teacher implementation phase may be especially helpful for improving social skills, which is why dosage for this part of the program was significantly associated with improvements in social skills. In either case, it should be noted that most teachers made it approximately through half of the program days, though all days were expected to be completed. This pattern indicates that teachers may need more support from coaches to enable full delivery of the intervention. The low proportion of FSN dosage may alternatively reflect that most teachers needed to recycle program days as described earlier in the introduction, where teachers repeat program days if a student does not earn enough points to complete a program day. Because of the way I calculated dosage, it is not possible to detect whether recycling of program days was related to teacher qualifications or how that influenced social skills improvement. I did statistically control for behavioral problems and special education support (IEP plans), which could be related to non-compliance with teacher directions and behavioral issues, thus influencing teachers' ability to reach optimal program dosage.

Smaller effects may also be due to the way I calculated the social skills outcome variable. Instead of examining the number of students who improved their social skills, I used a reliable change index (RCI; Jacobson & Truax, 1991) as a threshold for defining clinically meaningful improvement (i.e., the shift of SSIS score from maladaptive into the normative range). This method is more stringent and produced fewer students as having improved. While more restrictive, RCI is a robust, standardized method for estimating response to intervention and determining what support is appropriate.

**Dosage as a Mediator of Teacher Effects**

Counter to my hypothesis, I did not find evidence of indirect effects for both teacher qualification indices on student socials skills through dosage. While teacher education was significantly related to dosage and dosage were significantly associated with social skills, I did not find a significant indirect effect on social skills through dosage. However, there was evidence to support an indirect effect of teacher education on student social skills improvement through dosage as, reflected in the reduction of effect of teacher education on social skills improvement when dosage was included in the model. Lack of significant direct or indirect effects of teacher education on students' social skills after controlling for covariates, maybe because my measurement of teacher education (i.e., graduate degree vs. bachelor's degree) did not provide enough variation to determine the effect of teacher education. Graduate training may allow teachers to deliver certain components that relate to their knowledge but might not generalize to all components of FSN that support its efficacy. For instance, a graduate degree might give teachers more familiarity with teaching concepts and a better foundation for explaining the FSN curriculum to students. However, this knowledge may not account for other factors that would engage students to change behavior, which would align with the small effect size of graduate education on dosage.

Contrary to my hypothesis, I did not find evidence that the effect of program dosage on student response to treatment would be stronger when teachers reported more satisfaction with the program or a better working relationship with the coach. In other words, the effect of program dosage was not amplified or reduced due to a teacher's satisfaction with the program or their collaboration with their coach. This finding is

promising for the generalizability of the program effects. It suggests that teacher satisfaction with the program and relationship with the FSN coach did not mitigate the effects of delivering the program.

These findings do not match the few studies that have found interaction effects between dosage and other process measures such as engagement in the program (Lippke et al., 2016) and teaching training (Reyes et al., 2012). Not detecting significant interaction effects is commonly caused by insufficient distribution, or variance, of the predictor variables included in the interaction. Though a tested model may make theoretical sense, field studies often lack sufficient distribution among observations for the interaction variables to detect significant effects, making these effects rare due to issues with power (McClelland & Judd, 1993). There were likely issues related to variable distribution in the current study, which may have made it harder to detect significant interaction effects. For example, the average score for the working alliance variable was between 4 ("almost always") and 5 ("always") on a 1-5 scale. The final dichotomous variable for working alliance would have added error to the measurement of interaction variables.

It is notable that satisfaction had a significant, albeit small, effect on social skills and speaks to the impact that social validity, even under controlled settings, can have on the efficacy of the FSN program. A direct effect of program satisfaction could mean that teachers who found the program easier to use in their classrooms could better manage other aspects of the classroom environment. Teachers with higher levels of satisfaction may have better relationships with their students, which would make it easier for teachers to focus on adhering to program protocol and produce student compliance.

Relatedly, the direct effect of working alliance on social skills improvement may reflect improvements in other factors related to the teachers' quality of program implementation. For instance, teachers who reported *always* working well with their coach could have also improved their skills in other areas that would affect student socials skills like better relationships with the target student. Coaching has been associated with improvements in teachers' application of program skills and student behavioral outcomes (Cappella et al., 2012).

It is also interesting that teacher education was no longer significant after including working alliance in the model ($p < .10$). This may be due to a shared link between teacher education, dosage, and coaching support, where there may be an indirect or interaction effect between teacher education and working alliance on dosage. This is supported by previous work that found a direct effect of working relationships with the coach on dosage and that teachers with graduate degrees required less coaching support (Johnson et al., 2018).

**Limitations and Future Directions**

For appropriate interpretation of the discussed findings, the following limitations must be considered. First, concerning the mediation analysis, the proposed mediator of dosage was not manipulated or examined in contrast to a control group comparison for dosage in the classroom. While this would inhibit causal interpretations of any significant indirect effects, the analysis satisfies the temporality requirement for mediation analysis.

A second limitation is the pre-and post-intervention measurement of social skills response. Without several timepoints of data for the outcome measure, I could not examine dose-response using multiple timepoints. Additionally, I was unable to examine

the effects that the two phases that came before the teaching phase (i.e., coach phase and transition phase) may have had on student outcomes. Future research could extend the evaluation of coaching and training effects on program outcomes, in addition to testing dosage-response thresholds (i.e., Rowbotham, 2019) to examine the long-term effects to substantiate evidence for the program's effectiveness.

Lastly, process measures (i.e., program satisfaction and working alliance) were collected at one timepoint post-intervention. As such, I could not examine the indirect or dynamic effects of working alliance on dosage and student social skills. Previous research has found that working alliance (measured as a latent factor) mediated the effect of coaching support and teacher and dosage, and that teachers with graduate degrees required less coaching support (Johnson et al., 2018).

Despite these limitations, this dissertation makes several important contributions. First, I have presented evidence to support the underlying theory of FSN, wherein more program dosage is associated with subsequent improvement in social skills. Second, I found evidence for teachers' ability to deliver dosage regardless of their years of experience at the school and that more support regarding teacher education (pedagogical knowledge) during pre-implementation training may boost the amount of dosage that teachers can deliver. Third, my study contributes to a growing body of work that tests the effects of implementation process measures other than dosage to assess the effects of teachers' experiences with program implementation on student outcomes. Expanding implementation process evaluations beyond dosage to include other experiential metrics for teachers allows further probing differences in program delivery procedures to inform future improvement of the program fit and possibly promote student outcomes. Finally, I

measured students' social skills improvement associated with program dosage using a reliable change index (RCI), a standardized method for assessing reliable improvement, and controlled for student-level program assignment, demographic factors, and behavioral needs. This is important because the RCI as a threshold for social skills improvement provides an estimate for potential observed reliable and statistically significant change in response to participating in FSN, regardless of individual student differences.

In light of the discussed findings, I offer several suggestions for researchers and implementers to extend the evaluation of FSN delivery. Future research could consider including more measurement of teachers' individual differences in knowledge and previous training in further testing its effects on the implementation process and student outcomes. Building on current findings regarding teacher education's effect on program dosage and students social skills improvement, future efforts could examine how a teacher's specific degree and educational background is linked to FSN implementation and efficacy. Looking at alternative measures of pedagogical knowledge, such as previous formal instruction and certification, could add to current understanding of how teachers' education levels contribute to FSN implementation.

Researchers may also want to pay attention to the coach's role in promoting working alliance and optimizing program satisfaction in relation to improving student outcomes. The independent positive associations that working alliance, program satisfaction, and dosage held with social skills improvement underscores the value of delineating the coach's part with supporting teachers in program implementation and efficacy. Empirical evidence for coaching interventions for teachers to promote SEL

outcomes in children is needed (Kraft et al., 2018) and would add to causal frameworks of implementation beyond teacher-student relationships (Berkel et al., 2011; Carroll et al., 2007). An in-depth examination of coach-teacher interactions throughout the 30-day implementation of FSN would be of particular interest regarding program adaptation. Specifically, future FSN evaluations could examine issues and adjustments that teachers and their coaches addressed with delivering the program in relation to student response to treatment.

Research focused on understanding variance among the coach, transition, and teacher phases of implementation may be helpful in informing future efforts to prepare coaches and teachers to implement and adjust the FSN curriculum structure as needed. Understanding the effects of coaching with delivering FSN could be beneficial to later stages of implementing FSN in effectiveness trials and scaling efforts. As coaches could be a lynchpin to successful implementation, carefully monitoring and quantifying their role will be needed to inform later training of future coaches that may not be as homogenous or accessible for training as with the collection of university-employed coaches in the trial that I examined in this study. Thus, careful examination of the influence that coaches might have is a priority for pre-implementation planning.

In practice, schools that are interested in piloting or adopting FSN would need a clear and thorough understanding of the resources necessary to manifest the expected results. Coaching support is the most costly component to implementing FSN, implementation teams and school systems would benefit from knowing the specific qualities of coaching that facilitate excellent working relationships with teachers, promote program satisfaction, and potentially adjust their technical support according to

teachers' level of education, for example. Producing a clear image of the coach's role thus requires that schools regularly invest in and monitor various supports for process outcomes (i.e., dosage, program satisfaction, and working alliance) as indicators for optimizing implementation. While teachers clearly matter to classroom program delivery, and their qualifications may support student outcomes, it seems that more information is needed about how teachers' education may translate into effective FSN implementation. School administrators who seek to promote the highest levels of program implementation for FSN may not focus on traditional teacher qualifications as precursors to reliably improving student social skills. Alternatively, perhaps decision-makers would better serve their schools and students by seeking resources that would amplify technical supports for teachers to deliver FSN successfully.

Future studies of FSN might also contribute to knowing how contextual factors influence implementation quality through testing factors related to teacher-child interactions and classroom dynamics. Current data showed a clear contrast between the teachers' and students' demographics, where teachers were almost exclusively white and female, and the majority of the target students – those with the highest risk of behavioral problems – were black and male, reflecting a race/ethnicity and gender mismatch between teachers and students. Race/ethnicity matching in classrooms is of interest to education researchers as it may moderate the effects of learning environment quality on student gains, where mismatched classrooms are associated with worse outcomes than those that are matched on race/ethnicity and gender despite the quality of their classroom (Dee, 2005; Rasheed et al., 2020). While this topic is outside the scope of my study, future studies that aim to move FSN trials to effectiveness in uncontrolled, real-world

settings should look at whether teacher-student matching on race/ethnicity and gender impacts FSN implementation and outcomes.

## Conclusion

Considering the broad application of SEL programs within schools' multitiered systems of service, programs such as FSN continue to show potential for promoting social skills among youth who struggle with behavioral problems. With a history of evidence for the efficacy of First Step programs and recent adjustments to improve the program, further examination of the predictors of implementation success is warranted to extend the reach and application of FSN (Walker et al., 2018; Lloyd et al., 2019). My dissertation presents an extension of research into how teacher characteristics predict program implementation (i.e., dosage) and how implementation outcomes affect student response to treatment.

My results support proposed models explaining variation in teachers delivering EBPs, while also validating the influence of implementation outcomes, program satisfaction, and working alliance. Future research would benefit from conducting a more detailed examination of the FSN curriculum components and implementation process measures to improve FSN's accessibility and effectiveness in addressing behavioral problems in classrooms.

REFERENCES CITED

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2001). *Ratings of Relations Between DSM-IV Diagnostic Categories and Items of the CBCL/6-18, TRF, and YSR* (pp. 1–9). University of Vermont.

August, G. J., Piehler, T. F., & Miller, F. G. (2018). Getting "SMART" about implementing multi-tiered systems of support to promote school mental health. *Journal of School Psychology*, *66*, 85–96. https://doi.org/10.1016/j.jsp.2017.10.001

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K.-K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the Positive Action program: A diffusion of innovations model. *Prevention Science*, *9*(4), 264–275. https://doi.org/10.1007/s11121-008-0100-2

Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the Pieces Together: An Integrated Model of Program Implementation. *Prevention Science*, *12*(1), 23–33. https://doi.org/10.1007/s11121-010-0186-1

Beyer, T., Postert, C., Müller, J. M., & Furniss, T. (2012). Prognosis and continuity of child mental health problems from preschool to primary school: Results of a four-year longitudinal study. *Child Psychiatry & Human Development*, *43*(4), 533–543. https://doi.org/10.1007/s10578-012-0282-5

Biggs, B. K., Vernberg, E. M., Twemlow, S. W., Fonagy, P., & Dill, E. J. (2008). Teacher adherence and its relation to teacher attitudes and student outcomes in an elementary school-based violence prevention program. *School Psychology Review*, *37*(4), 533–549. https://doi.org/10.1080/02796015.2008.12087866

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral Interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, *12*(3), 133–148. https://doi.org/10.1177/1098300709334798

Burke, J. D., Waldman, I., & Lahey, B. B. (2010). Predictive validity of childhood oppositional defiant disorder and conduct disorder: Implications for the DSM-V. *Journal of Abnormal Psychology*, *119*(4), 739–751. https://doi.org/10.1037/a0019708

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, *47*(8), 791–800. https://doi.org/10.1111/j.1469-7610.2006.01636.x

Cappella, E., Hamre, B. K., Kim, H. Y., Henry, D. B., Frazier, S. L., Atkins, M. S., & Schoenwald, S. K. (2012). Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools. *Journal of Consulting and Clinical Psychology*, *80*(4), 597–610. https://doi.org/10.1037/a0027725

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*(40), 1–9. https://doi.org/10.1186/1748-5908-2-40

Chafouleas, S. M., Briesch, A. M., Riley-Tillman, T. C., & McCoach, D. B. (2009). Moving beyond assessment of treatment acceptability: An examination of the factor structure of the Usage Rating Profile—Intervention (URP-I). *School Psychology Quarterly*, *24*(1), 36–47. https://doi.org/10.1037/a0015146

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*(1), 23–45. https://doi.org/10.1016/S0272-7358(97)00043-3

de Leeuw, R. R., de Boer, A. A., & Minnaert, A. E. M. G. (2020). The proof of the intervention is in the implementation; a systematic review about implementation fidelity of classroom-based interventions facilitating social participation of students with social-emotional problems or behavioural difficulties. *International Journal of Educational Research Open*, *1*, 100002. https://doi.org/10.1016/j.ijedro.2020.100002

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review*, *95*(2), 158–165.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., Romanelli, L. H., Leaf, P. J., Greenberg, M. T., & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive Interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, *1*(3), 6–28. https://doi.org/10.1080/1754730x.2008.9715730

Domitrovich, C. E., Li, Y., Mathis, E. T., & Greenberg, M. T. (2019). Individual and organizational factors associated with teacher self-reported implementation of the PATHS curriculum. *Journal of School Psychology*, *76*, 168–185. https://doi.org/10.1016/j.jsp.2019.07.015

Downer, J. T., Locasale-Crouch, J., Hamre, B., & Pianta, R. (2009). Teacher Characteristics Associated with Responsiveness and Exposure to Consultation and On-line Professional Development Resources. *Early Education and Development*, *20*(3), 431–455. https://doi.org/10.1080/10409280802688626

Durlak, J. A. (2015). Studying program implementation Is not easy but it is essential. *Prevention Science*, *16*(8), 1123–1127. https://doi.org/10.1007/s11121-015-0606-3

Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, *46*(3), 333–345. https://doi.org/10.1080/0305764X.2016.1142504

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

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

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, *36*(2), 103–112. https://doi.org/10.1207/S15326985EP3602_5

Ennett, S. T., Ringwalt, C. L., Thorne, J., Rohrbach, L. A., Vincus, A., Simons-Rudolph, A., & Jones, S. (2003). A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prevention Science*, *4*(1), 1–14. https://doi.org/10.1023/A:1021777109369

Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the National Institute of Child Health and Human Development Study of Early Child Care. *Developmental Psychology*, *46*(5), 1159–1175. https://doi.org/10.1037/a0020659

Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, *1*(1), 26–36. https://doi.org/10.1007/s12310-008-9002-5

Frey, A. J., Kuklinski, M. R., Bills, K., Small, J. W., Forness, S. R., Walker, H. M., Feil, E. G., & Seeley, J. R. (2019). Comprehensive Cost Analysis of First Step Next for Preschoolers with Disruptive Behavior Disorder: Using Real-World Intervention Data to Estimate Costs at Scale. *Prevention Science*, *20*(8), 1219–1232. https://doi.org/10.1007/s11121-019-01035-z

Frey, A. J., Lee, J., Small, J. W., Walker, H. M., & Seeley, J. R. (2017). Motivational Interviewing Training and Assessment System (MITAS) for school-based applications. *Report on Emotional & Behavioral Disorders in Youth*, *17*(4), 86–92.

Frey, A. J., Small, J. W., Feil, E. G., Seeley, J. R., Walker, H. M., & Forness, S. (2015). First Step to Success: Applications to Preschoolers at Risk of Developing Autism Spectrum Disorders. *Education and Training in Autism and Developmental Disabilities*, *50*(4), 397–407.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(8), 1036–1045. https://doi.org/10.1097/00004583-199708000-00011

Gil-Madrona, P., Gutiérrez-Marín, E. C., Cupani, M., Samalot-Rivera, A., Díaz-Suárez, A., & López-Sánchez, G. F. (2019). The effects of an appropriate behavior program on elementary school children social skills development in physical education. *Frontiers in Psychology*, *10*(1998), 1–8. https://doi.org/10.3389/fpsyg.2019.01998

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*(6–7), 466–474. https://doi.org/10.1037/0003-066X.58.6-7.466

Gresham, F. M. (2004). Current status and future directions of school-based behavioral interventions. *School Psychology Review*, *33*(3), 326–343. https://doi.org/10.1080/02796015.2004.12086252

Gresham, F. M. (2005). Treatment integrity and therapeutic change: Commentary on Perepletchikova and Kazdin. *Clinical Psychology: Science and Practice*, *12*(4), 391–394. https://doi.org/10.1093/clipsy.bpi048

Gresham, F. M. (2007). Response to intervention and emotional and behavioral disorders: Best practices in assessment for intervention. *Assessment for Effective Intervention*, *32*(4), 214–222. https://doi.org/10.1177/15345084070320040301

Gresham, F. M., & Elliott, S. N. (2018). *Social skills improvement system: Rating scales manual*. NCS Pearson.

Hastings, P. D., Zahn-Waxler, C., Robinson, J., Usher, B., & Bridges, D. (2000). The development of concern for others in children with behavior problems. *Developmental Psychology*, *36*(5), 531–546. https://doi.org/10.1037/0012-1649.36.5.531

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation and Accountability*, *21*(3), 209–234. https://doi.org/10.1007/s11092-009-9074-2

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology*, *67*(3), 300–307.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Johnson, S. R., Pas, E. T., Bradshaw, C. P., & Ialongo, N. S. (2018). Promoting teachers' implementation of classroom-based prevention programming through coaching: The mediating role of the coach-teacher relationship. *Administration and Policy in Mental Health*, *45*(3), 404–416. https://doi.org/10.1007/s10488-017-0832-z

Kraemer, H. C., Frank, E., & Kupfer, D. J. (2006). Moderators of treatment outcomes: Clinical, research, and policy importance. *JAMA, 296*(10), 1286–1289. https://doi.org/10.1001/jama.296.10.1286

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Kurth, J. A., McQueston, J. A., Ruppar, A. L., Toews, S. G., Johnston, R., & McCabe, K. M. (2019). A description of parent input in IEP development through analysis IEP documents. *Intellectual and Developmental Disabilities*, *57*(6), 485–498. https://doi.org/10.1352/1934-9556-57.6.485

Lendrum, A., Humphrey, N., & Wigelsworth, M. (2013). Social and emotional aspects of learning (SEAL) for secondary schools: Implementation difficulties and their implications for school-based mental health promotion. *Child and Adolescent Mental Health*, *18*(3), 158–164. https://doi.org/10.1111/camh.12006

Lippke, S., Corbet, J. M., Lange, D., Parschau, L., & Schwarzer, R. (2016). Intervention engagement moderates the dose–response relationships in a dietary intervention. *Dose-Response: An International Journal*, *14*(1), 1–10. https://doi.org/10.1177/1559325816637515

Lloyd, B. P., Bruhn, A. L., Sutherland, K. S., & Bradshaw, C. P. (2019). Progress and priorities in research to improve outcomes for students with or at risk for emotional and behavioral disorders. *Behavioral Disorders, 44*(2), 85–96. https://doi.org/10.1177/0198742918808485

Marti, M., Melvin, S., Noble, K. G., & Duch, H. (2018). Intervention fidelity of Getting Ready for School: Associations with classroom and teacher characteristics and preschooler's school readiness skills. *Early Childhood Research Quarterly, 44*, 55–71. https://doi.org/10.1016/j.ecresq.2018.02.010

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, *114*(2), 376–390. https://doi.org/10.1037/0033-2909.114.2.376

Mrazek, P. B., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. National Academy Press. http://site.ebrary.com/id/10071508

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). Robust interference using weighted least squares and quadratic estimating equations in the latent variable modeling with categorical and continuous outcomes. *Unpublished Manuscript, University of California, Los Angeles, USA*.

Muthén, L. K., & Muthén, B. O. (2020). *Mplus* (8.5) [Computer software]. Muthén & Muthén.

Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy for evidence-based practices across multiple contexts: Unpacking the black box. *Administration and Policy in Mental Health and Mental Health Services Research*, *40*(6), 439–450. https://doi.org/10.1007/s10488-013-0502-8

O'Connor, R. E., & Klingner, J. (2010). Poor responders in RTI. *Theory Into Practice*, *49*(4), 297–304. https://doi.org/10.1080/00405841.2010.510758

Pas, E. T., Bradshaw, C. P., Becker, K. D., Domitrovich, C., Berg, J., Musci, R., & Ialongo, N. S. (2015). Identifying patterns of coaching to support the implementation of the Good Behavior Game: The role of teacher characteristics. *School Mental Health*, *7*(1), 61–73. https://doi.org/10.1007/s12310-015-9145-0

Patterson, G. R. (1982). *Coercive family process* (Vol. 3). Castalia publishing company.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social interactional approach: Antisocial boys*. Oregon Social Learning Center & Castilian Publishing Company.

Pérez, D., Van der Stuyft, P., Zabala, M. del C., Castro, M., & Lefèvre, P. (2015). A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. *Implementation Science, 11*(19), 1–11. https://doi.org/10.1186/s13012-016-0457-8

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437–448. https://doi.org/10.3102/10769986031004437

R Core Team. (2021). *R: A language and environment for statistical computing* (3.6.3) [R]. R Foundation for Statistical Computing. https://www.R-project.org/

Rasheed, D. S., Brown, J. L., Doyle, S. L., & Jennings, P. A. (2020). The Effect of Teacher–Child Race/Ethnicity Matching and Classroom Diversity on Children's Socioemotional and Academic Skills. *Child Development*, *91*(3), e597–e618. https://doi.org/10.1111/cdev.13275

Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on Targeted Student outcomes for the RULER approach to social and emotional learning. *School Psychology Review*, *41*(1), 82–99. https://doi.org/10.1080/02796015.2012.12087377

Rowbotham, S., Conte, K., & Hawe, P. (2019). Variation in the operationalisation of dose in implementation of health promotion interventions: Insights and recommendations from a scoping review. *Implementation Science*, *14*(56), 1–12. https://doi.org/10.1186/s13012-019-0899-x

Rubie-Davies, C. M., Flint, A., & McDonald, L. G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships?: Teacher efficacy, expectations, and goal orientation. *British Journal of Educational Psychology*, *82*(2), 270–288. https://doi.org/10.1111/j.2044-8279.2011.02025.x

Sanetti, L. M. H., Collier-Meek, M. A., Long, A. C. J., Kim, J., & Kratochwill, T. R. (2014). Using implementation planning to increase teachers' adherence and quality to behavior support plans: Implementation planning. *Psychology in the Schools*, *51*(8), 879–895. https://doi.org/10.1002/pits.21787

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward Developing a Science of Treatment Integrity: Introduction to the Special Series. *School Psychology Review*, *38*(4), 445–459.

Sprague, J., & Golly, A. M. (2013). *Best behavior: Building positive behavior support in schools* (2nd ed.).

Stormont, M., Reinke, W. M., Newcomer, L., Marchese, D., & Lewis, C. (2015). Coaching teachers' use of social behavior interventions to improve children's outcomes: A review of the literature. *Journal of Positive Behavior Interventions*, *17*(2), 69–82. https://doi.org/10.1177/1098300714550657

Sumi, W. C., Woodbridge, M. W., Javitz, H. S., Thornton, S. P., Wagner, M., Rouspil, K., Yu, J. W., Seeley, J. R., Walker, H. M., Golly, A., Small, J. W., Feil, E. G., & Severson, H. H. (2013). Assessing the effectiveness of First Step to Success: Are short-term results the First Step to long-term behavioral improvements? *Journal of Emotional and Behavioral Disorders*, *21*(1), 66–78. https://doi.org/10.1177/1063426611429571

Sun, X., Briel, M., Walter, S. D., & Guyatt, G. H. (2010). Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ: British Medical Journal*, *340*(7751), 850–854. https://doi.org/10.1136/bmj.c117

Sutherland, K. S., Conroy, M. A., McLeod, B. D., Algina, J., & Kunemund, R. L. (2018). Factors associated with teacher delivery of a classroom-based tier 2 prevention program. *Prevention Science*, *19*(2), 186–196. https://doi.org/10.1007/s11121-017-0832-y

Sy, A., & Glanz, K. (2008). Factors influencing teachers' implementation of an innovative tobacco prevention curriculum for multiethnic youth: Project SPLASH. *The Journal of School Health*, *78*(5), 264–273. https://doi.org/10.1111/j.1746-1561.2008.00299.x

Vartuli, S., & Rohs, J. (2009). Assurance of outcome evaluation: Curriculum fidelity. *Journal of Research in Childhood Education*, *23*(4), 502–512. https://doi.org/10.1080/02568540909594677

Walker, H., Feil, E. G., Frey, A., Small, J. W., Seeley, J., Golly, A., Crosby, S., Lee, J., Forness, S., Sprick, M., Coughlin, C., & Stiller, B. (2018). First Step Next: An updated version of the First Step to Success early intervention program. *Perspectives on Early Childhood Psychology and Education*, *3*(1), 89–109.

Walker, H. M., Seeley, J. R., Small, J., Severson, H. H., Graham, B. A., Feil, E. G., Serna, L., Golly, A. M., & Forness, S. R. (2009). A randomized controlled trial of the first step to success early intervention: Demonstration of program efficacy outcomes in a diverse, urban school district. *Journal of Emotional and Behavioral Disorders*, *17*(4), 197–212. https://doi.org/10.1177/1063426609341645

Walker, H. M., Small, J. W., Severson, H. H., Seeley, J. R., & Feil, E. G. (2014). Multiple-gating approaches in universal screening within school and community settings. In R. J. Kettler, T. A. Glover, C. A. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools* (pp. 47–75). American Psychological Association.

Wanless, S. B., Rimm-Kaufman, S. E., Abry, T., Larsen, R. A., & Patton, C. L. (2015). Engagement in Training as a Mechanism to Understanding Fidelity of Implementation of the Responsive Classroom Approach. *Prevention Science*, *16*(8), 1107–1116. https://doi.org/10.1007/s11121-014-0519-6

Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and Protecting Youth Mental Health Through Evidence-Based Prevention and Treatment. *American Psychologist*, *60*(6), 628. https://doi.org/10.1037/0003-066X.60.6.628

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., ten Bokkel, I., Tate, K., & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: A meta-analysis. *Cambridge Journal of Education*, *46*(3), 347–376. https://doi.org/10.1080/0305764X.2016.1195791

Williford, A. P., Wolcott, C. S., Whittaker, J. V., & Locasale-Crouch, J. (2015). Program and teacher characteristics predicting the implementation of banking time with preschoolers who display disruptive behaviors. *Prevention Science*, *16*(8), 1054–1063. https://doi.org/10.1007/s11121-015-0544-0