

10-13-2017

# The Impact of Implementation Fidelity on Middle School Student Outcomes in the Life Skills Training Program

Enya B. Vroom

*University of South Florida*, [evroom@mail.usf.edu](mailto:evroom@mail.usf.edu)

Follow this and additional works at: <https://scholarcommons.usf.edu/etd>

 Part of the [Social and Behavioral Sciences Commons](#)

---

## Scholar Commons Citation

Vroom, Enya B., "The Impact of Implementation Fidelity on Middle School Student Outcomes in the Life Skills Training Program" (2017). *Graduate Theses and Dissertations*.  
<https://scholarcommons.usf.edu/etd/7450>

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

The Impact of Implementation Fidelity on Middle School Student Outcomes in the Life Skills  
Training Program

by

Enya B. Vroom

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science in Child and Adolescent Behavioral Health  
with a concentration in Research and Evaluation  
Department of Child and Family Studies  
College of Behavioral and Community Services  
University of South Florida

Major Professor: Oliver T. Massey, Ph.D.  
Bruce Lubotsky Levin, Dr.P.H., M.P.H.  
Svetlana Yampolskaya, Ph.D.

Date of Approval:  
October 10, 2017

Keywords: Social and emotional learning, Life Skills Training, quality of implementation,  
multilevel modeling, school-based prevention

Copyright © 2017, Enya B. Vroom

## **Dedication**

Thank you to my loving parents, Ernie and Evelyn Vroom, who have sacrificed their time, effort, and personal finances to allow me to pursue an advanced degree. You both have been my strength and guidance throughout this entire process.

To my best friend, Matthew Pedrero, who has provided me with love, encouragement, and support throughout all of my higher education. I cannot thank you enough for your patience, kindness, and understanding.

## **Acknowledgments**

I am grateful for the knowledge, guidance, and time that Dr. Oliver T. Massey, my graduate advisor and major professor, has provided me throughout my graduate career. I want to thank you for your feedback and support throughout the construction of this study and thesis.

I would like to gratefully acknowledge Dr. Bruce L. Levin, and thank him for continuously challenging me to become a better writer, student, and person. Thank you for your time and assistance in editing this thesis.

Without the help of Dr. Svetlana Yampolskaya, the analyses of this study would not have been possible. Thank you for your willingness to instruct, advise, and for your unfaltering patience. I am thankful for all of the knowledge you have shared with me.

## Table of Contents

List of Tables .....	ii
List of Figures .....	iii
Abstract .....	iv
Introduction.....	1
Fidelity .....	3
Importance of Fidelity.....	4
Methods.....	6
Life Skills Training in South Florida .....	6
Setting .....	7
Participants.....	8
Procedure .....	9
Measures .....	9
Life Skills Training Pretest/Posttest Measure.....	9
Botvin Life Skills Training Fidelity Checklist-Middle School Level 1.....	10
Analytic Approach .....	12
Results.....	13
Descriptive Statistics.....	13
Multilevel Analysis .....	14
Level 1 .....	14
Level 2 .....	16
Discussion.....	19
Limitations .....	22
Implications for Behavioral Health.....	22
Implications for Research .....	22
Implications for Practice .....	23
Implications for Policy.....	24
References.....	26
Appendix: Institutional Review Board (IRB) Determination.....	31

## **List of Tables**

Table 1: Student Pretest and Posttest Outcomes.....	13
Table 2: Competency Domains Outcomes from Pretest to Posttest .....	14
Table 3: Adherence (Levels 1&2).....	15
Table 4: Participant Responsiveness (Levels 1&2) .....	16
Table 5: Quality of Delivery (Levels 1&2).....	17
Table 6: All Predictors Included (Levels 1&2).....	18

## **List of Figures**

Figure 1: Hierarchy of Sample.....	8
------------------------------------	---

## **Abstract**

Social and emotional learning (SEL) programs have become increasingly popular during last 20 years, and have shown to reduce risky behaviors (i.e., substance use), improve communication skills, academic performance, and relationships among students of all ages when implemented in schools (Weissberg, Durlak, Domitrovich, & Gullotta, 2015; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Although the benefits of SEL programs are significant, the issue of implementation fidelity often arises. The purpose of this study was to assess the impact of implementation fidelity in the Life Skills Training program (LST) implemented with middle school students of a large South Florida school district. A not-for-profit-organization (NFPO) facilitated the training of teachers who implemented LST within the school district and provided the materials necessary to carry out the program's lessons. Fidelity was assessed by eight observers from the NFPO by utilizing the Botvin Life Skills Training Fidelity Checklist-Middle School Level 1. Three core elements of implementation fidelity, adherence, quality of delivery, and participant responsiveness, were assessed. The school district was tasked with modifying and conducting the LST Pretest/Posttest Measure to assess student behavior gains.

Multilevel modeling was used to assess the effect of individual-level (gender, race/ethnicity, and socioeconomic status [SES]) and classroom-level characteristics (adherence, participant responsiveness, and quality of delivery) on student behavior outcomes measured at posttest. Results from this study indicated that on the individual level, students' race/ethnicity and SES were significantly associated with predicting student behavior gains at posttest. On the



classroom level, participant responsiveness was significantly associated with predicting student behavior gains at posttest. The findings from this study make a unique contribution to the literature as it examined frequently overlooked core elements of fidelity such as participant responsiveness and quality of delivery.

## **Introduction**

Social and emotional learning (SEL) has become increasingly popular during the past two decades. Schools, families, researchers, and policy makers have come to the realization that a child's social and emotional well-being are important (Weissberg et al., 2015) and the skills gained from SEL curriculums can potentially have a positive effect on outcomes later in life (e.g., mental health and substance use) (Klapp et al., 2017). A child lacking certain skills to comprehend and manage his/her emotions can disrupt his/her optimal cognitive and social development. Youth with inadequate emotional skills may fail to feel empathy for others and have difficulties focusing on learning and controlling their behavior (Brackett, Elbertson, & Rivers, 2015). Social and emotional capacity can influence a youth's ability to "meet the demands of the classroom", and if they are well-equipped to learn what is required and "to benefit from instruction" (Zinsser & Dusenbury, 2015).

SEL can be described as the process of youth gaining and properly employing the attitudes, knowledge, and skills that are necessary to comprehend and manage emotions, display and feel empathy for others, build and maintain positive relationships, set and achieve goals, and make responsible decisions (CASEL, 2016; Weissberg, Durlak, Domitrovich, & Gullotta, 2015). SEL programs are now housed in many schools across the United States, and are starting to extend into after school settings and within community-based organizations (Weissberg et al., 2015).

The majority of SEL programs in the United States utilize the Collaborative for Academic, Social, and Emotional Learning's five competency domains: 1) self-awareness; 2) self-management; 3) social awareness; 4) relationship skills; and 5) responsible decision-making (Weissburg et al., 2015). The first two domains, self-awareness and self-management, focus on teaching children to be cognizant of their own strengths/abilities, be conscious of other's feelings, be aware of their own feelings in the moment, and encourages tenacity through challenges and to achieve one's goals (Weissburg et al., 2015).

The next two domains, social awareness and relationship skills, focus on teaching children how to build their ability to understand that each individual is diverse (background or culturally) and to empathize and feel compassion for other's situations, build the comprehension of social norms for behavior, guides how to create and sustain stable and supportive relationships, and models conflict resolution skills (Weissburg et al., 2015). The last domain, responsible decision-making, focuses on children gaining the ability to assess safety and ethical concerns and proper behavior norms for risky behavior, to make sensible and accurate appraisals of consequences that are associated with certain behaviors, as well as take their self and other's health and well-being into consideration (Weissburg et al., 2015).

Many diverse SEL programs exist, and the approaches used in conducting SEL programs can also vary. Some programs can have a heavy focus on preventing and changing attitudes toward drug and alcohol use, violence, and risky sexual behaviors, while other programs focus on social and emotional skills or self-esteem and positive self-concept. However, many programs incorporate both approaches (Sklad, Diekstra, De Ritter, Ben, & Gravesteyn, 2012). Currently, over 500 evaluations of various SEL programs have been conducted. The majority of these evaluations have concentrated on school-based programs (Weissburg et al., 2015).

A major meta-analysis on universal school-based SEL programs (including students in kindergarten through high school) was conducted in 2011 (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). The SEL participants demonstrated significant improvement in social and emotional skills, behaviors, attitudes, and academic performance as compared to the control groups (Durlak et al., 2011). Although most of these interventions showed significant results, many of them did not monitor the implementation of the program. Forty-three percent of the studies had to be excluded because they did not employ any technique to monitor the quality of implementation (Durlak et al., 2011). The present study aims to address the gap in the literature surrounding implementation fidelity within SEL programs.

### **Fidelity**

For the purpose of this study, *fidelity* can be defined as the key components of a program or practice that are essential for programmatic impact and are definitively responsible for the intervention's effects (Allen, Linnan, & Emmons, 2012). Fidelity is an important concern when discussing evidence-based practices/programs and can assist investigators in determining why an innovation succeeds or fails (Dusenbury, Brannigan, Falco, & Hansen, 2003). If a program lacks quality implementation, the chances of producing significant and positive outcomes among its participants are vastly reduced.

Over time, the definition of fidelity has evolved to include five core elements: 1) adherence; 2) dosage; 3) quality of delivery; 4) participant responsiveness; and 5) program differentiation (Allen et al., 2012; Dusenbury et al., 2003; Carroll et al., 2007). This study focused on adherence, quality of delivery, and participant responsiveness because the measures utilized in this study focused on those three elements. *Adherence* can be defined as whether a program is being implemented as it was originally developed and the critical elements of the

program are being presented or addressed (Carroll et al., 2007; Dusenbury et al., 2003). *Quality of delivery* can be defined as the manner in which the teacher delivers a program. This has less to do with how he/she follows the guidelines and reads from a script, and more to do with how he/she acts as a facilitator and/or coach (Carroll et al., 2007; Dusenbury et al., 2003). *Participant responsiveness* addresses how participants (in this study, students) are engaged, involved, or respond to a program (Carroll et al., 2007).

Although Life Skills Training programs (LST) have been extensively adopted in schools, research has shown that implementation fidelity varies widely by classroom teachers and the program may occasionally be conducted with weak fidelity (Botvin & Griffin, 2004).

### **Importance of Fidelity**

It is important for evidence-based programs to be implemented with fidelity. Research has shown that high quality implementation is strongly associated with positive outcomes (Durlak, 2015). However, although the empirical results produced by researchers in a controlled setting suggest a program is generalizable, the need for adaptations may still be evident when administered in real-world settings (Wright, Lamont, Wandersman, Osher, & Gordon, 2015). Often, adaptations are needed in order to ensure the program is appropriate and can serve to improve the impact and “fit” between the program and the specific population and setting (Durlak, 2015; Allen et al., 2012). For example, how fidelity interacts with certain demographic variables (e.g., race/ethnicity) may be particularly important to a minority group and their outcomes, and certain modifications may need to be made in future efforts to obtain the best possible outcomes.

The goals of this study were: 1) to assess the impact of implementation fidelity in the LST program on student outcomes; 2) to examine whether the fidelity corresponds with student

outcomes; and 3) if there is a differential effect of fidelity at the classroom and school-level. Based on the core elements, it was hypothesized there will be significant associations between the core elements including adherence, participant responsiveness, and quality of delivery and an improvement in student behavior. The three evaluation questions were as followed: 1) Does the quality of implementation affect students' behavior gains?; 2) Does the implementation fidelity of the program impact students' behavior gains differently while controlling for demographic characteristics?; and 3) Do any core elements of implementation fidelity as defined by adherence, participant responsiveness, or quality of delivery predict students' behavior gains differently?

## **Methods**

### **Life Skills Training in South Florida**

The LST program, originally developed by Gilbert J. Botvin, is a SEL program that teaches social and emotional skills as well as drug resistance skills to middle and high school students (Botvin & Griffin, 2004). A not-for-profit organization (NFPO) facilitated the implementation of LST in a large South Florida school district in which the study was conducted. LST was implemented in 48 middle schools within the district. The program includes 15 core lessons that are mandatory, and can also incorporate one to three extra lessons that are content specific or expand upon a prior lesson (e.g., violence).

The staff of the NFPO was tasked with training elective teachers who implemented LST (e.g., art, band, and foreign language) in the middle schools. The elective teachers were required to attend a 2-day workshop where they became familiar with the structure, content, and goals of the LST program. The elective teachers were given a manual that had detailed lesson plans, as well as the goals and objectives for each lesson (Botvin & Griffin, 2004). Lessons ranged from 45-50 minutes in length. At each school, there was one “program champion” that was in charge of facilitating and coordinating the LST program with the school’s elective teachers. They offered support and resources to the teachers, reported back to NFPO with any questions or concerns, and were also required to attend the 2-day training workshop.

In the fall semester, 2016, teachers delivered one LST lesson every school day for 3-4 weeks. This was a different strategy than what has been used and reported in the research

literature. Usually, the first stage of LST is supposed to be given over a four month time-span or longer. The first 15 lessons being offered in 6<sup>th</sup> or 7<sup>th</sup> grade, and booster lessons would be taught in the following two school years. However, the feedback that the NFPO has received from the elective teachers in the last two years suggests that shortening the first stage of the program would be beneficial for the students and elective teachers.

### **Setting**

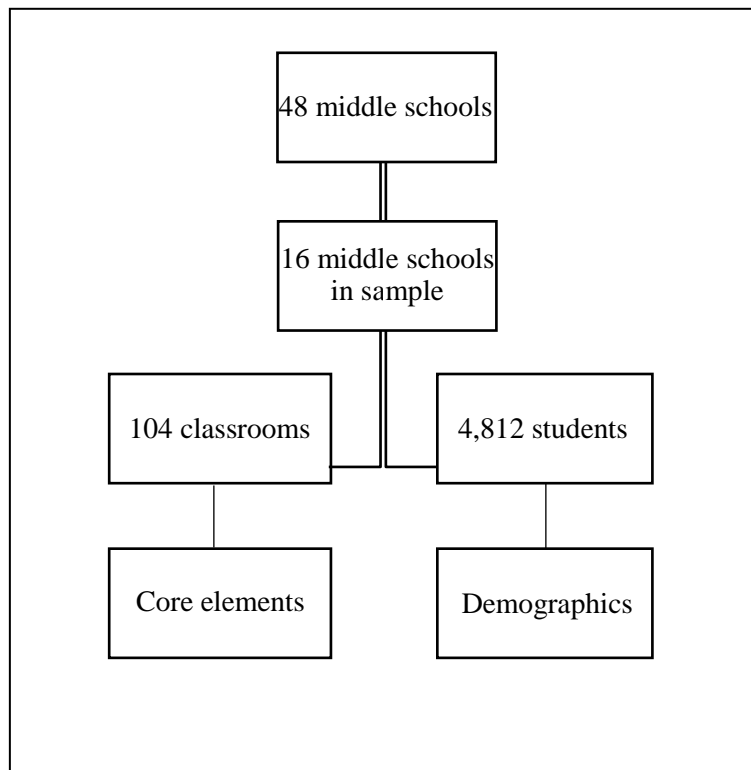
The agency responsible for providing the monetary means to conduct LST in the school district, Blueprints for Health Youth Development (Blueprints), required the NFPO to do observations and fidelity checks of the program in order to assess the quality of implementation of the participating teachers. Blueprints is an organization that provides a registry of evidence-based practices that are designed to develop the well-being and health of youth. (Blueprints, n.d.). Blueprints is based out of the Center for Study and Prevention of Violence at the University of Colorado, Boulder.

Blueprints was responsible for training the NFPO's staff and administered the training via a Skype conference call. In total, eight of the staff were trained on how to utilize the fidelity checklist. The individuals who performed the observations and completed the fidelity checks were also required to have attended at least the initial teacher training in the LST curriculum that was held at one of the participating schools. The observers coordinated with the program champions of each school to put in place the teachers' periods and schedules of when they would be teaching LST. Although the teachers were aware the observers would be coming, they did not know the exact day they would be assessed.



## Participants

The study design included a multilevel analysis of middle school student and teacher data. Participants were 4,812 6<sup>th</sup> grade students attending middle school in a large South Florida school district, as well as 104 teachers who were observed by the staff of the NFPO. The student sample consisted of 53% male and 48% females. With respect to race, 40.1% of students identified themselves as White, 33.7% identified as Hispanic, 16.5% identified as Black, 5% identified as Multiracial, 4.5% identified as Asian American. Over half of the study population was on free and reduced lunch status (55%), and 45% were not. Students' free and reduced lunch status was used as an indication for socioeconomic status (SES).



**Figure 1. Hierarchy of Sample**

## **Procedure**

The school district was responsible for modifying, disseminating, and matching the students' pretest and posttest measures. Due to the lack of resources, the NFPO was not required to conduct multiple observations in all 48 schools where LST took place. Sixteen middle schools were randomly chosen, and each elective teacher conducting LST in the chosen schools was observed one time only. Between 2 and 12 observations and fidelity checks took place at each school. The student pretest and posttest measures were matched and de-identified before the analysis. No names or identifiable information was collected from the students. The teacher names associated with the fidelity checks were also de-identified, and all fidelity checks were then attached to the students' pretests and posttests with which they correspond.

The study was not considered human subjects research by the Institution Review Board at the University of South Florida, and was approved by the school district's review board.

## **Measures**

**Life Skills Training Pretest/Posttest Measure.** The Life Skills Training Pretest/Posttest Measure (LSTM) was adapted from the original Life Skills Training Questionnaire- Middle School developed by Botvin (Botvin et al., 1994; Botvin et al., 1997). The LSTM is divided into two sections that assess knowledge and behavior. The 11 questions of the knowledge section were in multiple-choice format, and focused on knowledge acquisition from the different lessons and techniques that were taught (e.g., What is the goal of the Mental Rehearsal technique?). These questions were scored based on accuracy. The knowledge section of the LSTM was modified and added to the measure by the school district after the pretest had already been administered. Therefore, the multilevel analyses of this study were solely focused on the

behavior section of the LSTM, because it cannot provide a comparison from pretest to posttest of the knowledge section.

The 29 questions of the behavior section were answered on a Likert scale assessing how often or how likely students would engage in a behavior (e.g., “I am comfortable giving compliments to others”). Other questions of the behavior section were answered on a Likert scale assessing how much students agreed or disagreed with a statement (e.g., “It is easy for me to make friends”).

The LSTM questions can also be broken down by the previously mentioned five competency domains (e.g., self-awareness) in the knowledge and behavior sections. Drugs and alcohol was added as a sixth domain because LST also addressed these topics in four lessons (e.g., alcohol and marijuana). The pretest and posttest measure was given online to the students within each participating classroom at baseline and immediately following the conclusion of the program. Higher students’ scores on the LSTM indicated better knowledge and behavior gain. The LSTM did not include any demographic information.

Demographic information of the students, such as gender, race/ethnicity, and socioeconomic status (SES (free/reduced lunch status) was collected and matched to the each individual student’s pretests and posttests by the school district.

**Botvin Life Skills Training Fidelity Checklist-Middle School Level 1.** The observation measure used by the NFPO was the Botvin Life Skills Training Fidelity Checklist-Middle School Level 1. Multiple checklists exist and they are differentiated by levels that indicate what grade(s) the students are in and the age appropriate material included in each lesson. This level 1 checklist is specifically utilized with students in 6<sup>th</sup> and 7<sup>th</sup> grades.

The checklist, used by third-party observers, assists in helping teachers determine whether they are delivering the program content adequately and are utilizing the proper materials given to them in their training (Botvin, n.d.). The checklist consists of 15 sections in total. Each section represents its own topic/lesson as well as the objectives and activities that should be present in the lesson (e.g., making decisions). To assess adherence, the observers check yes/no on the multiple items that should be included in the lesson, indicating if the item was present or not.

Attached to the checklist is a form that includes multiple items assessing participant responsiveness and quality of program delivery. To assess participant responsiveness, the observers rated how well the students responded, understood, and engaged in the lesson. To assess the quality of program delivery, the observers are required to rate (on a scale of one to five) the different attributes of the teacher's delivery of LST (e.g., how clear were the instructions given and to what extent did the presentation of materials seemed rushed or hurried). Higher scores on the checklists indicated better implementation fidelity.

Three individual-level (student-level 1) variables were included in the multilevel analysis: 1) gender (male 0, female 1); 2) race/ethnicity (1=White, 2=Black, 3=Hispanic, 4=Asian, 5=Multiracial); and 3) whether they were or were not on free and reduced lunch status (i.e., SES) (0=No, 1=Yes). Student total outcomes were represented by the total mean sums of the students' posttest measures. Students' posttest outcomes were considered the dependent variable of this study. Three classroom-level (level 2) variables were included in the analysis: 1) adherence; 2) participant responsiveness; and 3) quality of delivery. *Adherence* was represented by adding the total number of Yes's (multiplied by two in order to account for missing items) and the total number of No's recorded by the observer and dividing that by the total number of

items answered by the observer in the lesson specific fidelity check. *Participant responsiveness* was represented by the mean sums of questions five, six, and eight, and *quality of delivery* was represented by the mean sums of questions 11 through 16 on the fidelity check measure.

### **Analytic Approach**

Several techniques were used to address the goals of the study. First, descriptive statistics were used to examine frequency distribution of the variables. Second, paired t-tests were employed to compare students' pretest and posttest outcomes overall, as well as within each individual competency domain, including drugs and alcohol. Finally, multilevel analysis, also known as multilevel modeling or hierarchical linear modeling (HLM), was utilized to assess the effect of individual (gender, race/ethnicity, SES) and classroom characteristics (adherence, participant responsiveness, quality of delivery) on student behavior outcomes measured at posttest. This method was chosen because the data used for this study had cluster structure (i.e., students were nested within classrooms and classrooms were nested within schools) and it allows for individual (i.e., within-persons) and contextual (i.e., between classrooms) variations. To account for the nested structure of the data related to classrooms nested within schools, a *school* clustering variable was included in the model. Thus, standard errors were computed taking into account non-independence of observations due to school clustering (Muthén, & Muthén, 1998-2017).

The data analyses was carried out in two steps. First, bivariate conditional models with one covariate at level 2 and all predictors at level 1 were examined. Secondly, a multivariate model with all predictors at level 2 was estimated. The outcome variable in this study was the students' posttest scores. Mplus statistical software v.7.4 was used to carry out the multilevel analyses (Muthén & Muthén, 1998-2012).

## Results

### Descriptive Statistics

To test for differences in group means of the students' pretests and posttests overall, between the five competency domains, and the added domain of drugs and alcohol, paired t-tests were utilized. There was no statistically significant difference between the mean scores from pretest and posttest. Student outcomes did not change significantly after participating in LST. Items were scored so that higher means indicated better outcomes on pretest to posttest.

**Table 1. Student Pretest and Posttest Outcomes**

<u>Outcome</u>	<u>Pretest</u>		<u>Posttest</u>		t	<u>df</u>
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>		
	89.06	11.17	89.23	11.26	1.024	4,111

*Note.* \*  $p < 0.05$ .

When the competency domains were examined, results of the paired t-tests indicated there was a positive statistically significant difference between the mean scores on pretest to posttest when self-awareness, social awareness, and decision-making competency domains were examined. In other words, the mean scores from pretest to posttest increased on those three domains. There was a negative statistically significant difference on mean scores from pretest to posttest when self-management and drugs and alcohol competency domains were examined. Mean scores on those two competency domains decreased from pretest to posttest. There was no significant difference between mean scores from pretest to posttest when the relationship skills competency domain was examined. Table 1 shows the results of paired t-tests for the students'

pretest and posttests overall, and Table 2 provides total results for the specific competency domains.

**Table 2. Competency Domains Outcomes from Pretest to Posttest**

Competency Domains	Pretest		Posttest		t	df
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>		
Self-awareness	9.01	1.63	9.08	1.59	2.46*	4,111
Social awareness	17.47	3.20	17.60	3.23	2.70*	4,111
Self-management	11.75	2.24	11.57	2.25	-5.14*	4,110
Relationship skills	17.27	3.19	17.35	3.24	1.80	4,107
Decision-making	15.09	2.72	15.33	2.79	5.42*	4,105
Drugs and alcohol	18.58	2.22	18.32	2.28	-7.08*	4,093

Note. \*  $p < 0.05$ .

### Multilevel Analysis

**Level 1.** Among the individual variables, both race/ethnicity and SES were found to be statistically significant predictors of students' posttest outcomes. Specifically, students who identified as Black and Hispanic on average scored a little over 2 points lower than students who identified as White (White students were used as the constant). Asian and Multiracial race categories and gender were not significant predictors. Finally, socioeconomic status as represented by students free and reduced lunch status was a significant predictor of students'

posttest outcomes. On average, students who indicated they were on free and reduced lunch status scored approximately four points lower than students who indicated they were not on free and reduced lunch status (see Tables 3-5).

**Table 3. Adherence (Levels 1 & 2)**

Level 1		
(Individual)		
Adherence (Within Level)	Estimate	
	for the Intercept	Standard Error
Gender	0.37	0.38
Race/ethnicity		
Black	-2.37*	0.53
Hispanic	-2.40*	0.56
Asian	1.30	0.72
Multiracial	-0.50	0.54
Free and reduced lunch status	-3.70*	0.32
Level 2		
(Classroom)		
Adherence (Between Level)	Estimate	
	for the Intercept	Standard Error
	0.98	1.22

*Note.* \*  $p < 0.05$ . \*\*White was used as a reference category.



**Table 4. Participant Responsiveness (Levels 1 & 2)**

Level 1		
(Individual)		
Participant responsiveness (Within Level)	Estimate	
	for the Intercept	Standard Error
Gender	0.37	0.38
Race/ethnicity		
Black	-2.45*	0.52
Hispanic	-2.36*	0.56
Asian	1.32	0.71
Multiracial	-0.48	0.55
Free and reduced lunch status	-3.63*	0.32
Level 2		
(Classroom)		
Participant responsiveness (Between Level)	Estimate	
	for the Intercept	Standard Error
	0.26*	0.05

*Note.* \*  $p < 0.05$ . \*\*White was used as a reference category.

**Level 2.** When examining classroom variables individually, participant responsiveness and quality of delivery were found to be statistically significant predictors of students' posttest outcomes. Adherence was not found to be a statistically significant predictor. Results also indicated that a significant association between adherence, participant responsiveness, and quality of delivery with students who identified as Black and Hispanic and being on free and

reduced lunch status was present. Tables 3-5 depicts the estimate of the intercepts and standard errors at the between level for all classroom-level predictors.

**Table 5. Quality of Delivery (Levels 1 &2)**

Level 1 (Individual)		
Quality of delivery (Within Level)	Estimate	
	for the Intercept	Standard Error
Gender	0.36	0.38
Race/ethnicity		
Black	-2.48*	0.53
Hispanic	-2.37*	0.56
Asian	1.28	0.72
Multiracial	-0.49	0.55
Free and reduced lunch status	-3.67*	0.33
Level 2 (Classroom)		
Quality of delivery (Between Level)	Estimate	
	for the Intercept	Standard Error
	0.05*	0.02

*Note.* \*  $p < 0.05$ . \*\*White was used as a reference category.

In the last step, all covariates on the classroom level were included. Results indicated that participant responsiveness was the only statistically significant predictor of students' posttest outcomes. In other words, higher scores given by the observers on the participant responsiveness items of the fidelity checks can predict higher rates of behavior gain of students at posttest.

Adherence and quality of delivery were not significant predicts of students' posttest outcomes (see Table 6).

**Table 6. All Predictors Included (Level 2)**

	Level 2	
	Estimate	Standard Error
All Predictors Included (Between Level)	for the Intercept	
Adherence	-0.96	1.26
Participant responsiveness	0.37*	0.16
Quality of delivery	-0.03	0.04

*Note.* \*  $p < 0.05$ . \*\*White was used as a reference category.

## **Discussion**

The current study was designed to examine the effects of individual-level (i.e., race/ethnicity) and classroom-level (i.e., adherence) predictors on student behavior gains after completing the LST program. To date, very few studies focusing on SEL have examined implementation fidelity, its elements (i.e., adherence), and how they can affect student behavior gains (Allen et al., 2012; Durlak et al., 2011). The majority of the literature that currently exists tends to examine adherence and dosage (Durlak, 2016). The current study makes a unique contribution to the literature because it assessed two less common elements of fidelity, participant responsiveness and quality of delivery.

Generally, the results of the study indicated student race/ethnicity and SES (i.e., free and reduce lunch status) were significantly associated with student outcomes. Past research on demographic characteristics affecting student behavior gain in SEL programs seem to be mixed. However, in many cases, the literature is consistent with the findings of this study that states Black, Hispanic, and lower SES students are likely to be at higher risk for obtaining poorer social and emotional outcomes when compared to White students (Garner, Mahatmya, Brown, & Vesley, 2014; Castro-Olivo, 2014). More specifically, in a study evaluating the effectiveness of the Resolving Conflict Creatively intervention program, Black students' prosocial behavior increased over time at a slower rate than White students, however, no significant differences were found between Hispanic and White students (Aber, Brown, Jones, 2003). Frequently, program effects are not disaggregated by demographic characteristics, which can make

determining if a SEL program can yield equal gains among its participants problematic (Garner et al., 2014).

Results of the study indicated that gender was not significantly associated with student outcomes. Generally, gender is thought to be associated with emotional intelligence and that girls tend to express more positive emotion and can regulate their emotions of that superior to boys (Garner, et al., 2014). However, the results tend to be inconsistent across SEL studies. This could have to do with biased gender stereotypes that are present in self-report ratings leading individuals to answer questions based on their gender (Lopez-Zafra & Gartzia, 2014).

Overall, Black, Hispanic, and lower SES students consistently scored lower when compared to White students while examining each core element separately, as well as when all elements were included in the multilevel model. This study took demographic characteristics into consideration and found that not all students were reporting the same positive outcomes. This suggests adaptations may need to be made to LST components and/or how the teachers are trained in order to increase these specific students' outcomes.

On the classroom level of the multilevel model, participant responsiveness and quality of delivery were significantly associated with student behavior gains when the core elements were evaluated separately. However, when all individual and classroom-level variables were included, only participant responsiveness was significantly associated with student outcomes. These findings are not consistent with the existing literature that most often states adherence to be the most significant indicator of student outcomes (Mihalic et al., 2008; Durlak, 2016). Given the unique design of this study, this may have made participant responsiveness more significant as compared to other core elements. More time and increased training for observers might be needed in order to accurately assess dosage and adherence.

Generally, it can be concluded that higher scores given by the observers on the participant responsiveness items of the fidelity checklist can predict higher rates of student behavior gain at posttest. While delivering LST, it is essential to ensure students are actively engaged, understand the material, and are participating in the lesson. Teachers may need to modify their delivery methods and/or examples used to explain a topic in order for it to be applicable to the students within their classrooms. Results from this study indicated that participant responsiveness is a key element to implementation fidelity and can be linked to what fits the students' experiences.

From a theoretical standpoint, quality of delivery can still be considered a key factor to implementation fidelity. Mihalic, Fagan, and Argamaso (2008) examined quality of implementation by looking at teaching techniques (e.g., discussion, skill demonstration, and behavioral rehearsal). The results indicated that teachers' use of interactive techniques was positively correlated with good student behavior, and therefore, could be more likely to lead to knowledge and behavior acquisition. Quality of delivery can be directly tied back into participant responsiveness. High quality delivery can potentially mediate student behavior or misbehavior and can lead to students being more actively engaged, have a better understanding of the material, and increase responsiveness during a lesson.

Research has shown there can be significant variability of implementation fidelity attained across providers (i.e., classrooms or teachers) within the same study (Durlak & DuPre, 2008). According to Durlak and DuPre (2008), it is not uncommon for implementation levels to vary 20 to 40% between sites and classrooms and teachers. More research is needed to narrow down how the core elements are measured in order to obtain a better consensus on results among LST sites and classrooms.

## **Limitations**

Limitations of the study should be noted. First, the study was not able to make an accurate comparison of knowledge gain of the students due to the knowledge section being added to the LSTM at posttest. Therefore, true knowledge acquisition was impossible to capture. Future research should address this limitation by ensuring the instruments of measurement are completed before being administered at baseline. Second, there seemed to be inconsistencies amongst observers in the way they completed the fidelity checklists. Moving forward, it would be beneficial to ensure all observers are completing the checklist in the same manner. Third, the competency domain subscales showed to have weak reliability when Cronbach's alpha was assessed. Therefore, the LSTM may benefit from modification targeting the competency domains to increase the reliability of the subscales. Fourth, the study sample only included 16 out of the total 48 schools where LST was implemented. Thus, it is unclear to what extent the results of this study can be generalizable to the entire school district and the general population.

## **Implications for Behavioral Health**

**Implications for Research.** The findings of this study offer useful information on future directions for research, practice, and policy. Future research would benefit from including all of the core elements of implementation fidelity and how they can affect student outcomes in LST. Adherence and dosage seem to be the most common focus, however, as this study indicated, other elements may be equally as important to student outcomes (i.e., participant responsiveness). Another valuable research initiative would be to examine how adaptations affect student outcomes. The literature indicates that adaptations can result in both positive and negative outcomes (Durlak, 2016). While it is essential to keep the core components of a SEL

program intact, making adaptations to obtain the right “fit” for students, teachers, and a school could potentially increase the likelihood of positive outcomes.

In addition to evaluating adaptations of the program, a closer examination of the individuals observing the classrooms and completing the checklists should be added to the research agenda, as well as having an universal method to measure the core elements of fidelity. Having the observers and the fidelity measure be consistent not only across the LST program, but also with other school-based SEL programs, could help combat the inconsistencies that are seen in the literature.

**Implications for Practice.** Poor implementation can result in large economic losses for schools and can make it nearly impossible to interpret results properly. In turn, this could hinder individuals, schools, and educational leaders from advocating for policy and guidelines that could bring effective programs into their districts (Durlak, 2016). More importantly, lack of implementation fidelity can also lead to poor student outcomes. It is essential to evaluate whether a school or school district is ready and committed to take on an SEL initiative. A school’s culture, climate, willingness to accept new innovation (i.e., buy-in), leadership, coaching, and understanding of the program goals and mission are important components to consider before implementation takes place (Vroom, 2016; Wanless & Domitrovich, 2015). Taking the initiative to facilitate these components before the program takes place can lead to higher quality implementation, as well as significant student outcomes.

It is also important to ensure the schools and teachers have enough time and resources to implement the program in its entirety. Shortening trainings and/or the length of the program could affect the quality of implementation and reduce the likelihood of knowledge and behavior acquisition among students. Teacher buy-in is essential and in many cases can be hard to obtain.



Since the passing of the No Child Left Behind Act in 2001, intense pressure has been put on teachers to ensure strong academic performances from their students. In turn, teachers can be more apt to focus on the core curricula, and a new SEL program can receive less attention (Vroom, 2016; Reyes, Brackett, Rivers, Elbertson, & Salovey, 2012). Finally, ensuring all teachers and program champions are trained properly and are provided assistance when needed is essential for LST to be successful. If the core components of the program are not presented correctly or at all, the quality of delivery and examples used do not fit the classroom, and participants are not actively responsive to the lesson, thus, the predicted behavior gains are less likely to materialize.

**Implications for Policy.** The majority of federal initiatives and policy surrounding SEL have narrowly focused on suspension and expulsion prevention, as well as substance use (i.e., problem behaviors) (Zaslow, Mackintosh, Mancoll, & Mandell, 2015; Office of Early Learning, n.d.). Although problem behaviors are an important component of SEL programs, it is expected that initiatives that have problem behaviors as the sole focus will not produce the full range of benefits when compared to approaches that are comprehensive and universal (Zaslow et al., 2015).

In January of 2015, the Supporting Emotional Learning Act (SELA) was introduced to the House of Representatives (Supporting Emotional Learning Act [SELA], 2015). Under this bill, the National Center for Educational Research would be required to carry out research on SEL education and educating teachers on evidence-based assessment tools and teaching methods. This bill was a good example of what SEL legislation should incorporate; however, no further action has been taken on this bill since 2015 (Zaslow, 2015; SELA, 2015). It is important that education and coaching for teachers and not-for-profits and consideration for

implementation fidelity are included in federal and state legislation. The adoption of widespread initiatives and effective prevention programs will have little to no effect until the quality of implementation by not-for-profits and teachers can be ensured (Mihalic, Fagan, & Argamaso, 2008).

## References

- Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology, 39*(2), 324-348.
- Allen, J. D., Linnan, L. A., & Emmons, K. M. (2012). Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. In R. C. Brownson, G. A. Colditz, & E. K. Proctor (Eds.), *Dissemination and implementation research in health: Translating science to practice* (pp. 281-305). New York: Oxford University Press.
- Blueprints for Healthy Youth Development. (n.d.). *About us*. Retrieved from <http://www.blueprintsprograms.com/about>.
- Botvin, G. J., Epstein, J. A., Baker, E., Diaz, T., & Ifill-Williams, M. (1997). School-based drug abuse prevention with inner-city minority youth. *Journal of Child & Adolescent Substance Abuse, 6*(1), 5-19.
- Botvin, G. J., & Griffin, K. W. (2004). Life Skills Training: Empirical findings and future directions. *Journal of Primary Prevention, 25*(4), 211-232.
- Botvin, G. J., Schinke, S. P., Epstein, J. A., & Diaz, T. (1994). Effectiveness of culturally focused and generic skills training approaches to alcohol and drug abuse prevention among minority youths. *Psychology of Addictive Behaviors, 8*(2), 116-127.
- Botvin LifeSkills® Training. (n.d.). *LST fidelity checklists*. Retrieved from [https://www.lifeskillstraining.com/lst\\_process\\_tools.php](https://www.lifeskillstraining.com/lst_process_tools.php).

- Brackett, M. A., Elbertson, N. A., & Rivers, S. E. (2015). Applying theory to the development of approaches to SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds), *Handbook of social and emotional learning: Research and practice* (pp. 20-32). New York: The Guilford Press.
- 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.
- Castro-Olivo, S. M. (2014). Promoting social-emotional learning in adolescent Latino ELLs: A study of the culturally adapted Strong Teens program. *School Psychology Quarterly*, 29(4), 567-577.
- Collaborative for Academic, Social, and Emotional Learning (CASEL). (n.d.). *What is SEL?* Retrieved from <http://www.casel.org/what-is-sel/>.
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, 46(3), 333-345.
- Durlak, J. A. (2015). What everyone should know about implementation. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds), *Handbook of social and emotional learning: Research and practice* (pp. 395-405). New York: The Guilford Press.
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327-350.
- 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.

- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237-256.
- Garner, P. W., Mahatmya, D., Brown, E. L., & Vesley, C. K. (2014). Promoting desirable outcomes among culturally and ethnically diverse children in social emotional learning programs: A multilevel heuristic model. *Educational Psychology Review, 26*, 165-189.
- Klapp, A., Belfield, C., Bowden, B., Levin, H., Shand, R., & Zander, S. (2017). A benefit-cost analysis of a long-term intervention on social and emotional learning in compulsory school. *The International Journal of Emotional Education, 9*(1), 3-19.
- Lopez-Zafra, E., & Gartzia, L. (2014). Perceptions of gender differences in self-report measures of emotional intelligence. *Sex Roles, 70*, 479-495.
- Mihalic, S. F., Fagan, A. A., & Argamaso, S. (2008). Implementing the LifeSkills Training drug prevention program: Factors related to implementation fidelity. *Implementation Science, 3*(5), 1-16.
- Muthén, L.K. and Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*(1), 85-112.
- Pike, G. R., & Rocconi, L. M. (2012). Multilevel modeling: Presenting and publishing the results for internal and external constituents. *New Directions for Institutional Research, 154*, 111-124.
- Office of Early Learning. (n.d.). *Position statement on expulsion and suspension prevention in early childhood*. Retrieved from

[http://www.floridaearlylearning.com/sites/www/Uploads/files/Statewide%20Initiatives/Health%20and%20Safety/Expulsion%20Policy%20Position%20Statement\\_ADA.PDF](http://www.floridaearlylearning.com/sites/www/Uploads/files/Statewide%20Initiatives/Health%20and%20Safety/Expulsion%20Policy%20Position%20Statement_ADA.PDF)

- 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.
- Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools, 49*(9), 892-909.
- Spoth, R., Gyll, M., Trudeau, L., & Goldberg-Lillehoj, C. (2002). Two studies of proximal outcomes and implementation quality of universal preventive interventions in a community-university collaboration context. *Journal of Community Psychology, 30*(5), 499-518.
- Supporting Emotional Learning Act of 2015-2016, H.R. 497, 114<sup>th</sup> Cong. (2015). Retrieved from <https://www.congress.gov/bill/114th-congress/house-bill/497/all-actions>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55.
- Vroom, E. B. (2016). *Implementation, barriers, and successes of social and emotional learning programs in schools*. Unpublished manuscript.
- Wanless, S., & Domitrovich, C. (2015). Readiness to implement school-based social-emotional learning interventions: Using research on factors related to implementation to maximize quality. *Prevention Science, 16*(8), 1037-1043.

- Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social and emotional learning: Past, present, and future. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds), *Handbook of social and emotional learning: Research and practice* (pp. 3-19). New York: The Guilford Press.
- Wright, A., Lamont, A., Wandersmand A., Osher, D., & Gordon, E. S. (2015). Accountability and SEL programs: The Getting to Outcomes® approach. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds), *Handbook of social and emotional learning: Research and practice* (pp. 500-515). New York: The Guilford Press.
- Zinsser, K. M., & Dusenbury, L. (2015). Recommendations for implementing the new Illinois Early Learning and Development Standards to affect classroom practices for social and emotional learning. *Early Childhood Research & Practice, 17*(1).
- Zaslow, M., Mackintosh, B., Mancoll, S., & Mandell, S. (2015). Federal policy initiatives and children's SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds), *Handbook of social and emotional learning: Research and practice* (pp. 549-565). New York: The Guilford Press.

**Appendix: Institutional Review Board (IRB) Determination**



RESEARCH INTEGRITY AND COMPLIANCE  
Institutional Review Boards, FWA No. 00001669  
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799  
(813) 974-5638 • FAX (813) 974-7091

April 10, 2017

Enya Vroom  
CFBH-Child and Family Behavioral Health  
Tampa, FL 33612

RE: Not **Human Subjects Research Determination**  
IRB#: Pro00030214  
Title: The Impact of Implementation Fidelity on Middle School Student Outcomes in the Life Skills Training Program

Dear Ms. Vroom:

The Institutional Review Board (IRB) has reviewed your application and determined the activities do not meet the definition of human subjects research. Therefore, this project is not under the purview of the USF IRB and approval is not required. If the scope of your project changes in the future, please contact the IRB for further guidance.

All research activities, regardless of the level of IRB oversight, must be conducted in a manner that is consistent with the ethical principles of your profession. Please note that there may be requirements under the HIPAA Privacy Rule that apply to the information/data you will utilize. For further information, please contact a HIPAA Program administrator at 813-974-5638.

We appreciate your dedication to the ethical conduct of research at the University of South Florida. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

A handwritten signature in black ink, appearing to read "Kristen Salomon", written over a horizontal line.

Kristen Salomon, Ph.D., Vice Chairperson  
USF Institutional Review Board