SOCIAL AND EMOTIONAL LEARNING AS A UNIVERSAL LEVEL OF SUPPORT: EVALUATING THE FOLLOW-UP EFFECT OF *STRONG KIDS* ON SOCIAL AND EMOTIONAL OUTCOMES

by

JASON E. HARLACHER

A DISSERTATION

Presented to the Department of Special Education
and Clinical Sciences
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2009

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Jason Harlacher

Title:

"Social and emotional learning as a universal level of support: Evaluating the follow-up effect of Strong Kids on social and emotional outcomes"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Special Education and Clinical Sciences by:

Kenneth Merrell, Chairperson, Special Education and Clinical Sciences
Robert Horner, Member, Special Education and Clinical Sciences
Tary Tobin, Member, Special Education and Clinical Sciences
Lynn Kahle, Outside Member, Marketing

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 13, 2009

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

An Abstract of the Dissertation of

Jason E. Harlacher for the degree of Doctor of Philosophy

in the Department of Special Education and Clinical Sciences

to be taken June 2009

Title: SOCIAL AND EMOTIONAL LEARNING AS A UNIVERSAL LEVEL OF SUPPORT: EVALUATING THE FOLLOW-UP EFFECT OF *STRONG KIDS* ON SOCIAL AND EMOTIONAL OUTCOMES

Approved: _____
Dr. Kenneth W. Merrell

The present study examined the initial and follow-up effect of *Strong Kids*, a social and emotional learning (SEL) curriculum, among a sample of 106 third and fourth graders. Students were assigned by classroom to either the treatment or wait-list condition, and completed questionnaires on SEL knowledge (*Strong Kids* Knowledge test) and perceived use of SEL skills (the Coping Scale, Social and Emotional Assets and Resiliency Scale) across 3 assessment periods (pre-testing, post-testing, and follow-up). The classroom teachers also completed a social functioning questionnaire (the School Social Behavior Scales- 2nd edition) on each student at each assessment period. The classroom teachers implemented 12 weekly lessons across a 3-month time period and 1

booster session approximately 1 month after the last lesson. They also promoted generalization of SEL skills by providing praise and pre-correction to students on the SEL skills they were learning. Analyses revealed that the treatment group had greater positive gains across all of the dependent measures from pre-test to post-test. These gains maintained at the 2-month follow-up period, providing preliminary evidence of the preventative quality of *SK*. The results are discussed within the broader framework of a three-tiered model of support for SEL, and the possibility of using *SK* as a universal level of support within school.

CURRICULUM VITAE

NAME OF AUTHOR: Jason Harlacher

PLACE OF BIRTH: York, PA

DATE OF BIRTH: September 15, 1977

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon; Eugene
Utah State University; Logan
Ohio University; Athens

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2009, University of Oregon
Master of Science, School Psychology, 2006, Utah State University
Bachelor of Arts, Psychology, 1999, Ohio University

AREAS OF SPECIAL INTEREST:

Systems-level Changes in Schools
School-wide Models of Support
Social and Emotional Learning
Psychoeducational Assessment

PROFESSIONAL EXPERIENCE:

School Psychologist
Sweet Home School District, Sweet Home, OR, Oct 2007-present
Lane Education Service District, Eugene, OR, Aug 2006-present
4J School District, Eugene, OR, April 2006-2007
Davis School District, Farmington, UT, Aug 2003-Aug 2005

Psychological Testing, Ogden School District, Ogden, UT, 2002-2003

Social Worker II, BA, Clark County Mental Health Services: Youth Challenges, Springfield, OH, 1999-2001

Teaching Assistant, University of Oregon, 2005-2007

Psychology Instructor, Utah State University, 2003-2005

Psychology Teaching Assistant, Utah State University, 2002-2003

Research Team Member, Oregon Resiliency Project, University of Oregon, Eugene, 2005-2008

Research Assistant, Early Intervention Research Institute, Logan, UT, 2001-2002

Crisis Hotline Volunteer, Careline, Athens, OH, 1998

GRANTS, AWARDS AND HONORS:

Association of School Psychology Students (ASPS) President, University of Oregon, 2007-2008

Project Integrate Training Grant, University of Oregon, 2006-2007

Student Advisory Board Representative, University of Oregon, 2005-2007

Effective Educational Practices Scholarship, University of Oregon, 2006

Ohio War Orphan's Scholarship, Ohio University, 1999

Dean's List Scholarship, Ohio University, 1999

Psi Chi National Psychology Honor Society, Ohio University, 1997-1999

Air Force Aid Society Grant, Ohio University, 1995-1999

PUBLICATIONS:

Harlacher, J. E., Ling, C., & Siler, C. E. (2008). A survey of K-12 staff perceptions of bullying: Influences of school level and bullying types with implications for interventions. Manuscript submitted for publication.

Barnes, A. C., & Harlacher, J. E. (in press). Response-to-Intervention as a set of principles: Clearing the confusion. *Education & Treatment of Children.*

Merrell, K. W., & Harlacher, J. E. (2008). Behavior rating scales. In R. P. Archer & S. R. Smith (Eds.), *Personality assessment* (pp. 247-280). New York: Routledge/Taylor & Francis.

Harlacher, J. E., Roberts, N. E., & Merrell, K. W. (2006). Classwide interventions for students with ADHD: A summary of teacher options beneficial for the whole class. *Teaching Exceptional Children, 39*, 6-12.

ACKNOWLEDGMENTS

I wish to express my sincere gratitude and appreciation first to Professor Merrell for his assistance and support throughout this process. He provided sound advice, timely feedback, and professional judgment in formulating this dissertation from the beginning to end. His support and enthusiasm helped to see this project through and I am thankful for his commitment. I also must thank Dr. Merrell and the Oregon Resiliency Project for the contributions and funding of this study, as it would not have been possible without that support.

Thanks also go out to Yolanda Elementary and the teachers who participated in the study. Thank you to Amy Paschal, Nicki Gorham, Jill Olson, and Lisa Gibson for allowing me to conduct a study within your respective classrooms. Your commitment to the study and implementation of *Strong Kids* has provided a valuable service to the Oregon Resiliency Project and the academic community at large. A special thank you to Julie Collins, who first responded to my requests to use Yolanda Elementary as part of the study.

A special thank you to fellow school psychology students and members of the Oregon Resiliency Project team who participated in data collection: Kelly Laugle, Nancy Nelson, Sarah Whitcomb, Brad Cohn, Candice Ling, Josh Felver-Gant, Katie Ravitch, Kara Tom, Chelsea Siler, Aaron Barnes, and Rebecca Briggs.

For my parents and family, for
they have pushed me to make them proud.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

PROBLEM STATEMENT

Students in today's schools are faced with difficult and challenging circumstances. For example, it's estimated that 13% of children are currently living in poverty (United States Census Bureau, 2000). Students are also coming to terms with growing numbers of mental health disorders, social-emotional problems, and more intense and difficult behavioral problems (Greenberg et al., 2003; Sprague & Horner, in press). In fact, 1 in 5 youths meet criteria for a diagnosable psychiatric condition, and 1 in 10 meet criteria for a severe emotional disturbance (Hoagwood & Johnson, 2003). Approximately 12 out of every 1,000 children experience one form of maltreatment in the form of neglect or physical, emotional, or sexual abuse (United States Department of Health & Human Services, 2006), and 50% of youths who meet criteria for one mental health disorder meet criteria for a second condition (Kazdin, 2004). These dire circumstances that many children face cannot be underestimated. Risk factors such as low SES, family discord, and weak attachment to school are linked to lower academic proficiency, increased rates of delinquency, and diminished capacity for wages (Doll & Lyon, 1998). The financial burdens single-parent homes face leads to less adult supervision as parents work more than one job and swing or graveyard shifts (Weissberg & O'Brien, 2004). Additionally, children who have been maltreated are at risk for developing various psychological problems (Turner, Finkelhor, & Ormrod, 2006), and

the presence of certain mental health diagnoses are associated with poor social relationships with peers, low academic performance, and difficulty maintaining employment later in life (American Psychological Association, 2004; Mash & Barkley, 1998). Perhaps most telling of how such stressors can impact our youth is the disheartening statistic that suicide is the third leading cause of death among teenagers (National Center for Injury Prevention and Control, 2004).

Naturally, the challenges that youths deal with places immense pressure on schools to provide more than just academic instruction (Greenberg et al., 2003). Rising health care costs, limited access to community services, and sporadic use of evidence-based practices are a few reasons students do not receive the care and support that they need (Greenberg et al.; Hoagwood & Johnson, 2003; Strein, Hoagwood, & Cohn, 2003). School personnel today are being asked, if not forced, to provide extensive services to students including after-school supervision, wrap-around services, and counseling and mental health treatment (Weisberg & O'Brien, 2004). Arguably, schools have become the "de facto" mental health delivery system, as of the mere 20% of children who receive mental health treatment, 70 to 80% of them receive treatment in the schools (Hoagwood & Johnson). As schools continue to evolve in the 21st century, the traditional view in which academics is the sole focus is becoming less and less.

While schools juggle the demands of a student body that is more difficult and challenging than in previous years (Merrell, Irvin, & Gimpel, 2006), they must also meet federal mandates of academic standards (No Child Left Behind, n.d.). The task of providing students with skills beyond academic instruction does not mean schools must

sacrifice one skill over another, however, as a solid social and emotional skill base can serve the foundation for better educational achievement (Zins, Weissberg, Wang, & Walberg, 2004). Healthy social and emotional skills are associated with improved academic outcomes (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; Catalano et al., 2003), better academic self-efficacy and self-determination (Doll, Zucker, & Brehm, 2004), and a stronger connection to school (Zins, Weissberg et al., 2004). In fact, some researchers view social and emotional learning (SEL) as a priority similar to academic proficiency (CASEL, n.d.). Teaching SEL concurrently with academics is considered the key to providing students with all they need to succeed in life. Arguably, it prepares them to be "lifelong" learners who are able to cope with the taxing stressors they may deal with (Greenberg et al., 2003; Zins, Weissberg et al., 2004). It is possible for schools to meet academic proficiency standards while also teaching social and emotional skills, thus equipping students with coping skills that can ensure *resiliency* (i.e., the ability to achieve goals or develop competence in the face of adversity; Doll & Lyon, 1998).

The phrase SEL is a new term, but includes programs that are called by various other names, such as *violence prevention*, *character education*, and *positive youth development*. All of these terms refer to a movement that focuses on increasing the social and emotional competencies of students and reducing the problematic behaviors students display (CASEL, n.d.). Programs currently exist that schools can adopt to teach children and adolescents skills that provide critical SEL instruction and are complimentary to academic success. Even a cursory look at the literature produces numerous books and

programs that promise social success for students. Unfortunately, many of these programs may not be evidenced-based and as schools take on more programs and extend their domain of service, they can become more fragmented in their service delivery. In turn, this leads to less effective practices, poor fidelity, and high staff turnover (Elias, Zins, Graczyk, & Weissberg, 2003). Schools have the capacity to meet all of the needs of students, but if they are to make worthwhile changes that maintain over time without burdening the system, schools should coordinate their services on a school-wide level and use a systems-level approach (Sugai & Horner, 2006). Recent research argues for a comprehensive, empirically-based approach to services, and the result of this push is the conceptualization of SEL programs housed within a three-tiered model of increasing support (Greenberg et al., 2003; Elias et al., 2003). As with Response-to-Intervention models and Positive Behavior Support, SEL programming can adopt a universal, "core" curriculum and then build upon that curriculum with increasing levels of support for students, creating a coordinated system that can sustain itself.

A promising program developed by researchers from the University of Oregon, the *Strong Kids (SK)* series, is a universal curricula that can provide the foundation of SEL skills from which schools can build a multi-tiered model of support (Merrell, Carrizales, Feuerborn, Gueldner, & Tran, 2007). Preliminary research has been favorable, as students who have been exposed to the *SK* series have shown improvements in crucial knowledge of social-emotional concepts and coping skills, and reductions in symptoms of negative affect (e.g., Castro-Olivo, 2006; Gueldner, 2006; Merrell, Juskelis, Tran, & Buchanan, 2006). However, more research is needed to better determine the exact impact

of the *SK* series as a universal level of support, particularly as a preventative and protective element in a child's life. No study to date has yet to explore how well the *SK* curricula can create beneficial gains in students over time. Improvements in outcomes have been documented at the conclusion of the *SK* curriculum, but no research has examined moderate or long-term follow-up.

As such, the primary goal of the current study was to determine if the use of a SEL curriculum (i.e., *SK*) leads to better social and emotional functioning over a moderate time period (i.e., approximately 6 months), thus providing insight into its protective quality against later negative outcomes. In determining this goal, the study investigated the (a) increase in knowledge of SEL skills, (b) the increase in self-perceived use of SEL skills, (c) the maintenance of those gains over time, and (d) the increase in social functioning over time in a treatment group versus a control group. The specific research questions were:

1. Did the use of a SEL curriculum lead to an increase in the knowledge of SEL skills?

2. Did the use of a SEL curriculum lead to an increase in the self-perceived use of SEL skills?

3. Did those gains in knowledge and perceived use of SEL skills maintain over time at a spring follow-up?

4. Did the use of a SEL curriculum result in any preventative effects for the students? Specifically, did the use of a SEL curriculum in the fall lead to better social functioning in the spring, relative to a group who did not receive the SEL curriculum?

CHAPTER II

LITERATURE REVIEW

The literature review section has two goals, the first of which is to define social and emotional learning (SEL) and to identify the specific skills that are considered essential to SEL. In doing so, any evidence linking SEL and its impact on various outcomes (e.g., internalizing & externalizing behaviors, social functioning, academic outcomes) are presented. The second goal of the literature review is to summarize the current state of delivery of school-based SEL programs. A three-tiered model of supported is discussed, with special attention devoted to implementing *Strong Kids* as a universal level of support for all students within a school. This review is not intended to be exhaustive, but to provide a basic framework for understanding the literature base related to the current study.

Social and Emotional Learning

Definition of Social and Emotional Learning

Because researchers vary slightly on the terminology used and ways to describe SEL, actually defining it may seem like a nebulous effort. At first glance, capturing the parameters of SEL may be overwhelming and this point is no more salient than when we compare definitions. McCombs (2004) emphasized that SEL is "*integrating thinking, feeling, and behavior to achieve important social tasks"* (p.27), yet Johnson and Johnson (2004) defined SEL as "*the mastery of interpersonal and small-group skills, and the*

internalization of prosocial attitudes and values" (p.40). An even broader definition is offered by the Collaborative for Academic, Social, and Emotional Learning (CASEL, n.d.), calling SEL "*the process of acquiring the skills to recognize and manage emotions, develop caring and concern for others, make responsible decisions, establish positive relationships, and handle challenging situations effectively*". Such variation need not be disheartening, however. It is true that researchers may emphasize differing aspects of SEL and identify different skill sets of SEL, but there is general agreement as to what SEL is (see Zins, Weissberg, et al., 2004). Simply put, SEL refers to the process by which people learn to manage and navigate the social and emotional aspects of their lives (Cohen, 2001; CASEL, n.d.). Even the most routine, everyday life activities can bring frustration, differing opinions, and set-backs. SEL is the process in which we learn how to react and deal with these obstacles in order to accomplish our goals.

With such a broad definition of SEL, identifying the core set of skills that make up SEL is an important first step. There is no shortage of identified skills that fall under the construct of SEL; in fact, six different researchers have identified 10 different SEL concepts, only 6 of which were commonly identified by more than one researcher (Cohen, 2001; Denham & Weissberg, 2004; Doll et al., 2004; Goleman, 1995; Payton et al., 2000; Zins, Bloodworth, Weissberg, & Walberg, 2004). With such range, how is the typical educator to know what components are essential to SEL and which ones are not? Fortunately, the breadth of identified skills does not hinder the conceptualization of SEL. When the totality of the research is considered, common elements that make up SEL do emerge, allowing one to make a clear definition of the concepts and skills that make up

SEL (Cohen; Denham & Weissberg, 2004; Goleman; Zins, Bloodworth et al., 2004). These common elements are presented next, followed by a section reviewing how those elements may impact a youth's overall level of functioning.

Table 1

Social and Emotional Learning Skills

Skill	Description	Example
Self-awareness	Ability to identify one's own emotions, values, and strengths.	A student recognizes that he is feeling frustrated while working on math, a subject he knows he struggles with.
Social awareness	Ability to recognize other people's emotions and how emotional expression affects certain situations.	A student playing basketball with friends realizes that a teammate plays harder and is happier when they cheer him on.
Self-management	Ability to monitor and control one's emotions, impulses, and behavior in order to achieve goals.	A student takes deep breathes while taking a test in order to relax and finish the test.
Social management	Managing one's emotions and behaviors in order to achieve goals in social settings.	Students use negotiation and cooperation in order to complete a group assignment before class ends.
Responsible decision making	Making decisions that lead to healthy and safe results.	A student plans out a timeline for a large essay due at the end of the school year.

Components of SEL

Within the SEL literature, five skills are consistently identified by various researchers (see Table 1). Emotional recognition and expression are central tenets to SEL

(Denham & Weissberg, 2004), so it's no surprise that *self-awareness* is the most common skill mentioned among SEL researchers. Self-awareness is anchored by the ability to identify and recognize one's emotions, but this awareness also includes recognizing one's own strengths, values, and efficacy (Doll et al., 2004; Zins, Bloodworth et al., 2004).

Similar to self-awareness is *social awareness*, the ability to read and recognize emotions in others and knowledge about how emotional expression can affect certain social situations (Denham & Weissberg, 2004). Although self-awareness and social awareness may involve similar components, such as recognizing and labeling emotions, the ability to be socially aware is a separate skill because it includes components distinct from self-awareness, such as assuming different perspectives and showing empathy and understanding for others (Goleman, 1995; Zins, Bloodworth et al., 2004).

After mastering social and self-awareness, it is logical that the next two SEL skills involve managing that awareness. *Self-management* is the ability to regulate and monitor one's emotions, behavior, and impulses in order to achieve desired goals (CASEL, n.d.; Doll et al., 2004; Weissberg & O'Brien, 2004). Children competent with this skill are able to delay impulses, foresee the consequences of their actions, and motivate themselves in light of mitigating factors, such as frustration while performing a task (Zins, Bloodworth et al., 2004). *Social management*, also referred to as "relationship skills" (CASEL; Weissberg & O'Brien, 2004) or "relationship management" (Zins, Bloodworth et al., 2004), is the process of establishing and maintaining rewarding relationships. Specifically, social management involves the ability to effectively communicate, cooperate and negotiate with others, manage and read one's own and

other's emotions in social settings, deal with and resolve conflicts that arise, and seek and provide help to others (CASEL; Zins, Bloodworth, et al.).

Finally, *responsible decision-making* is the ability to make healthy and responsible decisions that lead to beneficial results or solutions. Although Cohen (2001) labels this skill problem-solving, responsible decision making is a broad category in which problem-solving is a part of. Responsible decision-making centers around taking responsibility for one's decisions and being able to evaluate factors and consequences associated with given decisions (Weissberg & O'Brien, 2004). This characteristic includes the ability to analyze situations, such as evaluating the danger of attending an unsupervised party, taking personal responsibility, such as owning up to a mistake, and making healthy and safe choices, such as being honest or avoiding the use of illicit drugs. Responsible decision-making involves more than a superficial knowledge of SEL skills and evokes the actual *use* of them, expanding upon identification of emotions, for example, and focusing on how that information can be used to make important and worthwhile choices (Denham & Weissberg, 2004).

To summarize, the five skills most consistently identified are: *self-awareness, social awareness, self-management, social management*, and *responsible decision-making*. Although each skill offers unique components, each one is complimentary to each other and students may even use all of the skills simultaneously when maneuvering through the social and emotional situations they face daily. Having presented the key components of SEL, the next section distinguishes the differences between SEL skills

and social skills. The section thereafter investigates the effect SEL programs can have on students.

Social Skills versus Social and Emotional Learning Skills

The difference between *social skills* (SS) and *social and emotional learning skills* (SELS) is slight, as there is considerable overlap between the skills captured within each domain. The most obvious difference may be that SELS includes a focus on emotions and the internal processes associated with developing social and emotional skills. Whereas SS are defined as the discrete behaviors that result in positive social interactions (Caldarella & Merrell, 1997; Gresham, 2002), SELS are broader, focusing on both discrete skills *and* internal processes. SELS include the developmental process by which individuals learn to understand and manage their relationships to others, including the cognitive, affective, and behavioral changes that accompany this process. A simple way to distinguish between the two is to conceptualize SEL skills as an umbrella under which SS (and emotional skills) fall.

An example is presented to distinguish between SS and SELS. Imagine a student who is extremely anxious and has difficulty initiating conversations with classmates. The social skills involved in this situation (ones the student lacks) would be those behaviors that lead to positive outcomes for the student, such as introducing one's self, choosing a topic of conversation, and perhaps even dressing a particular way. Social skills instruction would therefore focus on teaching those behaviors. The SELS involved include those behaviors defined as social skills (subsumed under the category of *social management* listed in Table 1), but it also includes awareness of and the subsequent

management of the emotional aspects of the situation (*self-awareness* and *self-management* in Table 1). Thus, SEL would incorporate awareness of the anxiety around starting a conversation and perhaps teaching relaxation and cognitive exercises to manage that emotion. Social skills are the discrete behaviors that lead to desirable social interactions (Gresham, 2002), but SELS is more board, concerned with both the social *and* emotional aspects of a situation (CASEL, n.d.).

SEL and Outcomes

Effect of SEL Programs on Outcomes

Researchers from CASEL and developers of various programs, such as the Promoting Alternative THinking Skills (PATHS) curriculum and the Resolving Conflicts Creatively Program, have been documenting the connection between SEL and various outcomes pertinent to healthy functioning for several years now (see Zins, Weissberg et al., 2004). This expanding scientific base is building momentum as more and more research continues to emerge indicating that SEL programs are able to promote resiliency, decrease problematic behaviors in students, and contribute to improved academic performance (Greenberg et al., 2003; Zins, Weissberg et al., 2004). The present study aimed to add to this evidence base, as its overarching goal was to investigate how using a SEL program (i.e., *SK*) can protect children in light of the risk factors and stressors that they may encounter. This goal was measured by documenting gains in knowledge and self-perceived use of SEL skills, and then investigating the course of social functioning over a moderate time period (i.e., 6 months). The next several

paragraphs briefly review recent research, providing a summary of the documented results of SEL programs thus far.

Review of Previous Meta-analyses of SEL Programs

To date, a few reviews and meta-analyses have concluded that school-based programs that promote SEL skills are effective in decreasing problematic behavior in students while simultaneously increasing academic achievement and social competencies (Durlak & Wells, 1997; Greenberg, Domitrovich, & Bumbarger, 2001; Wilson, Gottfredson, & Najaka, 2001). Durlak and Wells (1997) identified 177 school-based prevention programs across three categories: *environment-centered* (e.g., parent training), *transition programs* (e.g., divorce support groups, first-time mothers), and *person-centered programs* (e.g., affective education, problem-solving). The authors concluded that prevention programs, in general, result in small gains in improving academic achievement and decreasing internalizing and externalizing behaviors (mean effect sizes (ESs) ranged from 0.30 to 0.32). (Cohen (1988) classifies 0.20, 0.50, and 0.80 as small, medium, and large effects, respectively). Examining the 46 programs identified as affective education (i.e., awareness of emotions and their expressions), the authors concluded that such programs resulted in robust gains in improving competencies (e.g., self-assertion, communication skills) and decreasing problematic behavior (e.g., anxiety/depression, externalizing behavior problems). However, these findings are moderated by age, as children 7 and under experienced moderate to large gains (ES of 0.85 for competencies and 0.69 for problematic behavior) and children 8 and up experienced modest gains (ESs ranged from 0.21 to 0.41). Although the analyses

conducted by Durlak and Wells provided general information on the effect of prevention programs and not more discrete information, they were able to document that school-based prevention programs that contain elements synonymous with *SK* can result in positive outcomes for students.

Wilson and colleagues (2001) add more credence to SEL programs impacting student outcomes with their meta-analysis of school-based prevention programs. The authors reviewed results from 165 prevention activities on four outcomes: delinquency, alcohol/drug use, nonattendance, and conduct behaviors (e.g., aggressive behavior, defiance, disrespect of others). They discovered that social competence programs that used cognitive-behavioral techniques and behavioral instruction resulted in positive gains for students (i.e., decreased delinquency, drug use, and conduct problems and improved attendance). Conversely, non-cognitive-behavioral counseling and social work resulted in negative effects. Wilson and colleagues concluded that programs that are based on sound learning principles hold promising results, but their results are cautioned by the size of the ESs found (their small ESs ranged from .04 to .29). Still, it is encouraging that SEL programs appear to decrease internalizing and externalizing behaviors present in students, while also increasing their academic achievement and attendance.

Based on these two meta-analyses, it is reasonable to expect that SEL programs can result in decreases in negative affect (i.e., internalizing behaviors and characteristics), improvement in social functioning (i.e., by decreasing externalizing symptoms), and improved academic performance for students. Studies on SEL programs published since then have continued to document the positive benefits of such programs. Catalano and

colleagues (2003) implemented a comprehensive prevention program (Raising Healthy Children) with first and second graders and found increases in teacher-reported academic performance, commitment to school, and social competency (e.g., ability to understand other people's emotions, cooperate with others), as well as decreases in antisocial behavior (e.g., breaking things, telling lies, fighting), compared to a control group. Riggs, Greenberg, Kusche, and Pentz (2006) implemented the PATHS curriculum, a universal SEL program that includes lessons taught three times a week (30-45 minutes each) and weekly consultation. Although the study focused on neuro-cognition as a mediational role on the effects of SEL on outcomes, the second and third graders who received the PATHS curriculum had lower teacher-reported internalizing and externalizing behaviors at a one-year follow-up, compared to a control group.

The SEL programs documented thus far provide promising hope that other SEL programs, such as *SK,* can result in similar gains. After all, *SK* includes affective education and problem-solving, two elements that have resulted in favorable gains identified by Durlak and Wells (1997), and the entire curriculum is centered on cognitive-behavioral techniques and behavioral instruction, which have resulted in superior gains relative to other intervention methods (Wilson et al., 2001). However, there are elements of the cited studies that make attributing certain outcomes to one factor or another difficult. For instance, Riggs and colleagues (2006) included consultation as part of their program, so it is hard to tease apart the exact impact the PATHS program would have had if consultation had not been a part of the package. Additionally, Catalano and colleagues (2003) included monthly booster sessions, classroom "coaches", and classroom

management techniques. Although they should be applauded for such a multi-element and coordinated effort, it is hard to determine which elements are critical to the benefits they obtained. Such variation in implementation and components of SEL programs leads to the question as to what factors are crucial for SEL success. The next section discusses such factors.

Critical Factors to SEL Success

Schools can't use just any SEL program and expect tremendous gains. The promising effects of such programs are not guaranteed, so schools must be wary of how they go about implementation. Using six different schools and a total sample of 350 first graders, Kam, Greenberg, and Walls (2003) implemented the PATHS curriculum with two major components: (1) twice-weekly lessons on problem-solving and emotion identification and (2) weekly consultation by the project coordinators for the teachers. Evaluating four outcomes (aggression, dysregulated behavior, social competence, attention), the study was largely ineffective. However, when *quality of implementation* and *leadership support* were examined, the authors discovered that those schools with better fidelity and higher principal support resulted in significant changes in aggression, behavior dysregulation, and social competence (attention was not affected). The results on implementation and administrative support are striking. Still, they are not surprising given earlier research in consultation literature and Positive Behavior Support (PBS) that point to the importance of these two factors (Elliott, Witt, Kratochwill, & Callahan-Stoiber, 2002; Sugai & Horner, 2006).

In addition to leadership support and fidelity, *generalization of skills* and the *length of the intervention* are two more factors critical to success. Significant changes in social and emotional functioning of students are modest when programs run less than 9 months (Greenberg et al., 2001). As an example, Kam and colleagues (2003), described earlier, implemented the PATHS curriculum for only 4 months, a length of time generally thought of as too short to result in any preventative effect (Greenberg et al., 2001; Wilson et al., 2001). In comparison, a longer effort by Catalano and colleagues (2003), cited earlier, found that results for the treatment group continued to improve over an 18-month period, whereas results for the control group did not change or actually had negative changes. Collectively, this suggests that SEL programs should target long-term implementation (i.e., 9 months at the least) and there is consensus in the research that a multi-year effort is ideal (Greenberg et al., 2003).

In terms of generalization of skills, the behavioral literature is replete with evidence of the importance of this factor (see Alberto and Troutman, 2006). As students acquire the cognitive knowledge of new skills in one setting, they need targeted and planned generalization for those skills to transfer to new settings. It is simply not enough to use a SEL curriculum and hope that students will apply new skills across settings. The success of social skills programs utilize generalization techniques as a critical component (Gresham, 2002), and some of the successful SEL programs have used reinforcement strategies and generalization techniques to facilitate acquisition of the newly taught skills (see Zins, Weissberg, et al., 2004). Therefore, any SEL program wishing to ensure acquisition of SEL skills should include a generalization feature.

Finally, and perhaps most important, is the ability of a school to implement another program or curriculum when they are already under pressure and time constraints. No doubt it is a lot to ask a school to include SEL as one of their primary objectives, and this is no more evident than in the efforts by Seifer, Gouley, Miller, and Zakriski (2004). Using the PATHS curriculum within one school, 150 children from 3 first-grade and 3 kindergarten classrooms received 3 lessons from the manual each week, but unfortunately, few effects were found. The authors found no differences between the treatment and control groups on social status, emotion identification, or negative feelings toward school, and only found significant effects for depressive symptoms and social competence. It is not surprising few effects were found, despite use of the curriculum over the course of two years. Many of the teachers reported straying from the scripts provided for each lesson, which decreases fidelity, and a few teachers had little motivation to use the program, as the implementation of a new literacy program and academic pressure from administrators during the same time period pressured teachers to focus on academics. Additionally, many of the teachers reported being dissatisfied with the amount of time needed for the program each week (a total of 90 to 135 minutes per week) and did not have access to someone knowledgeable about SEL or the specific program to provide guidance (there is no mention in the article of weekly support or consultation that has been utilized in previous research with PATHS). One of the main themes that emerged from the efforts of Seiger and colleagues is that the school was torn between wanting to provide time for SEL and achieving mandated academic goals. This brings us to the question of how can schools coordinate SEL with academics in a manner

that does not tax the system? With students needing academic instruction, but also needing SEL instruction because they are faced with immense risk factors, what can schools do to coordinate both of these goals? The next section examines the use of a three-tiered model that can help schools meet and sustain SEL goals of their students.

Tiered Support

A Three-tiered Approach to SEL

Although Siefer and colleagues (2004) did not find the PATHS curriculum an effective program, lessons can be taken from their research. It is fair to say that schools want to implement SEL, but that they need long-term efforts that do not jeopardize time for instruction. Efforts must ensure good implementation and generalization techniques, but just as important is support and encouragement from the administration. Coordinating all of these facets is not a simple task, and the amount of time it takes has led to a history of schools in which one program is picked up, tried for a bit, and then turned away as the next "fad" comes along (Merrell, Irvin et al., 2006). Schools need to be able to implement a program that can be sustainable after it is up and running. The innovation of tiered-levels of supports and a systems-level focus can help schools achieve goals and sustain them (Hunter, 2003). A description of the three-tiered model is presented next, followed by its application to SEL programming.

The three-tiered model, commonly called the "triangle", is a multi-level support paradigm that has its roots in prevention science and the field of public health (Hunter, 2003; Merrell & Buchanan, 2006). Within this model, each student is provided an increasing level of support as needed based on their presenting issue of concern. Within

the *primary* level, the base of the triangle, the model aids and assists 80% of the student population. The goal of this universal support is to prevent problems from arising. About fifteen percent of the population is targeted within the *secondary* level and the goal here is to lessen the severity of the problem and prevent the development of more chronic problems. Finally, the *tertiary* level focuses on approximately 5% of the population and deals with providing intensive problem-solving strategies to identify effective interventions and courses of actions for those students. The goal here is remediation and stabilization of the issues students have (Horner, Sugai, Todd, & Lewis-Palmer, 2005).

Use of this model has been successful with behavior in reducing the number of students who move up the triangle into the second and third tiers. Taylor-Greene and colleagues (1997) established school-wide expectations (e.g., be respectful, be responsible, be there- be ready, follow directions, hands and feet to self) in a middle school (grades 6 to 8) and students were reinforced for following them. Within one year, the number of office disciplinary referrals dropped a dramatic 42% from the fall to the spring. The same framework is also being applied to literacy and Response to Intervention models with promising results (Good, Gruba, & Kaminski, 2002; Kame'ennui, Good, & Harn, 2005). Within these models, the number of students "rising up" is decreased, freeing up the system to work more smoothly, more efficiently, and with less burden.

Although no research to date has specified applying the "triangle" to SEL programming, previous research has coordinated layers of support for SEL. For instance, the Conduct Problems Prevention Research Group (CPPRG; 1999) implemented a SEL

curriculum plus tiered support with approximately 198 first grade classrooms (along with 180 control classrooms). All students in the treatment condition received 57 lessons in SEL skills, such as problem-solving and feelings identification over the course of the year, but those students who were identified "early starters" of conduct problems were provided additional support. Ten percent of the treatment group received parenting support classes, small-group social skills interventions, academic tutoring, and weekly visits. Thus, CPPRG provided two tiers of supports to students: a universal curriculum to all students and additional support to a much smaller percentage of students identified as high risk. This combined effort resulted in positive changes at the classroom level within one year. The treatment group compared to the control group had lower peer ratings of aggression and disruptive behavior, whereas independent observations noted improved classroom atmosphere (i.e., decreases in level of disruption and increases of students' engagement in academic activities). Although CPPRG only involved two tiers of support, they give considerable hope that SEL programming can enjoy the same type of success that a multi-tiered model has demonstrated with behavior and academics. A more detailed model of SEL support within a three-tiered model is presented next.

Hypothetical Model of SEL Support.

Primary level. At the *primary* level, all students would receive a preventative curriculum that focuses on the five key SEL skills outlined earlier in this review (see Table 1) (Merrell, Levitt, & Gueldner, in press). The use of such a curriculum would provide all students with a common language and explicit instruction on emotional and social skills that are linked to healthy development (Cohen, 2001; Denham & Weissberg,

2004). In addition, the primary level would include reinforcement of those skills, perhaps in the form of tickets and verbal praise, as used in PBS models (Taylor-Greene et al., 1997). There would be a focus on generalization of the skills between settings so that students would learn to apply the skills outside of the time reserved for the curriculum. This could be accomplished through integrating cooperative activities and learning into academic lessons (Doll et al., 2004), the use of homework sheets as part of the lesson (Merrell et al., 2007), or having students role-play the skills in various settings throughout the school (Zins, Weissberg, et al., 2004). An intriguing example by Elias (2004) discusses a framework that outlines the SEL skills that students can use to examine characters in stories or historical figures they read about. Examples of universal curriculum include the *SK* curricula, the PATHS program (Zins, Weissberg et al., 2004), and *Fast Track* (CPPRG, 1999). All of these curriculums include fairly regular (weekly, twice-weekly) lessons that target SEL skills.

An important feature of this level is using data to monitor students' performance and growth on SEL skills. Although there is not a unified system for SEL as developed as there is in other systems that can provide information uniformly between levels (such as formative assessment in academics or office discipline referrals for behavioral models), the data-collection would likely look similar to the multiple gating framework outlined by Merrell (1999). At the primary level, this would most likely involve screeners or questionnaires that provide staff with information on students who are displaying challenges with respect to social and emotional functioning. Some examples include having teachers rate students on brief measures that capture the global constructs of SEL

skills, such as peer relationships scales or emotion regulation measures, or perhaps collecting information on risk factors students are experiencing (e.g., socio-economic status, recent losses or traumatic events, office referrals). The data system should provide the staff with information on those students that are displaying difficulties with SEL skills, allowing them to determine if students are receiving enough support given their SEL functioning.

Secondary level. Students that continue to display social and emotional concerns, or who score high on screeners used by the school, are then provided additional support within the *secondary* level of support. The purpose here is to prevent problems from becoming more serious, so students who are at-risk for developing more serious and chronic social and emotional concerns are targeted. The interventions delivered within this phase aren't necessarily individualized, but they are more targeted than primary interventions, providing particular attention to one of the SEL skills or risk factors students are having difficulty with. A feature of secondary interventions is that they can reach a large group of students in an efficient manner.

The interventions used could be relatively abundant, from using a "check-in/check-out" format of SEL skills to small-group interventions that target more in-depth SEL skills and topics, such divorce or anger management. Merrell and colleagues (in press) point out that students within this level experience cognitive and problem-solving deficits, so interventions that target these domains could be used (e.g., *First Steps to Success* for externalizing concerns, *Coping with Stress Course* for internalizing concerns). Other examples include the *Friend-2-Friend* program that targets relationship

skills and reducing relational aggression (Leff et al., 2007), providing parenting classes (CPPRG, 1999), using a peer mentoring or "big brother/big sister" program, or programs that target children of divorced parents. Schools do not necessarily have to use a packaged published program and instead, could intensify the existing universal curriculum by providing additional lessons or modifying the content to suit a group of students' needs.

Tertiary level. Students that do not respond to the first two levels of support fall into the *tertiary* level. Within this level, students may have identified disorders and severe emotional and behavioral concerns, so the goal is to remediate the problem and reduce its negative impact. During this level, a more concentrated and individual effort is made to support the child. Ideally, students would receive an in-depth evaluation of their social and emotional needs that may include direct observations, interviews, functional behavior assessment, and rating scales. This information would allow the school to design a "treatment plan" that may include individual counseling, parent-teacher collaboration, and linking parents and students with community supports. Some examples of interventions in the literature include school-based health clinics, (Merrell et al., in press), parent training (Hembree-Kigin & McNeil, 1995), mediation for bullies and victims (Olweus, 2001), weekly visits between the school and parents (CPRRG, 1999), and linking parents with community supports (i.e., "wrap around" services) (Merrell et al., in press). What separates this level from the previous two is its individualized focus and the intensity of the services provided.

SK *as a multi-tiered approach to SEL.* It is possible to use SEL programming within a three-tiered model, but the success of this model is dependant on both the features reviewed that are important (e.g., fidelity, length of implementation, administrative support, generalization) *and* the effectiveness of the universal curriculum. The foundation of any tiered level of support is the most critical, as the universal level is only as good as the number of cases it prevents from rising to higher tiers. Any program that cannot prevent cases from moving higher up into the system does not provide the foundational structure for the system to operate effectively.

Preliminary research with *SK* has indicated gains in knowledge of social-emotional concepts and coping skills and reductions in negative affect with a 6th grade sample (Gueldner, 2006). Castro-Olivio (2006) demonstrated similar gains in knowledge with an adapted version of the curriculum with high-school Latino students, as well as good treatment fidelity and high social validity. There is great promise in *SK* as a universal level of support within a three-tiered model, but documentation of its impact on social and emotional functioning *over time* is needed next. The present study addressed this need as it evaluated *SK* on a small scale to determine the extent of its preventative qualities, based on four research questions.

Research Questions and Hypotheses

1. Did the use of a SEL curriculum lead to an increase in the knowledge of SEL skills (as measured by the SK knowledge test)?

It was anticipated that students within the treatment (i.e., those exposed to the *Strong Kids* curriculum) would score higher on a post-test measure of knowledge relative to a control group that was not exposed to the curriculum.

2. Did the use of a SEL curriculum lead to an increase in the use of SEL skills (as measured by number of coping skills and social-emotional assets)?

It was expected that students within the treatment group would report using more coping skills when presented with a coping scale and score higher on a social-emotional assets measure.

3. Did those gains in knowledge and use of SEL skills maintain over time, at a spring follow-up?

A desirable outcome would be for the differences between the treatment and control group at post-test maintained at a follow-up in the spring time. However given the dearth of research on long-term impact of brief SEL programs such as *SK*, the research base was not sufficient to make a directional hypothesis.

4. Did the use of the SEL curriculum in the fall lead to better social functioning in the spring, relative to a group who did not receive the SEL curriculum?

Again, the limited research on long-term use of brief SEL programs eschews making a directional hypothesis, but a desirable outcome was that students in the treatment group compared to the control group would score better on a measure of social functioning.

CHAPTER III

METHODS

Participants

Recruitment

Principals from elementary schools in the Pacific Northwest (i.e., Eugene and Springfield, OR) were initially contacted via email about the study. A total of 17 principals were informed of the study and 3 expressed interest in participating. After discussing with the principals more details about the study, each principal identified teachers from grades 3 to 5 to inquire about their participation in the study. The researcher discussed the study with teachers either in person or via email. Eight teachers from 3 schools expressed interest, however, two teachers from one school declined participation citing difficulty with managing a new reading curriculum and the study simultaneously, and two other teachers from another school were unable to participate because they could not be randomly assigned to the treatment conditions. Four teachers from one school agreed to participate.

Final Sample

The final sample consisted of 106 students from 4 classrooms in grades 3 and 4 (two 3rd grade classrooms, one 4th grade classroom, and one classroom consisting of 3rd and 4th graders). Fifty-four percent of the sample was girls and 46% was boys. The sample was largely Caucasian (84%) with a mean age of 8 years, 5 months for 3rd graders

and 9 years, 5 months for 4th graders. Information on office discipline referrals was collected as another means of describing the participants. Teachers 1, 2 and 3 had an average of approximately 1.5 referrals per month between September and February, but Teacher 4's classroom had an average of 3.2 (10 referrals alone were given in February). Appendix M provides more detailed information on the referrals from each classroom. Table 2 summarizes information on the ethnicity, gender, and mean age of the sample.

Procedure

Design

The current study employed a quasi-experimental design with 1 qualitative independent variable (i.e., treatment or wait-list condition) and 4 quantitative dependent variables. Although it would have been desirable to assign students at the individual level to conditions, this was not possible because students were in pre-existing groups (i.e., their respective classrooms). Furthermore, students had not been randomly assigned to classrooms when the school was developing its rosters. Instead, students were placed into their classrooms by their previous year's teacher (e.g., second grade teachers chose classrooms for next year's third graders). Students were assigned by matching the students' strengths and needs (based on previous teacher's judgment) with the teachers' strengths, as well as academic ability, social maturity, and wanting to keep certain students in separate classrooms (e.g., students who engage in problematic behaviors with each other were "broken up" between classrooms). One exception is the classroom that consisted of both 3rd and 4th graders. The 4th graders had been placed in that classroom prior to their 3rd grade year, with the expectation that they would stay in that classroom

for their 4th grade, so approximately two-thirds of that classroom was chosen based on parent request from a year ago. The remaining third of the students in that particular classroom were assigned by their teachers from last year or were randomly assigned based on the numbers of students in other classrooms.

Because students could not be randomly assigned at the individual level, each classroom was randomly assigned by both grade level and teacher to either the treatment or wait-list condition, resulting in one 3rd and one 4th grade classroom in the treatment group and one 3rd grade and one 3rd/4th grade classroom in the wait-list group. The number of students in each condition was evenly split, with 54 students assigned to the treatment condition and 52 assigned to the wait-list condition (see Table 2).

Data Collection

Passive consent forms, approved by the University of Oregon Human Subjects IRB, were sent home with students from four classrooms (see Appendices A and B). Parents/guardians were given two weeks to return the form if they *did not want* their child participating in the study. Only 2 of the 110 parents who received forms declined participation, resulting in an acceptance rate of 98%. Data were collected during the regular school day. Two to three researchers entered the classroom and obtained assent from each student (see Appendices C and D) and only 2 students declined participation. The researchers then administered 3 measures (The Coping Scale, The Social-Emotional Assets and Resilience Scale for Children, The *Strong Kids* Knowledge Test; see Appendices E through H) in both whole-class and small-group format. Students were given a small reinforcer (e.g., fruit snack, eraser, pencil) for completing the measures.

Each teacher completed the School Social Behavior Scale on each student over a one-week time period for monetary compensation. The same data collection procedure was used for all of the data collection periods (i.e., pre-testing, post-testing, follow-up testing).

Table 2

Demographical Information of Participants (N = 106)

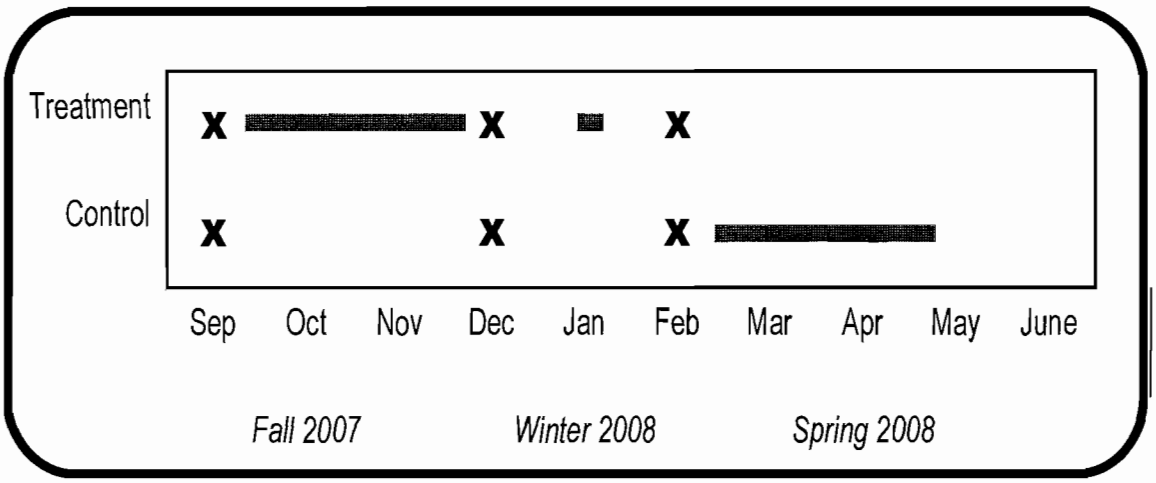
	Treatment		Wait-List		*Total*	
Sample	N	% [a]	N	% [a]	N	% [b]
Race/Ethnicity						
African American	2	4	0	-	2	2
Asian American	0	-	2	4	2	2
Caucasian	46	85	38	73	84	79
Native American	2	4	1	2	3	3
Hispanic	2	4	9	17	11	10
Other	2	4	2	4	4	4
Grade						
Third	27	50	39	75	66	
Fourth	27	50	13	25	40	
Gender						
Boys	24	44	25	48	49	46
Girls	30	55	27	52	57	54
Total	*54*		*52*		*106*	

Note: [a] Represents percentage within condition; [b] represents percentage of total sample.

Implementation

Teachers in the treatment group implemented the *Strong Kids* curriculum once a week for 12 weeks beginning in September of 2007. Teacher 2 implemented all of the lessons herself, but Teacher 1 had a student teacher that implemented half of the lessons. The curriculum ended in December and a booster session was conducted in February of 2008. At that time, the wait-list group implemented *SK* in their respective classrooms following the same format of one lesson per week for 12 weeks (they did not implement a booster session). Unfortunately, Teacher 1 did not implement the booster session because of snow days and a lack of planning. This issue has bearing on the results and is discussed in the Discussion chapter. A visual representation of the study is displayed in Figure 1.

The teachers were trained prior to beginning implementation (treatment group in September, wait-list group in February). The training consisted of a one-hour session in which the researcher went through a lesson from the curriculum (Lesson 6) and highlighted the general format of the lessons. Each teacher was provided with all of the handouts and overheads for each lesson, as well a copy of the curriculum. Part of the training included discussing the pre-correction and providing of social reinforcement of the skills students would learn. Teachers were given examples of ways to pre-correct students and specific examples of behaviors to praise. Appendix I outlines the agenda used for the trainings.



Note: "X" denotes assessment period; gray bar denotes use of curriculum.

Figure 1. Timeline of Study

Fidelity Checks

Fidelity of implementation was assessed through a series of observations of the lessons. Fidelity refers to how well the treatment was implemented as intended (Elliot et al., 2002). This was measured by recording the number of components of the curriculum that were presented to students. An observation checklist was created that described key behaviors that each lesson calls for, such as use of the overheads and handouts, and certain ideas or key points of the lesson. (Appendix J contains the fidelity checklists used for the study). A lesson was observed and each behavior was marked as either implemented or not and the teacher was given a "percentage of components implemented" score. Figure 2 displays the percentage of components implemented for each teacher in the treatment condition. Overall, the treatment condition had a mean rate of 85% of components implemented, with Teacher 1 implementing 79% on average and

Teacher 2 with 91%. Nearly one-third of the total lessons were observed for fidelity (7 out of 24 lessons between both teachers).

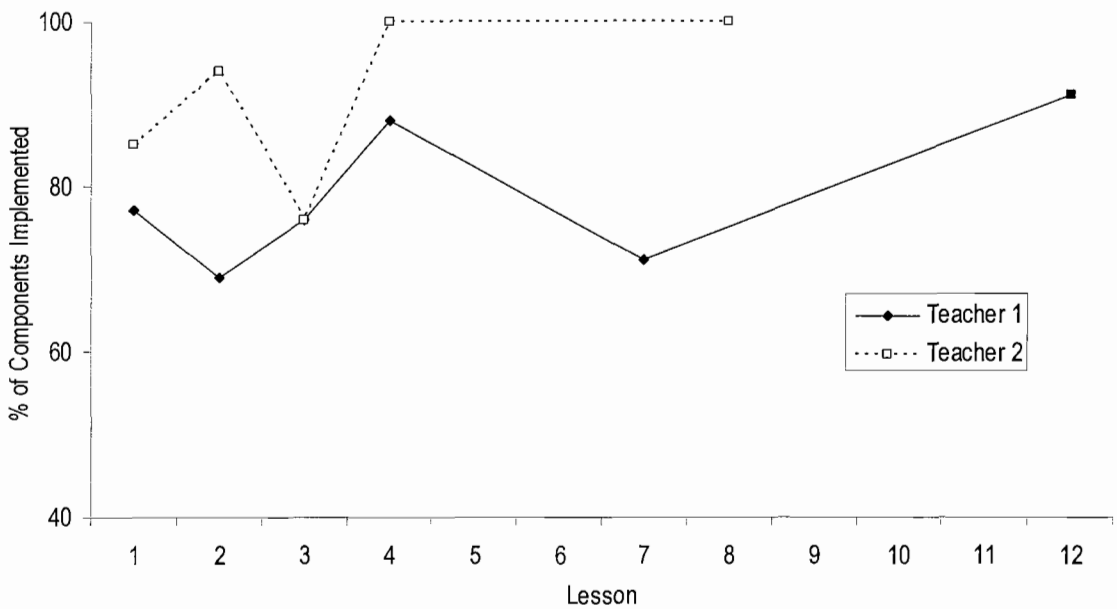


Figure 2. Percentage of Components Implemented Between Teachers in Treatment Condition

Inter-rater reliability checks were obtained for 3 lessons for each teacher in the treatment condition. Two research assistants conducted simultaneous observations using the fidelity observation sheets described above. Following the interval agreement method, inter-rater reliability was calculated by taking the number of components each observer agreed upon and the number of components they disagreed upon (e.g., one observer recorded a component, but the other observer did not). Using the formula: $[A / (A + D) * 100]$, where A = the number of agreements and D = the number of disagreements (Kennedy, 2005), the total agreement for Teacher 1 was 94% and 100% for Teacher 2.

Generalization of Skills

An important feature of the current study was promoting generalization of the skills taught in *SK*. This generalization was achieved through (1) pre-correction of the skills (e.g., reminding students to use *SK* skills prior to engaging in various school activities) and (2) providing verbal and social praise for use of the skills. For example, the teachers reported that they reminded students of the importance of discussing their feelings prior to a student disclosing about a death in his family and to use their newly learned anger management skills prior to recess in which some students have a history of being competitive. The teachers also reported praising students when they identified thinking errors throughout the school day.

Teachers were asked to pre-correct and reinforce students as often as they could and to record each time they (1) pre-corrected and (2) socially praised a student or group of students for using a SEL skill from the *SK* curriculum. The teachers recorded their data using a tally system on a data sheet provided by the researcher. No minimum number of pre-corrections or praises was set because the generalization feature was designed to be low-key, flexible, and easy to implement. Additionally, pre-correction was promoted though the display of two posters (one from Lesson 6: Thinking Clearly and one from Lesson 4: Managing Anger) that illustrated some of the skills learned by the students (see Figures 3 and 4 for depictions of the posters used).

As seen in Figures 5 and 6, Teacher 2 provided a steady rate of pre-corrections and social praise throughout the lessons, whereas Teacher 1 had decreasing rates of generalization over time. Teacher 1 provided an average of 2.08 pre-corrections per

week, compared to 2.5 for Teacher 2. In terms of verbal or social praise, Teacher 1 provided 3.25 praise statements per week, compared to 4.58 for Teacher 2. Given that there was no expectations placed on how often teachers should pre-correct and praise students, it was encouraging that the teachers were able to provide steady rates of pre-correction and praise.

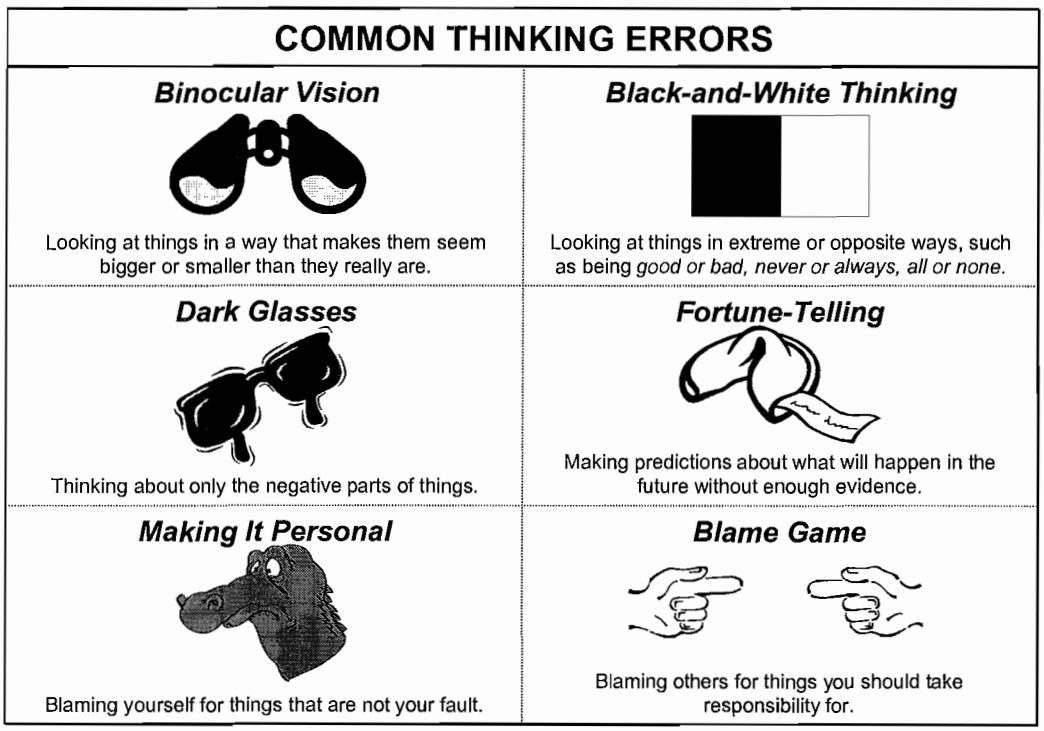


Figure 3. "Common Thinking Errors" Poster

Compensation for Teachers

It was a lot to ask a teacher to give up instructional time for a study, so each teacher was provided compensation as a thank you. At the conclusion of the study, each teacher received a complimentary gift bag with a gift card to Target and various snack

items. Each teacher also received monetary compensation at each data collection point for completing behavior measures on their students.

Compensation for Students

Students received small reinforcers each time they completed measures during the data collection periods. In addition, each classroom received a free Subway lunch party for their participation.

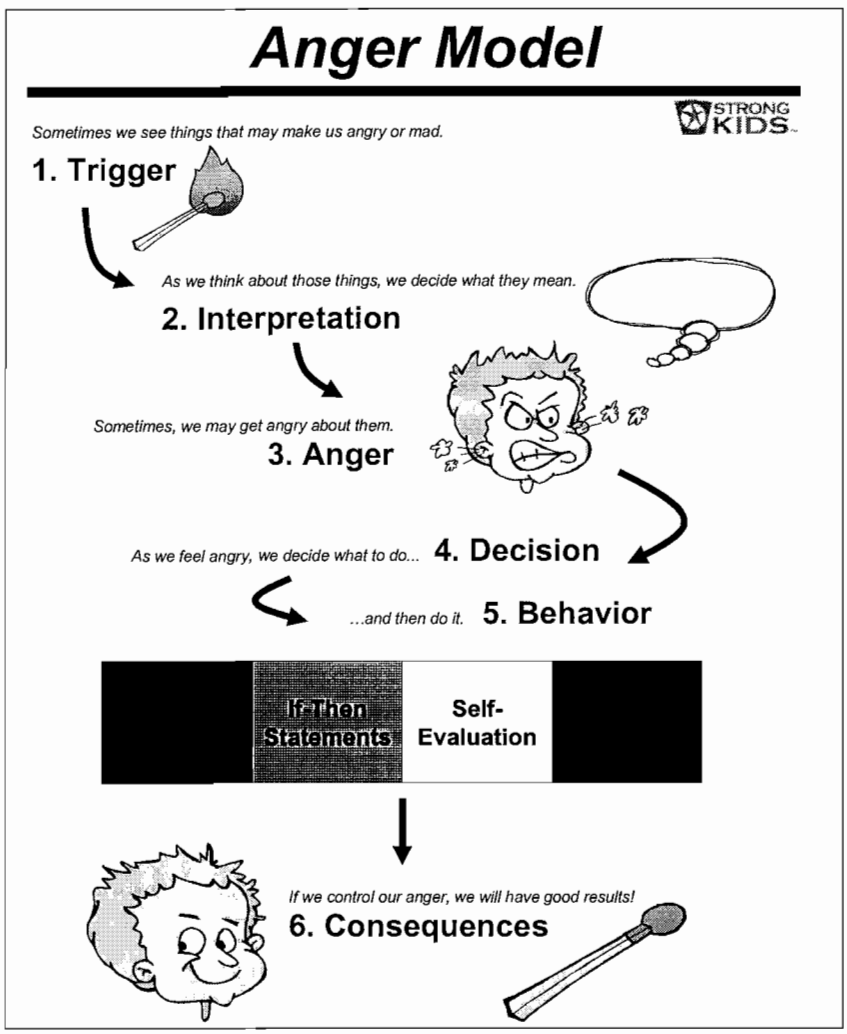


Figure 4. "Anger Model" Poster

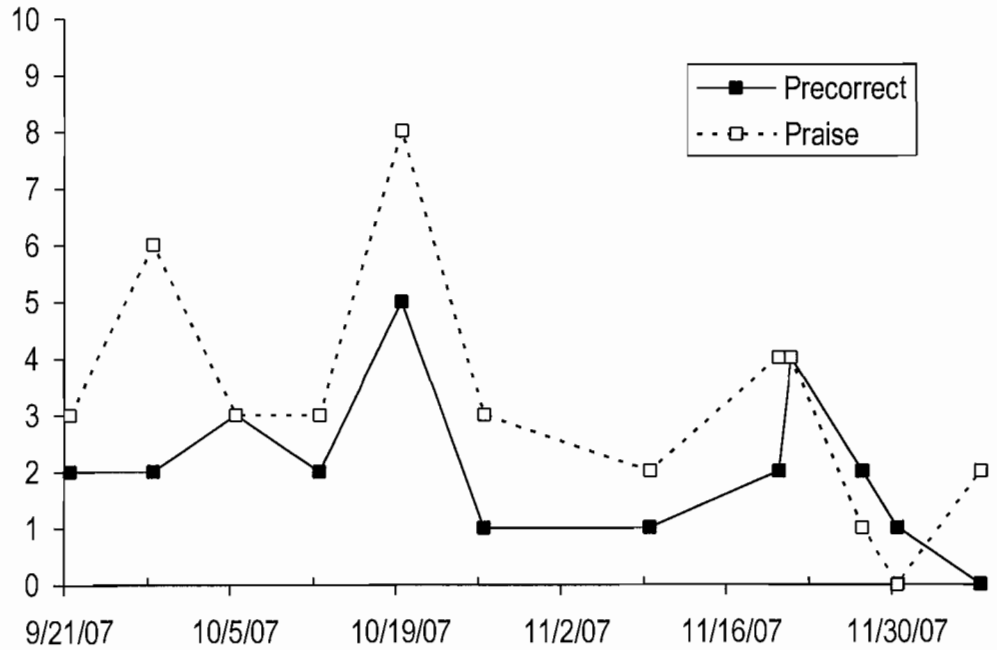


Figure 5. Occurrences of Precorrection and Social Praise Provided by Teacher 1

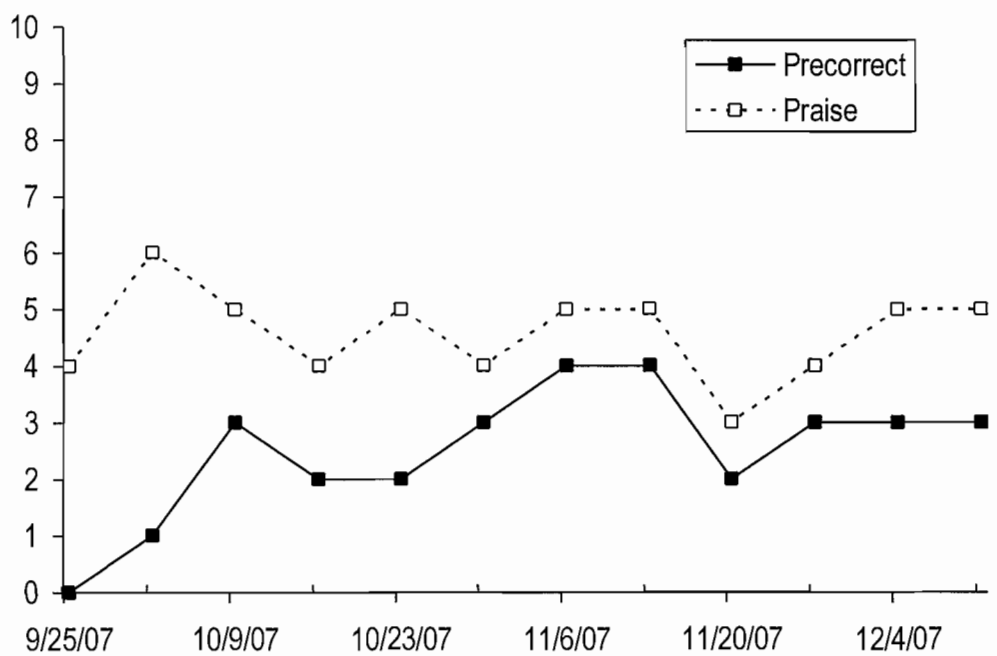


Figure 6. Occurrences of Precorrection and Social Praise Provided by Teacher 2

Overview of Independent and Dependent Variables

Independent Variable: Strong Kids *Curriculum*

The IV was whether or not the student is exposed to the *SK* curriculum (Merrell et al., 2007). *SK* is a brief, universal curriculum that teaches social, emotional, and mental health skills to children in grades 3 to 8. *SK* is based on the notion of *psychological wellness* (i.e., the concept of developing competencies that promote healthy functioning alongside reducing pathology and risk factors), and specifically targets the five pathways to wellness outlined by Cowen (1994). The curriculum enlists cognitive change, affective education, and behavioral skills in order to produce social and emotional competency gains and resiliency. Topics covered include emotion-identification, anger management, thinking errors, problem-solving, setting goals, and managing stress.

There are 12 individual lessons and 1 booster session that last approximately 40 to 50 minutes each. Each lesson is scripted, although teachers and staff who employ the curriculum are free to provide their own examples as part of their instruction, and are encouraged to do so when it is appropriate. However, each lesson rests on sound instructionally strategies, as they consist of teacher-led instruction of the skill and its components, modeling of the skill, opportunities to practice the skill with feedback, and generalization of the skill through targeted homework activities. In addition, each lesson includes a review of the previous lesson, group discussion and activities, and review of the homework for that lesson.

Dependent Variables

The study included 3 dependent variables. They were SEL knowledge (measured by the *SK* Knowledge Test), perceived use of SEL skills (measured by both the Coping Scale and the SEARS-C), and social functioning (measured by the SSBS-2). Each measure is discussed next.

SEL Knowledge: Strong Kids *Knowledge Test*

The *SK Knowledge Test* is a 20-item measure with 5 True-False and 15 multiple choice items that each student completes. It provides a measure of how much information and concepts from the *SK* curriculum students learned and retained. Items ask students' knowledge of definitions (i.e., "What is an emotion?") and their application of concepts (e.g., "Devin's gym teacher tells him to try out for the basketball team. Devin thinks he is too short and won't make it, so he decides not to try out. What thinking error is described here?").

Previous research with this measure has demonstrated adequate internal consistency for a brief measure used for research and administrative purposes (e.g., .60s to .70s range), and strong sensitivity to change following treatment. In addition, some studies have shown interesting patterns of convergent and discriminant validity between the knowledge test and other social-emotional outcome measures. Each student completed the questionnaire in a small-group in which a researcher read the items aloud. A total score of number of items correct was tallied to represent each participant's level of SEL knowledge.

Reliability with current sample. Reliability was assessed using the Kuder-Richardson-20 (KR-20) formula because the items were scored dichotomously (0 = incorrect, 1 = correct). (KR-20 provides an estimate of reliability similar to Cronbach's alpha, but is used when items are dichotomous, Salvia & Ysseldyke, 2004). Reliability coefficients below .60 are considered unacceptable based on Salvia and Ysseldyke's guidelines for group data. The KR-20 estimates for the *SK* Knowledge Test were as follows: .48 at pre-testing, .66 at post-testing, and .70 at follow-up for the treatment group, and .63, .58, and .65 for the control group. Test-retest reliability was calculated with Pearson's *r*, and Cohen's (1988) guidelines of .10, .30, and .50 as weak, moderate,

Table 3

Internal Consistency Alpha Coefficients for Dependent Measures

Questionnaire	Pre-testing	Post-testing	Follow-up
Knowledge Test			
Treatment	.48	.66	.70
Wait-list	.63	.58	.65
Coping Scale			
Treatment	.82	.82	.91
Wait-list	.88	.86	.89
SEARS-C			
Treatment	.92	.93	.95
Wait-list	.95	.95	.95
SSBS-2			
Treatment	.97	.98	.98
Wait-list	.96	.96	.97

and strong was used to determine the strength of association between scores. Using only the wait-list condition to calculate test-retest reliability, test-retest at both pre-testing (3-months; $r = .76$) and follow-up (5-months; $r = .59$) was strong. Table 3 displays the internal consistency coefficients for all of the dependent measures, and Table 4 displays the coefficients for the test-retest reliability between the data collection points.

Perceived Use of SEL Skills: Causey and Dubow's Coping Scale

A shortened version of the *Coping Scale* (Causey & Dubow, 1992) was used as one measure of a student's perception of their use of SEL skills. The Coping Scale is a 34-item self-report questionnaire that measures a students' perceived use of several coping skills in response to two vignettes. Students are presented with a social scenario (i.e., "When I get a bad grade in school, one worse than I normally get, I usually…") and then respond to items using a 5-point Likert scale (1 = never to 5 always) that ask how often they perform a certain skill, such as "tell a friend or family member what happened" and "forget the whole thing". Items represent one of five coping skills that

Table 4

Test-retest Reliability Alpha Coefficients for Dependent Measures

Questionnaire	3 months	5 months
Knowledge Test	.76	.59
Coping Scale	.57	.52
SEARS-C	.63	.56
SSBS-2	.87	.87

were derived from factor analysis (e.g., seeking social support, problem-solving, distancing, internalizing, externalizing). Students then read to a second scenario (i.e., "When I have an argument or fight with a friend, I usually…) and respond to the same items.

The Coping Scale has demonstrated good reliability among fourth to sixth graders across all five coping skills (Cronbach's alpha scores ranged from .66 to .84), but some of the individual items inter-correlation fell below .60. Test-retest reliability is adequate, with two-week reliability scores ranging from .58 to .73 with previous research. In terms of validity, the Internalizing subscale was positively correlated to the Total Anxiety scale score on the Revised Child Manifest Anxiety Scale ($r = .40$), and the Problem Solving subscale was related to "perceived control" (as measured by the Self-Perception Profile for Children; $r = .64$). In general, the Coping Scale has modest to moderate relation to corresponding measures that target similar constructs (Causey & Dubow, 1992).

In order to reduce the length of the measure and improve the reliability, the two scales with the highest reliability were used to measure a student's self-reported ability to use social-emotional skills (the "Seeking Social Support" and "Self-reliance/Problem-solving" scales each had $\alpha = .84$). These two scales combined were hypothesized to measure the SEL skill *responsible decision making* (see Table 1). The end result was a 16-item questionnaire in which students were presented with the social situation "When I get in a fight or argument with a friend, I…" and then rated how often they perform certain behaviors that were listed, such as "…try to think of different ways to solve it" and "…get help from a friend" on a 5-point scale (i.e., 1 = "never", 2 = "almost never", 3

= "sometimes", 4 = "almost always", 5 = "always"). A total score was calculated, which represented their self-perception of their use of SEL skills.

Reliability with current sample. Internal consistency was measured using Cronbach's alpha. The alpha coefficient was above the .60 criterion suggested by Salvia and Ysseldyke (2004) at pre-testing (α = .82 for treatment, .88 for wait-list), post-testing (α = .82 for treatment, .86 for wait-list), and at follow-up (α = .91 for treatment, .99 for wait-list) (see Table 3). Test-retest reliability was r = .57 at 3-months (post-testing) and .52 at 5-months (follow-up) (see Table 4).

Perceived Use of SEL Skills: Social-Emotional Assets and Resilience Scale for Children (SEARS-C)

The SEARS-C (Merrell, 2007) is a recently developed experimental 52-item self-report questionnaire, designed by the senior author of the Strong Kids programs, and currently in the process of a national standardization field trial. The SEARS-C measures children's perception of their social-emotional competencies and their social assets. It was specifically developed to measure gains in the skills and competencies that the *SK* curriculum is centered upon. Although the questionnaire does not have any psychometric information to date, the item content was developed through a four-step process that began with approximately 120 items and included a cluster analysis. Readability analyses of the SEARS-C items and directions has indicated a Flesch-Kincaid Grade Level readability score of 2.5, and a Flesch Reading Ease score of 92%.

Students were asked to rate how true a behavior was of them on a 4-point scale (i.e., 0 = "never true", 1 = "sometimes true", 2 = "often true", 3 = "always true"). A total score was tallied, which represented their self-perception of their SEL skills. Although

there is likely overlap between the constructs measured by the SEARS-C and the Coping Scale, the SEARS-C is believed to measure all 5 of the SEL skills listed in Table 1 (and was designed in part for that purpose) whereas the Coping Scale measures more specific SEL skills in response to a specific situation (i.e., responsible decision making; see Table 1).

Reliability with current sample. Internal consistency was high with the SEARS. Alpha coefficients were more than adequate as they were consistently above .90 for all three data collection points for both treatment conditions (see Table 3). As with the other two measures discussed so far, test-retest reliability was within the strong range (r = .57 at 3-months and r = .62 at 5-months) (see Table 4).

Validation of SEARS-C. Because the SEARS-C is a new measure, it was important to check how it correlates with the other DVs. If the SEARS-C is measuring similar constructs of SEL skills as the other measures are, then a small to moderate correlation can be expected. As seen in Table 5, the SEARS-C correlates significantly with the Coping Scale (r = .47), the Knowledge Test (r = .21) and the SSBS-2 (r = .24). Using Cohen's (1988) conventions for estimating effect sizes, the SEARS-C has small to moderate correlations with the other measures (r = .10, .30, and .50 as small, moderate, and large ESs, respectively).

Social Functioning: School Social Behavior Scales- Second Edition

In order to measure social functioning, a scale from the School Social Behavior Scales (SSBS-2; Merrell, 2002) was completed by the teacher. The SSBS-2 is a 64-item questionnaire that measures *social competence* and *antisocial behavior*. Respondents read

items (e.g., behaves appropriately at school; remains calm when problems arise) and rate the behavior's occurrence from 1 (never) to 5 (frequently). Reliability is impressive for the SSBS-2. Internal consistency reliability coefficients were .91 and above, and test-retest reliability coefficients at 1- and 3-weeks were $r \geq .86$ and $\geq .60$, respectively. Validity has been established with the SSBS-2 by its moderate to strong correlations with other measures of social functioning, such as the Conners Teacher Rating Scale and Social Skills Rating System, and by its ability to distinguish among various groups of students who would be expected to differ in terms of their teacher-rated social competence (e.g., students with and without learning disabilities).

The 14-item Peer Relations subscale was used for this investigation. This scale has an internal consistency of $\alpha = .96$ and best represents social functioning compared to the other dimensions of the SSBS-2, given that it accounted for the largest percentage of variance and the highest Eigenvalue in the initial and confirmatory factor analyses that were conducted on the scales. The teacher completed the 14-items of the Peer Relations subscale of the SSBS-2 on each student at each data collection period.

Table 5

Correlations among Dependent Measures as Represented by Pearson's r

Questionnaire	1	2	3
1 Knowledge Test	-		
2 Coping Scale	.11	-	
3 SEARS-C	.21*	.47*	-
4 SSBS-2	.30*	.03	.24*

Reliability with current sample. As with the other DVs, internal consistency was measured using Cronbach's alpha. Reliability for the shortened SSBS-2 was impressive for both treatment conditions, as it was $\alpha \geq .90$ at all of the assessment periods. Test-retest reliability was also impressive at 3-months (post-testing, $r = .90$) and 5-months ($r = .90$) (see Table 4).

Variables for Each Participant

To summarize, each student was assigned to one of two levels for the independent variable (i.e., treatment group or wait-list group). Each participant also had 4 dependent variables scores: (1) total score on the Knowledge Test, (2) total score on the Coping Scale, (3) total score on the SEARS-C, and (4) total score on the SSBS-2.

CHAPTER IV

RESULTS

Data Cleaning

The data set was cleaned using guidelines outlined by Tabachnick and Fidell (2001). Although no participants were deleted, analysis of the pre-testing data set resulted in 3 scores being deleted because they were considered outliers (i.e., their respective z-scores were more than 3 standard deviations from the group mean). No cases or variables were deleted due to missing data, as less than 5% of any participant's responses were missing. The final data included 106 participants. The variance of each dependent variable was analyzed and each variable passed tests of normality, had non-significant skewness and kurtosis, and had homogeneity of variance.

Pre-test Differences

An important step in conducting research with control-groups is examining the extent to which the groups in each condition are similar prior to conducting analyses. Pre-existing differences should be identified so that any potentially confounding variables can be identified. In addition, having a good grasp of the patterns in the data can avoid misattributing the effects of the IV to an extraneous factor or pre-existing difference (Tabachnick, & Fidell, 2001). The next section explores any differences on the dependent measures based on demographic variables (e.g., age, gender).

Age, Race, Grade, and Gender as Differences on Pre-test Measure

To examine age-related effects, age was constructed as a qualitative variable with 3 levels (8, 9, and 10 years old). A one-way MANOVA was conducted with age as the IV and the scores on the dependent measures as the DVs. The MANOVA was non-significant [Wilk's $\Lambda = .91$, $F(8, 182) = 1.07$, $p = .39$; power = .49], indicating that age was not related to scores on the dependent measures. An additional one-way MANOVA was conducted with race as the IV (race was a qualitative variable with 6 groups: Caucasian, African-American, Asian-American, Native American, Hispanic/Latino and Other) and the 4 aforementioned total scores on the dependent measures as the DVs. The results were non-significant [Wilk's $\Lambda = .81$, $F(20, 292.83) = 0.96$, $p = .51$; power = .61], indicating that race was not related to any of the scores on the dependent measures.

A final two-way MANOVA was conducted with gender and grade entered as IVs and total scores on the dependent measures as the DVs. The omnibus test was significant for grade [Wilk's $\Lambda = .89$, $F(4, 90) = 2.80$, $p = .03$; power = .75], but non-significant for gender [Wilk's $\Lambda = .93$, $F(4, 90)$, $p = .16$; power = .50] and the interaction [Wilk's $\Lambda = .99$, $F(4, 90)$, $p = .88$; power = .11]. Post-hoc testing revealed that 4th graders scored higher on the SEARS and the SSBS-2 compared to 3rd graders. Fourth graders scored higher on the SEARS ($d = 0.52$) and SSBS-2 ($d = 0.50$) by a medium magnitude (based on Cohen's (1988) conventions of 0.20, 0.50, and 0.80 as small, medium, and large ESs). The means for each grade are presented in Table 6. Because one classroom includes both grades 3 and 4, additional analyses were conducted to determine if the grade effect was partly due to a classroom effect.

Table 6

Pre-testing Means and Standard Deviations of the Dependent Measures Between Grade Level

	Grade Level			
	Third		Fourth	
Variable	M	SD	M	SD
SK Knowledge Test	9.69	3.21	10.70	2.75
Coping Scale	37.02	11.24	41.74	8.13
SEARS-C	110.41	21.00	118.58	17.10
SSBS-2	30.45	12.67	36.28	13.20

Classroom Differences

Pre-test differences were further analyzed by examining the differences on the dependent measures between classrooms. The goal of these analyses was to determine if any classroom effects could account for, or contribute to, the grade-related effects on the dependent measures. A one-way MANOVA was conducted with classroom entered as an IV (4 levels) and total scores on the dependent measures entered as DVs. The results were significant [Wilk's $\Lambda = .65$, $F\,(12, 238.41) = 3.59$, $p = .00$] and post-hoc testing revealed a classroom effect for the total scores on the SEARS and the SSBS-2 (as found earlier with grade). Appendix K lists the means of each classroom between each testing period.

To determine if the differences between classrooms on the SEARS and SSBS-2 were due to classroom effects, grade effects, or both, a two-way MANOVA could be conducted to analyze the interaction effects. Unfortunately, there were not enough

degrees of freedom for such an analysis. Consequently, a 'dummy' variable was created to represent the interaction (i.e., a student's classroom and grade level). This inclusion resulted in a qualitative variable with 5 levels (teacher 1/grade 4; teacher 2/grade 3; teacher 3/grade 3; teacher 4/grade 3; teacher 4/grade 4). A one-way MANOVA was conducted with the newly constructed "grade and classroom" variable as the IV and total scores on the SEARS and SSBS-2 as the DVs. The MANOVA was significant [Wilk's Λ = .61, F (16, 272.54), p = .00; power = .98], and the respective means of each grade and classroom are displayed in Table 7. A visual display of the grade-level means and the means for each classroom/grade level are displayed in Figures 7 and 8.

Table 7

Means and Standard Deviations Between Grade Level and Classroom on the SEARS-C and the SSBS-2

	SEARS-C		SSBS-2	
Grade	M	SD	M	SD
Third				
Teacher 2	109.17	21.20	36.37	11.81
Teacher 3	117.29	19.58	27.61	12.22
Teacher 4	96.63	17.85	21.45	8.51
Fourth				
Teacher 1	123.18	15.68	38.74	13.09
Teacher 4	108.26	16.15	31.15	12.39

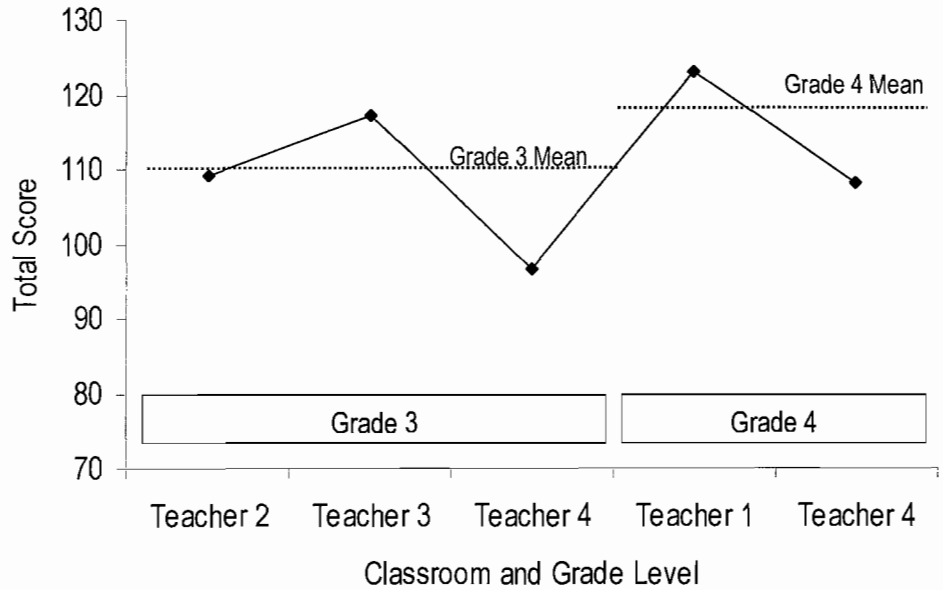


Figure 7. Means of Constructed "Classroom and Grade Level" Variable on the SEARS-C

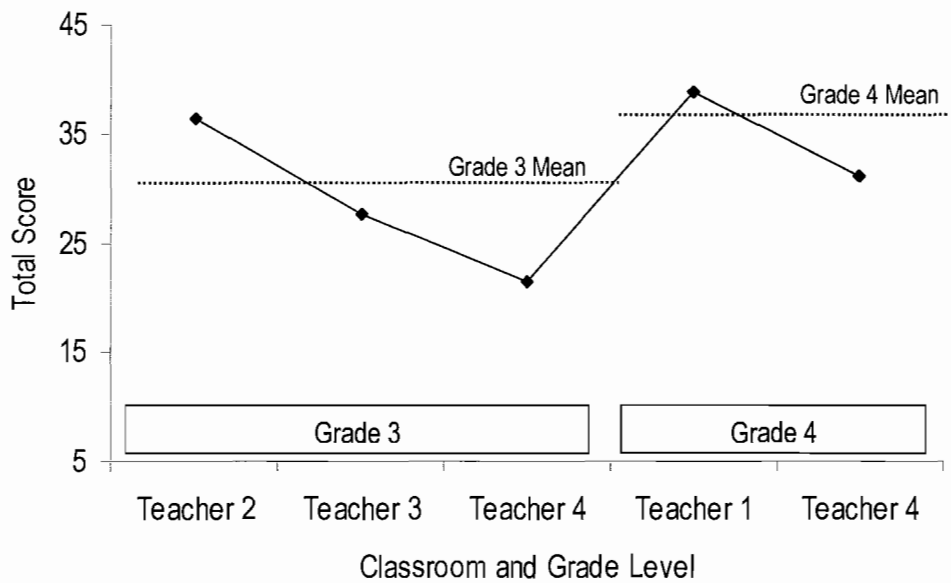


Figure 8. Means of Constructed "Classroom and Grade Level" Variable on the SSBS-2

Post-hoc testing revealed that participants in teacher 4's third grade classroom (teacher 4/grade 3) scored statistically significantly lower on the SEARS compared to participants in teacher 3's third grade classroom (teacher 3/grade 3; d = -1.10) and teacher 1's fourth grade classroom (teacher 1/grade 4; d = -1.58). In terms of the SSBS-2, the third graders in teacher 4's room scored statistically significantly lower than teacher 2's third grade classroom (teacher 2/grade 3; d = -1.45) and teacher 1's fourth grade classroom (d = -1.57). Additionally, participants in teacher 3's third grade classroom (teacher 3/grade 3) scored lower on the SSBS-2 compared to teacher 2's third grade classroom (d = -0.73) and teacher 1's fourth grade classroom (d = -0.89).

Because the differences among classrooms were not consistent among grade levels, it is reasonable to conclude that the study has a "nesting" effect (Killip, Mahfoud, & Pearce, 2004). Nesting and its impact on the present study are discussed next.

Intraclass Correlation

"Nesting" is an issue with research methodology if participants are already in pre-existing groups (such as a classroom) and when participants are assigned to conditions at a group level, yet statistics are conducted at the individual level. Although it is appropriate to conduct statistical analyses at the level of random assignment, this is not often easy to do in applied research in schools. Students are already in pre-existing groups and it is not feasible to move students into newly structured classrooms, nor is it particularly feasible to obtain a large enough sample size at the group level. A clustering issue is problematic because participants within an existing cluster tend to be more similar to each than to a group chosen purely at random (i.e., a group not within a pre-

existing cluster). As a result, the variation in their responses is reduced, leading to a reduction in the within-group variance. This homogeneity among responses of participants within clusters can magnify the differences between participants on the outcome measures. Consequently, clustering can over-estimate the effects of the independent variable (Killip et al., 2004).

In order to account for this clustering effect, an *intracluster correlation coefficient* (ICC) can be calculated and used to adjust the sample size (the result is the calculation of the *effective sample size* (ESS)). The ICC is discussed first, followed by a section on the ESS.

Intracluster Correlation Coefficient

The intracluster correlation coefficient represents how related clustered data is by comparing the amount of within-level variance to the between-level variance. It is represented with:

$$(ICC)\rho = \frac{s_b^2}{(s_b^2 + s_w^2)}$$

If participants respond similarly within clusters, then ρ (rho) approaches 1 and there is more and more of a clustering effect (mathematically within the equation above s_w^2 would equal 0). This results in no variance within a cluster; therefore, analyses should be conducted at the group level. Conversely, if ρ approaches 0, this indicates that there is no correlation between participants' responses and the cluster or group they are in and therefore, there is no nesting or cluster effect. Values of ρ typically range between 0.05 and 0.15, but this varies with the outcome being measured and the group itself.

Much of social science research indicates values between 0.01 and 0.05, but ICC values can range between 0.15 and 0.25 in studies in which no covariate is accounted for (Hedges and Hedberg, 2007).

Once the ICC is known, a researcher can use it to determine the *effective sample size* (ESS). The ESS indicates the sample size one has from a statistical perspective. Although the current sample is 106 participants (an average of 25.5 participants in each cluster), the magnitude of the ICC may lower the sample size from a statistical efficiency standpoint, resulting in a loss in power. Despite having 106 participants, the analyses conducted in the presented study functioned under the ESS, as this accounts correctly for the intraclustering effect (Killip et al, 2004).

An ESS was calculated for each dependent variable using the ICC (represented by the Greek symbol *rho* (ρ)). The ESS was calculated using the following equations from Killip and colleagues (2004):

$$ESS = \frac{mk}{DE}$$

$$DE = 1 + \rho(m-1)$$

where m = the number of participants in each cluster, k = the number of clusters, ESS = effective sample size, DE = design effect (i.e., the correction factor used to calculate the ESS), and ρ = intracluster correlation coefficient.

As seen in Table 8, the ρ values of the present study ranged from 0.00 to 0.17. Within the Coping Scale, there was no intracluster effect so the resultant ESS is the actual sample size of 106. Nearly half of the sample size was reduced with the Knowledge Test (ESS = 58) and unfortunately, over three-quarters of the sample size was lost with both

the SEARS and the SSBS-2. This is a critical issue in terms of statistical power of the study, so the next section will discuss ways to maintain a fair amount of power when sample size, the effect being sought, and the ICC are static.

Table 8

Intracluster Correlation Coefficients (ρ) and Effective Sample Size for each Dependent Measure

Measure	ρ	ESS
SK Knowledge Test	0.03	58
Coping Scale	0.00	106
SEARS-C	0.16	21
SSBS-2	0.17	20

Level of Significance and Power Estimation

Traditionally, psychological research has used an alpha level of .05 and below when conducting tests of significance. (Alpha represents the conditional probability of making a Type I error, or in other words, concluding that an effect exists when in fact there is no effect). Having an alpha at .05 and below has been used to keep the risk of making a Type I error at 5% and less (Howell, 2002). However, alpha levels equal to or below .05 are argued by some to be arbitrary and can be adjusted in order to increase power (Casico & Zedeck, 1983). Increasing alpha can increase the power of the test, or the ability of the study to detect actual effects (i.e., rejecting the null hypothesis when it is in fact false). There are instances in which increasing the alpha level higher than .05 is reasonable, such as when the study is examining small effects or when the sample size is

small. When adjusting the value of alpha, factors such as sample size, size of the effect that is desirable to achieve, and the relative risk of Type I and Type II errors should be considered (Casico & Zedeck).

Determining the size of effect the researchers expects to find is a first step in determining the alpha level for a study. One method for judging how robust of an effect to expect from a given study is by examining previous research (Casico & Zedeck, 1983). Previous studies with the *Strong Kids* series has primarily resulted in small, but meaningful effect sizes (ESs) (Castro-Olivio, 2006; Gueldner, 2006), and earlier research of various SEL programs have found similar effects in magnitude. For example, a meta-analyses by Durlak & Wells (1997) on mental health prevention programs found a mean ES of $d = 0.34$, which is small in magnitude based on Cohen's (1988) suggested standards. A more recent meta-analysis found ESs for universal school-based SEL interventions that ranged from $d = 0.23$ to 0.43 on SEL skills (Durlak & Weisberg, 2005). Given these findings, it seems reasonable to expect at least a small ES with the current study.

Knowing that a small effect is reasonable, the power of the study can be determined. Using the software program GPower (Faul & Erdfeldner, 1992), the power of the current study was calculated. Assuming a small ES ($d = 0.20$ to 0.49) and considering the sample size of 106, the current power of the study is .27 (note that this example ignores the clustering effect discussed in the previous section). This is disappointing, as the current study would detect a true effect only 27% of the time. That means nearly 75% of the time, a false null hypothesis would fail to be rejected and a

potentially beneficial result to the field would be missed. Although an increase in sample size can increase the power of a study, it would require a sample size of 620 to detect a small ES with an alpha level of .05 (Power would be equivalent to 0.80, a reasonable guideline set forth by Cohen (1988)). Obtaining such a large N is difficult to do, as Casico and Zedeck (1983) pointed out that researchers are often faced with too small of a sample instead of one large enough to have strong statistical power. Consequently, other methods of increasing power should be explored.

Because the sample size and ES sought are relatively set, traditional alpha setting at the .05 or lower level is not reasonable with the currently low power of the study. Given this, a *compromise power analysis* can provide a reasonable alternative to increasing power. This approach to alpha setting examines the relative risk of making a Type I error (i.e., false positive: finding an effect that is *not* actually there) versus a Type II error (i.e., false negative: failing to find an effect that *is* actually there) (Erdfeldner, 1984). Within most research, making a Type I error is considered more costly, as researchers would prefer to not find an effect than to falsely find one (Casico & Zedeck, 1983). Alpha is set stringently low, which can make the chances of making a Type II error relatively high, depending on the power of the study. When examining research that has serious consequences having a low alpha level is logical and ethical; say for instances in which a new drug with severe side effects is being tested for its effectiveness. However, Erdfeldner (1984) argues that within basic psychological research, the relative risk of Type I to Type II error should be equivalent (i.e., q = beta/alpha = 1/1) so that the statistical power of a study is not severely compromised, particularly when the study is

examining non-life-threatening factors (Casico & Zedeck, 1983). Within the current study, with alpha set at $p = .05$ and power at .27, the risk of a Type II error is: $\beta = 1\text{-}.27 = .73$. As one can see, the chances of making a Type II error are approximately 14 times more likely than making a Type I error. Given the non-threatening nature of the current study and the small ES presumed to occur, it makes sense to use compromise power analysis to set alpha and beta equal. Although this will raise alpha beyond the conventional .05 level, Hallahan (1999) argued that there is no rational basis for a .05 alpha, so to raise it should not be viewed as "irrational". On the contrary, leaving alpha at .05 and having a beta level equal to .73 seems irrational.

Given the current sample size of 106, the estimated small effect between the treatment and wait-list conditions, and the ratio of alpha to beta of 1 (q = beta/alpha = 1/1), GPower was used to determine the appropriate alpha level, A separate compromise power analysis was conducted for each dependent variable, given the changes in the ESS between them and because separate post-hoc analyses were conducted on the DVs. To compare the changes in power and alpha level, Table 9 displays each dependent measure's power, alpha level, and beta level within each power analysis approach. Using the SK Knowledge Test as an example, on one hand, a traditional alpha level of .05 resulted in a Type II error that is 16 times more likely relative to a Type I error, with less than a 20% chance of finding a true effect. On the other hand, the compromise power analysis resulted in equivalent chances of an alpha and beta error and raised the power to nearly 70%. Overall, the power of the study was approximately tripled for half of the measures and over five times greater for the other two. Although the resulting power for

each measure was still below the .80 convention set by Cohen (1988), the increase in power was a considerable improvement from the .11 to .27 range of power with $p = .05$.

Table 9

Alpha, Beta, and Power between a Standard Power Analysis and a Compromise Power Analysis

	Standard Power Analysis			Compromise Power Analysis		
Measure	α	β	Power	α	β	Power
SK Knowledge Test	.05	.81	0.19	.35	0.35	0.65
Coping Scale	.05	.73	0.27	.30	0.30	0.69
SEARS-C	.05	.89	0.11	.40	0.40	0.59
SSBS-2	.05	.89	0.11	.41	0.41	0.58

Hypothesis Testing

Prior to conducting the analyses for the hypotheses, multicollinearity among the DVs was assessed using Pearson's *r*. Multicollinearity occurs when variables are highly correlated with other (.90 and above) and thus become redundant (Tabachnick & Fidell, 2001). Fortunately, the highest correlation among the pretesting scores on the dependent measures was $r = .47$ (see Table 5). Although the DVs were correlated with each other, they were not redundant and there was no multicollinearity among them.

Because all of the research questions ask the degree to which there were changes among the DVs between treatment conditions, a mixed MANOVA was conducted first to determine if there were statistically significant changes among the groups and their respective scores on the DVs. An advantage of using a MANOVA is that it allows the

researcher to examine multiple DVs at once instead of conducting several univariate analyses. Conducting a MANOVA can avoid the inflated Type I error due to repeated univariate analyses, and it can also have more statistical power to detect an effect by taking into account the correlations among the DVs (Tabachnick & Fidell, 2001). To determine if the groups differed across time on the DVs, a two-way mixed effects MANOVA was conducted with condition (treatment or wait-list) as a between-subject factor and time of assessment (pre-test, post-test, follow-up) as a within-subject factor. After the MANOVA was conducted, pairwise comparisons (i.e., univariate tests) were analyzed to determine if the specific research questions could be answered.

MANOVA Results

Post-hoc power analysis using Wilk's Λ revealed that power was .98, indicating that a true effect would be detected 98% of the time. Consequently, with an alpha level of $p = .05$, beta would equal $\beta = .02$. Because of these low incidences of Type I and Type II error, the traditional alpha level of .05 was used instead of adjusting power with a compromise power analysis. The interaction of condition and time of assessment for the mixed effects MANOVA was significant [Wilk's $\Lambda = .72$, $F(8, 77)$, $p \leq .01$], indicating that the participants' scores on the dependent measures varied as a result of the condition they were assigned to and the time of the assessment. Follow-up univariate tests were examined to investigate the specifics of the interaction and to answer the research questions. The descriptive means and standard deviations for each dependent measure are displayed in Table 10.

Research Question 1

Question 1 asked the extent that the treatment group would have greater gains than the wait-list condition of knowledge of social-emotional learning (SEL) skills, as measured by the SK Knowledge test. By looking at the data "horizontally" (i.e., comparing each group's scores at each assessment period to each other) and "vertically" (i.e., comparing each groups' respective changes in scores over time), it was possible to determine if the results were in the expected direction. (For a table of the means, standard deviations, and the magnitude of difference between the treatment and wait-list at each assessment period, the reader is referred to Table 10). First, groups were compared at both pre-test and post-test (i.e., looking at the data in Figure 9 vertically). The Bonferroni procedure was used to control for alpha slippage (Keppel & Zedeck, 1989). Using the alpha established by the compromise power analysis of $p = .35$, the critical p-value for the pairwise comparisons was $p = (.35/3) = .12$ (the p-value was divided by the total number of comparisons, which was a comparison between groups at each assessment period). The groups did not differ at pre-test, as the pairwise comparison was non-significant ($p = .92$). This indicated that the control and treatment group had similar scores on the SK Knowledge test at that time. Next, scores at post-test were examined and there was a statistically significant difference ($p = .01$), with the treatment group having a higher mean than the wait-list that was moderate in magnitude ($d = 0.73$).

Table 10

Means and Standard Deviations for each Dependent Measure between Conditions

	Pre-testing			Post-testing			Follow-Up		
Questionnaire	M	SD	ES	M	SD	ES	M	SD	ES
SK Knowledge Test									
Treatment	10.39	2.76	0.10	13.27	3.19	**0.73** [b]	13.46	3.30	**0.68** [b]
Wait-list	10.08	3.25		11.00	2.99		11.25	3.17	
Coping Scale									
Treatment	40.20	10.40	0.13	42.25	9.00	**0.67** [b]	40.57	11.86	**0.58** [b]
Wait-list	38.87	9.44		35.69	10.55		33.67	12.03	
SEARS-C									
Treatment	116.31	19.72	**0.29**[a]	120.84	22.55	**0.81**[c]	118.60	25.39	**0.76** [b]
Wait-list	110.56	19.92		101.05	26.07		99.45	25.20	
SSBS-2									
Treatment	37.56	12.40	**0.84**[c]	42.97	12.36	**0.82**[c]	44.54	11.69	**1.13** [c]
Wait-list	27.19	11.85		33.17	11.43		30.96	12.35	

Note: [a] small ES, [b] medium ES, [c] large ES. Bolded numbers indicate a statistically significant result. The ES is a comparison between each condition at that assessment period.

Table 11

Effect Sizes and Difference in Group Means on each Dependent Measure as a Comparison of each Assessment Period

	Pre to Post		Post to Follow		Pre to Follow	
Questionnaire	M	ES	M	ES	M	ES
SK Knowledge Test						
Treatment	2.88	**0.96**[c]	0.19	0.05	3.07	**1.01**[c]
Wait-list	0.92	0.29[a]	0.25	0.08	1.18	0.36[a]
Coping Scale						
Treatment	2.25	0.21[a]	1.68	0.16	0.37	0.03
Wait-list	-3.18	-0.32[a]	-1.98	-0.18	-5.20	**-0.48**[a]
SEARS-C						
Treatment	4.53	0.21[a]	-2.24	-0.09	2.29	0.10
Wait-list	-9.51	**-0.41**[a]	-1.60	-0.06	-11.11	**-0.49**[a]
SSBS-2						
Treatment	5.41	**0.44**[a]	1.57	**0.13**	6.98	**0.58**[b]
Wait-list	5.98	**0.51**[b]	-2.21	-0.19	3.77	**0.31**[a]

Note: [a] small ES, [b] medium ES, [c] large ES; Bolded numbers indicate a statistically significant result. The ES is the change within each group between the two assessment periods.

To determine if the groups differed at post-test because of an increase or decrease in scores, the changes in scores for each group from pre- to post-test were examined (examining the data in Figure 9 from a horizontal perspective). (Table 11 lists all of the effect sizes and differences in each respective group's mean scores for each dependent measure at each assessment period). The Bonferonni-adjusted alpha level was $p = (.35/6) = .06$ (the total number of comparisons was 6 because groups were compared from pre- to post-, post- to follow-up, and pre- to follow-up). The pairwise comparison was significant ($p \leq .01$), as the treatment group had gains of about 3 questions correct on the Knowledge test resulting in a large ES ($d = 0.96$). By contrast, the wait-list group did not significantly change scores ($p = .15$). All together, the treatment group demonstrated positive gains on the SK Knowledge test that the wait-list group did not after they scored similarly at pre-test (see Tables 10 & 11 and Figure 9).

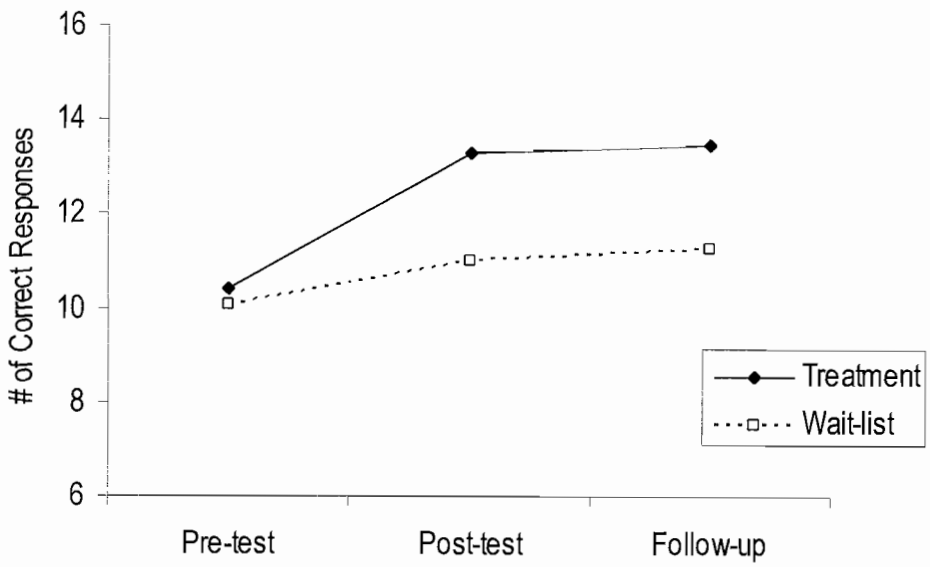


Figure 9. Means of *SK Knowledge Test* for Treatment and Wait-list Condition across Assessment Periods

Research Question 2

Coping Scale. Question 2 examined pre- to post-test gains, but instead of knowledge of SEL skills the question asked about use of SEL skills. It asked if students in the treatment group made larger gains in their self-perceived use of SEL skills compared to the wait-list group. This was measured using both the Coping Scale and the SEARS-C. Looking at the Coping Scale first, the alpha level under the Bonferroni correction method was $p = (.30/3) = .10$. The groups did not differ at pre-test ($p = .46$), but did differ at post-test ($p = .01$). As displayed in Table 10, the treatment group compared to the wait-list group scored higher on the Coping Scale by a medium ES ($d = 0.67$) at post-test (also see Figure 10).

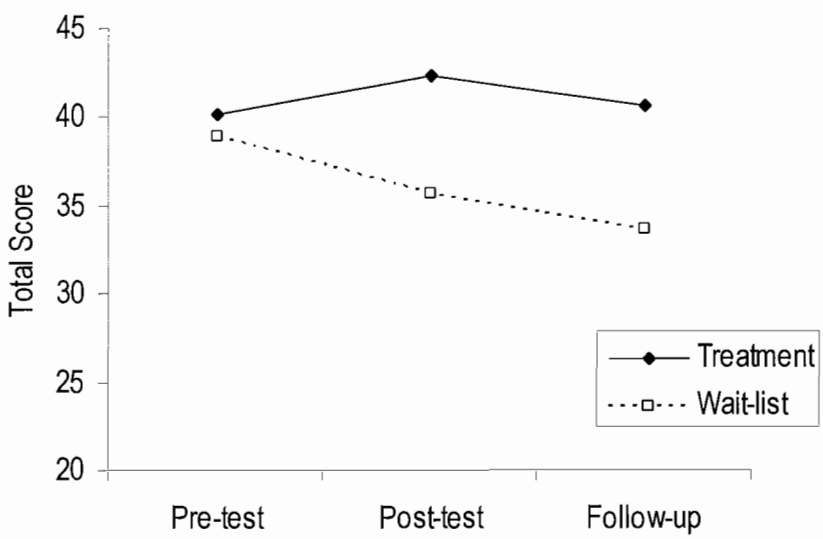


Figure 10. Means of *Coping Scale* for Treatment and Wait-list Condition across Assessment Periods

Examining the pre- to post-test scores for each group respectively [alpha level under Bonferonni method was $p = (.30/6) = .05$], the treatment group actually did not significantly change scores from pre- to post- ($p = .62$), nor did the wait-list group ($p = .06$). However, each group had a change in scores equivalent to a small ES. As displayed in Table 11, the treatment group *increased* scores ($d = 0.21$) and the wait-list group *decreased* scores ($d = -0.32$) from pre- to post-testing, resulting in differences between the two groups at post-testing that was statistically significant and moderate in magnitude (see Table 10). Thus, changes in opposite directions were found with the treatment group having a slight increase in self-perceived skills and the wait-list having a slight decrease, resulting in a statistically significant difference between the groups at post-test (even though they did not differ at pre-test). As to why the significance test did not detect this small change in scores in magnitude, perhaps the power of the study was responsible. For the Coping Scale, power was estimated at 65%.

SEARS-C. For comparisons between conditions based on scores on the SEARS-C, the Bonferroni-adjusted alpha level was $p = (.40/3) = .13$. The pairwise comparisons revealed that the treatment group, at pre-testing, scored slightly higher than the wait-list ($p = .11$) by a small magnitude ($d = 0.29$) (see Table 10 and Figure 11). This makes interpreting the results a bit tricky, as ideally the groups being compared would be similar on the DVs. Other ways to compare the two groups is discussed later on, but analyzing the changes in the ESs can allow one to determine if the changes are above and beyond the pre-existing small differences in magnitude. Not surprising, the groups did differ significantly at post-test ($p \leq .01$), but the difference was more robust, as the treatment

group scored higher on the SEARS-C by a large magnitude ($d = 0.81$). It appears that the difference between the groups changed over time, growing from a small effect to a large effect (see Table 10).

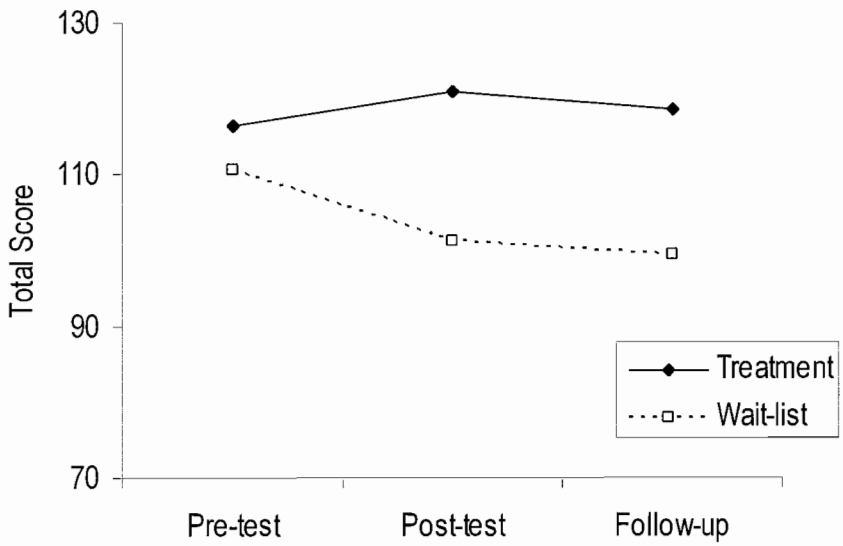


Figure 11. Means of *SEARS-C* for Treatment and Wait-list Condition across Assessment Periods

Next, pre- to post-test changes were examined within each group to determine the directions (e.g., gains or losses) that the groups experienced on the SEARS-C (see Table 11 for the change in each group's mean score). Using a Bonferonni-adjusted alpha level of $p = (.40/6) = .07$, the treatment group *did not* have statistically significantly changes in their scores ($p = .09$). However, they did have a positive gain in scores that was a small in magnitude ($d = 0.21$). The wait-list group *did* have statistically significant changes in scores ($p = .05$), but in a negative direction that was small in magnitude ($d = -0.41$) (see Table 11). Thus, a small *increase* in scores for the treatment group and a small *decrease*

for the wait-list group resulted in a large difference between the groups at post-test (even though both groups had a small difference between them at pre-test).

Given the small difference in scores between the two groups at pre-testing and the relative changes between the conditions over time, more analyses were conducted to determine if the gain for the treatment group was meaningful relative to the change for the wait-list group. To further examine the scores on the SEARS-C, an analysis on the changes in scores was conducted, discussed next.

ANOVA on change in scores. Knowing that the groups had a greater difference between them at follow-up than at pre-test, a reasonable question to ask is the extent to which the changes between the groups over time is significant. Given the small difference between them at pre-test, is the large difference at follow-up a meaningful change? Although earlier research with similar designs of pre- to post- have used an analysis of covariance (ANCOVA) as a way to control for pre-existing differences, this is not an appropriate use of the analysis (Jamieson, 2004; Wright, 2006). Because ANCOVA follows the assumption of random assignment, it assumes that any pre-existing differences are due to chance and not due to the group that they are in. However, without random assignment, any pre-existing differences may *not* be due to chance and instead, are attributable to group membership. Using ANCOVA on quasi-experimental designs, such as the current study, can result in biased results, as the ANCOVA assumes that both groups should have similar pre-test means (in reality, groups may differ on pre-testing means because of a real difference, not random chance) (Jamieson).

Given the inability to use an ANCOVA, an alternative is to calculate a difference score for each participant and then conduct an ANOVA on this score. Despite the difference in pre-testing mean scores, an analysis of the gains score can answer the question, is the average gain in scores different for the two groups? A variable was constructed to represent the *change in scores* by subtracting the pre-test score on the SEARS-C from the post-testing score for each participant. An ANOVA was conducted by entering the condition as the IV and the score on the *change in scores* as the DV. Using a traditional alpha level of $p \leq .05$, the results were significant [$F(1, 99) = 6.74$, $p = .01$; power = .73] as the treatment group had an average change in scores of 2.93 ($sd = 21.47$) and the control group an average change of -7.74 ($sd = 19.82$).

Overall, it appears that the treatment group had small increases in magnitude in their self-perceived use of SEL skills on both the Coping Scale and the SEARS-C, relative to the small decreases in magnitude for the wait-list group. It is possible that the lack of significance for the treatment group from pre- to post-testing is because of the less than ideal power level (power was .59 for the SEARS-C and .69 for the Coping Scale, based on the power analysis using GPower). Still, it is encouraging that the treatment group experienced changes in the anticipated direction. With both groups moving in opposite directions, the pre- to post-test differences changed from a small difference to a large one.

On a side note, the interactions between grade and classroom also make the above interpretation of the SEARS-C scores a bit tentative. In Appendix L, the same analyses

are conducted to determine if the changes in the SEARS-C are consistent across grade and classrooms.

Research Question 3

Question 3 asks if any of the positive gains in the scores for the treatment group maintain at a follow-up assessment period. For the gains made by the treatment group, a non-significant comparison between the post-test and the follow-up test would indicate that any changes from pre- to post-test were not lost (assuming there were changes from pre- to post). The reader is referred again to Table 11 for the changes in each group's mean scores and the accompanying effect size.

SK Knowledge Test. For the SK Knowledge test, the change from post-testing to the follow-up testing was non-significant for the treatment group ($p = .93$) and for the wait-list group ($p = .98$) (critical alpha level was $p \leq .06$). Therefore, it is reasonable to conclude that the gains (or lack thereof) made from pre-testing to post-testing still existed for both groups at the 2-month follow-up. The changes based on ESs were not meaningful for the treatment ($d = 0.05$) or for the wait-list ($d = 0.08$).

Coping Scale. For the Coping scale, the pairwise comparisons from post-testing to follow-up were non-significant for both the treatment ($p = .68$) and wait-list group ($p = .10$) (critical alpha level was $p \leq .05$). As with the SK Knowledge Test, the gains (or decreases) in scores on the Coping Scale maintained at the 2-month follow-up. Also, these changes were not meaningful in terms of magnitude (treatment $d = 0.16$; wait-list $d = -0.18$).

SEARS-C. As with the Knowledge Test and Coping Scale, changes from post-testing to follow-up were non-significant for both the treatment ($p = .38$) and wait-list group ($p = .52$). The critical alpha level was the same as previously used ($p \leq .07$). Changes in ESs were not meaningful either (treatment $d = -0.09$; wait-list $d = -0.06$).

Research Question 4

Question 4 asked the extent to which the *SK* curriculum can provide a preventative effect for students exposed to it. Specifically, do those students who use the curriculum have greater social functioning at the follow-up date than they did at pre-testing, relative to a control group (as measured by the SSBS-2)? This question is confounded by the pre-existing differences between the groups at pre-testing. Using the Bonferroni-adjusted alpha level of $p = (.41/3) = .13$, the treatment and wait-list group differed significantly ($p \leq .01$) by a large magnitude ($d = .84$) at pre-testing. To fully understand the pattern of data, the pairwise comparisons are first discussed, followed by an ANOVA on the change in scores. Figure 12 graphically displays the means for each condition.

Pairwise comparisons. Looking at the differences between the groups at post-testing, the large difference between them at pre-testing maintained at post-testing ($d = 0.82$) and was statistically significant ($p \leq .01$) (see Table 10). At follow-up, the groups were still statistically significantly different ($p \leq .01$), but the difference grew in size from $d = 0.81$ to 1.13. That means the average participant in the treatment ended with greater gains on their SSBS-2 from pre-testing to follow-up compared to the average

participant in the wait-list, even though there was an initial large difference between the two groups.

Next, the data is examined within each group to see how the scores changed over time. Both the treatment ($p \leq .01$) and wait-list group ($p \leq .01$) had statistically significant changes from both pre-testing to post-testing (see Table 11). (The critical alpha level was $p \leq .07$). The treatment had an increase in total score on the SSBS-2 by a small magnitude ($d = 0.44$), and the wait-list group had an increase moderate in magnitude ($d = 0.51$). From post-testing to follow-up, the wait-list group did *not* experience a significant change in scores ($p = .07$), but the treatment group did ($p = .01$). The treatment group had an increase in scores, but interestingly enough, this effect was not meaningful in size ($d = 0.13$) (see Figure 12 and Table 11). Because research

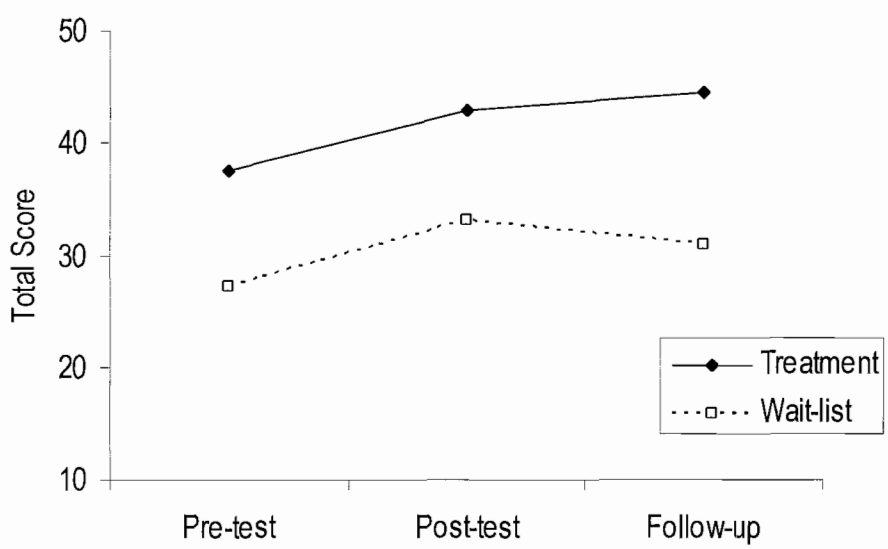


Figure 12. Means of *SSBS-2* for Treatment and Wait-list Condition across Assessment Periods

question 4 compares the groups at pre-testing to follow-up, the main change of interest are those scores from pre-testing to follow-up. Both groups had a statistically significant increase in scores from pre-testing to follow-up, with the treatment group ($p \leq .01$) having a moderate increase in scores ($d = 0.58$) and the wait-list ($p \leq .01$) having a small increase in scores ($d = 0.31$).

To further examine the scores on the SSBS-2, an analysis of the changes in scores was conducted, discussed next.

ANOVA on change scores. Because research question 4 focused on social functioning at the follow-up period, relative to the pre-testing period, a difference score was constructed for each participant by subtracting the *pre-testing score* from the *follow-up score* on the SSBS-2. A one-way ANOVA with condition as the IV and the newly constructed difference score as the DV was conducted. Using a traditional alpha level of $p \leq .05$ to control for Type I error, the results were significant [$F(1, 100) = 6.16, p = .02$; power = .69]. The average mean change in scores for the treatment group was 7.06 (sd = 5.50), relative to 4.19 ($sd = 6.14$) for the control group, which constituted almost a moderate difference in magnitude ($d = 0.49$). Collectively, the pairwise comparisons from the MANOVA and the ANOVA on difference scores indicate that although there were pre-existing differences between the two groups at pre-test, the treatment group experienced greater changes on the SSBS-2 than did the wait-list group.

In addition, Appendix L also analyzed scores on the SSBS-2 by taking into account the apparent interaction effect between grade and classroom. The same analyses

are conducted to determine if the changes in the SSBS-2 are consistent across grade and classrooms.

Summary of Results

Table 12 provides a concise summary of the results. Overall, there were positive gains for the treatment group on the majority of the DVs, as that group had statistically significant increases on the Knowledge Test and on the SSBS-2. The treatment group did not have statistically significant gains on the Coping Scale or the SEARS-C, but did have increases that were small in magnitude. Interestingly enough, the wait-list group had statistically significant *decreases* on those measures. The wait-list had no significant changes on the Knowledge Test, but did have statistically significant increases on the SSBS-2.

Table 12

Summary of Results

Construct	Changes within Groups	Changes between Groups
Knowledge of SEL skills	• Large increase for treatment group, no change at follow-up • No gains for wait-list, no change at follow-up	• Same at pre • Medium ES at post • Medium ES at follow-up
Perception of use of SEL skills (Coping Scale)	• Small increase for treatment group, no change at follow-up • Small decrease for wait-list, no change at follow-up	• Same at pre • Medium ES at post • Medium ES at follow-up
Perception of use of SEL skills (SEARS-C)	• Small increase for treatment group, no change at follow-up • Small decrease for wait-list, no change at follow-up	• Small ES at pre • Large ES at post • Medium at follow-up
Social functioning	• Small increase for treatment group, small increase at follow-up • Medium increase for wait-list, no change at follow-up	• Large ES at pre • Large ES at post • Large ES at follow-up

CHAPTER V

DISCUSSION

The present study examined both the initial and short-term follow-up effects of a social and emotional learning (SEL) curriculum. The SEL curriculum was generally effective, as the treatment group compared to the wait-list group scored higher on measures of SEL knowledge, perceived use of SEL skills, and teacher-rated social functioning at both post-testing and at a two-month follow-up. The study is consistent with prior research that has demonstrated positive effects from school-based SEL programs (see Durlak & Weissberg, 2005 for a review), particularly research using *Strong Kids* (Gueldner, 2006; Merrell et al., 2006). It also extends the research base on *Strong Kids* by demonstrating that the gains made from using the curriculum maintain over a two-month period after post-testing.

Use of SK as a Universal Level of Support

If any SEL curriculum is to be used as a universal level of support for students, then the effects should maintain for a reasonable extent over time. By providing students with relatively permanent and positive changes in their SEL functioning, fewer students will require more intense SEL support within a three-tiered model (Merrell et al., in press). Establishing the use of *SK* as a universal level of support requires first knowing its impact over the long-term. The present study examined this particular issue and asked whether or not those students exposed to the *SK* curriculum would have positive gains

that maintained at a follow-up testing period. The answer to this question proved to be affirmative, as the gains made by the treatment group at post-test held constant at the follow-up testing period. This finding provides initial evidence that *SK* can be used as a universal level of support within a tiered SEL model and provides preliminary evidence that the gains made from the curriculum are retained over time. Earlier research with *SK* has documented some positive effects from pre- to post-test (Gueldner, 2006; Merrell et al., 2006), but the current study is the first to demonstrate a maintenance effect with *SK*. Previous research of SEL curriculums in general has documented the effectiveness of such curriculums over the course of several years (Catalano et al., 2003; CPPRG, 1999; Riggs et al., 2006), so it is possible that the gains in the current study will extend beyond the initial two-month follow-up. Certainly a next step for future research is to lengthen the follow-up period in order to ascertain how long students retain their positive gains.

An interesting finding was that the wait-list group had negative changes on their self-perceived uses of SEL skills (i.e., decreases on the Coping Scale and SEARS-C). Whereas the treatment group demonstrated small increases in scores on the perceived use of SEL measures (albeit non-significant), it was the combination of those gains and the decrease in scores of the wait-list that resulted in rather robust differences between both groups. It was anticipated that the treatment group would have robust increases in scores and that the wait-list would have non-significant changes across the testing periods, but this was not fully supported by the data. One explanation for the decrease in scores of the wait-list is that the groups actually differed and the changes in the scores represented natural changes. After all, students were assigned to classrooms by their previous

teachers based on which teacher they would be a better "match" with, so perhaps the changes were a reflection of students with similar competencies being grouped together. However, the groups did not differ at pre-testing on *all* of the measures, as scores on the Knowledge Test and Coping Scale revealed no meaningful differences between groups at pre-testing. It is curious that any differences among the groups were not apparent in all of the dependent measures, particularly given the moderate correlations among the dependent measures (see Table 5). This points to another explanation for the negative change in scores of the wait-list.

Another possible explanation is some sort of testing effect of the measure itself. It is possible as students were tested repeatedly, they became more aware of their SEL skills. This, in turn, influenced their ratings during the subsequent testing periods (Shadish, Cook, & Campbell, 2002). Perhaps asking questions such as "I try to help others when they need help" (item 5 on the SEARS-C), and "When I have an argument or fight with a friend, I get help from a family member" (item 16 on the Coping Scale), made students more knowledgeable and more critical of their SEL skills over time. Thus, changes on the measures may have been a result of students simply being more aware of their behavior than of actual changes in their perceived use of SEL skills. However, by this logic, one might expect the changes on the measures to be similar across study groups. It is a rather coincidental finding that the treatment group had small gains on the Coping Scale and SEARS-C (albeit non-significant gains) and the wait-list group had small decreases if the changes were due to random testing effects. It seems more

reasonable that the testing effect would have been similar between conditions, so the change in scores may be a reflection of the impact of the curriculum.

SK as a preventative influence. As opposed to a testing effect that explains the changes in scores between the groups, it seems more plausible that the curriculum contributed a protective influence from the stress that students typically experience from school. We know that school can be a stressful experience for students. For example, several researchers have documented that school work and social interactions are among the most stressful events to students in grades K to 12 (Romano, 1997), with tests, school work, oral presentations, and peer relationships as the most stressful events for 4th and 5th graders (Bauwens & Hourcade, 1992; Romano, 1997). We also know that students with more SEL skills and social relationships are at less risk to experience maladaptive outcomes (Abela, 2001; Hilsman & Garber, 1995; Zins, Weissberg et al., 2004). In fact, among a sample of 4th and 5th graders, Dubow and Tisak (1989) documented that academic and behavioral problems are moderated by the degree of social support and problem-solving abilities students have. They concluded that students can build personal resources (e.g., coping/problem-solving skills) and environmental resources (e.g., social networks) that can aid them in dealing with life stressors. Such a process may be at work within the current study. As students in the treatment group participated in the curriculum, they may have developed more SEL skills, social support, and problem-solving skills to manage the stressors they experienced throughout the school year. Conversely, those students in the wait-list (who received the curriculum after post-testing) did not develop corresponding skills. As a result, perhaps they were less prepared

to manage the stress that they encountered. Although the current study did not directly assess stressors within participants, a mechanism at work may be an intra-child process of managing stress and events using skills acquired from the *SK* curriculum.

Still, the differences between the groups were more a result of the wait-list *decreasing* on measures of SEL skills than it was the result of large *increases* of the treatment group. Thus, based on these findings, it appears that the SK curriculum, as the very least, may serve to buffer against or prevent the loss of SEL skills rather than to instill large gains in students (although the treatment group did have small gains in magnitude). Perhaps the curriculum allows students to avoid losing SEL skills or a narrowing of their repertoire when faced with stressors over time. Previous research has illustrated that students dealing with stressors focus their skills on using one or two methods of coping or responses. For example, Beaver (1997) examined elementary students' use of coping skills in response to various stressors and found that students tended to pick only 1 or 2 ways to deal with the stressor. It is possible that the differences between the groups are a result of narrow versus broad methods of dealing with conflicts and stress. Those students in the treatment group may have learned how to manage stressors through a variety of means, but the wait-list did not develop a similarly diverse skill base and instead, relied repeatedly on fewer methods of coping. This issue could account for a decrease in scores on the Coping Scale certainly, as it measures the range of responses students have when faced with a specific stressor.

The idea of a buffering or preventative effect from the use of *SK* has been raised before. Castro-Olivo (2006) implemented a modified version of *SK* with Latino high

school students. Students experienced significant increases in SEL knowledge, and even though they had non-significant changes in their report of acculturation stress, questions were raised about the change in the trajectory of that outcome. Students were assessed twice prior and once after implementation of the curriculum and a visual analysis of the data appeared to show a change from an increase in acculturation stress prior to implementation to a flatter trajectory at post-testing. Whereas it would be ideal for students to experience great gains in SEL skills after receiving the curriculum, perhaps the curriculum does more in the way of preventing the loss of skills and allows students to deal with stress they experience in a better way. Students may experience multiple stressors that place them at risk for various adverse outcomes. With the coping skills and SEL concepts that *SK* teaches, students may learn new and effective ways to manage the stress they experience. This new knowledge and way of handling stress could stave off any negative effects the stress they experience may cause. This "buffering" provided by the curriculum seems plausible in both the current study and the study conducted by Castro-Olivio.

This notion parallels the bullying literature as well. Victims of bullying tend to have a restricted range of responses to bullies over time, as they respond in ways that provide immediate beneficial consequences for them, regardless of the long-term impact of the situation. For instance, victims will give into the bully and become passive in order to end the bullying interaction quickly. However, they do not realize that over time, this actually reinforces the bully's behavior and reinforces their status as a victim (Kochenderfer-Ladd & Skinner, 2002; Olweus, 2001). As they are victimized more and

more, victims respond in more rigid and static means. They essentially lose the ability to respond in different ways that may actually stop them from being victimized again. Perhaps a similar process occurs with students who are not explicitly taught SEL skills. They are faced with stressors over time and begin to narrow their focus of which ones to use, opting for ones that provide short-term relief, but may not necessarily effective in the long-term. Over time, this may limit their range of skills and their ability to manage stress. Even though it is a rather large generalization to apply the findings from the bullying literature to non-victimized populations, the difference in scores between the two groups suggests that the treatment group has a broader range of SEL skills. This notion of stressors and SEL skills provides some interesting fodder for future research, which is discussed later in the chapter.

Observing Social Behavior

The results on the measures of perceived SEL skills raises an interesting question of how best to measure a student's actual use of SEL skills. Use of rating scales are ideal because they capture a broad range of skills and ones that are not easily observed, but they suffer from potential error related to the subjective judgment of raters (Merrell, 1999). Although direct observation of social behaviors is difficult and requires numerous observations to achieve validity (Hintze, 2005; Hintze & Matthews, 2004), there are some possible avenues to explore. From a tiered-model of SEL support, educators would need a brief and efficient means of measuring SEL skills. Although it would likely be too cumbersome to observe all students in a school in order to "benchmark" students into respective levels of support, educators could allocate some resources to observing the

tertiary level and possibly the secondary level of support. The following paragraphs discuss the nature of what to observe in order to best capture SEL.

Social involvement. One possible behavior to observe would be how often a student interacts with peers. The amount of positive social interaction a student engages in with peers is linked to better social competence and school functioning, and researchers assert that effective interactions are the basis for social development (Brown, Odom, & Holcombe, 1996). In addition, social interaction, particularly positive interactions centered around games or activities, arguably encompass all of the SEL skills (Denham & Weissberg, 2004). When interacting with peers, students must manage their own social behaviors and emotions (*self-awareness, self-management)*, as well as navigate others' behaviors and reactions (*social awareness, social management*) in order to achieve desirable ends (*responsible-decision making*). Therefore, it is possible that observing social interactions, particularly cooperative interactions, would capture the entire construct of SEL and its accompanying skills.

One observation system to potentially use is the Social Withdrawal Observation Form (SWOF; Walker, Todi, Block, & Severson, 1988). Moroz and Jones (2002) used the SWOF to measure social involvement by observing students in grades 2 and 3. They observed 3 students at recess for 30-minutes using 10-second partial-intervals and coded any social interaction or engagement between the student and others as *social involvement*, resulting in a percentage of total intervals engaged with others. This observation provided a dynamic and sensitive measure, as Moroz and Jones were able to

document immediate increases in social involvement with peers after they implemented a peer-based praise system.

To further refine observing SEL skills, educators might consider what Brown and colleagues (1996) identified as social goals, behavioral strategies, and the success of interactions among children. In observing social involvement, educators could measure whether or not students social goals are met (e.g., did a student obtain cooperative play with peers after asking to join in on a game?), the behavioral strategies used to achieve that goal (e.g., asking to join a group versus aggressively joining in), and how successful they were at meeting that goal (e.g., did the student's attempts result in the goal he or she set out to obtain?). Observing these behaviors would allow educators to determine how successful students are at interacting with others, and it would provide a clue as to the amount of behaviors they use to engage peers. For example, if a student's request to play a game with others is denied, do they try other methods (e.g., suggesting a new game altogether?) or do they find others to play with? If the idea that *SK* prevents a loss of SEL skills is accurate, or more generally, if SEL curriculums enable students to engage in a broad range of competencies, then an observation system that measures the breadth of skills makes sense to consider. Because this is a more complex observation system requiring more training and proficiency in its use, it may be more suited for students in the tertiary level of support.

SEL skills and GOM. Observing social interactions raises the question of *general outcome measures* (GOMs). GOMs are measures that provide overall indicators of a student's competency within a given skill, and the repeated measurement of that

competency enables educators to measure a student's trajectory of growth (McConnell, McEvoy, & Priest, 2002). One of the most well-researched and established GOMs is that of oral reading fluency for reading (Shinn, 2002). Although there are not as well-established GOMs for SEL competencies as there are in academics, there are initial studies that have investigated the possibility of using direct observations to create a GOM for social behaviors. Cummings, Kaminski, and Merrell (in press) researched the idea of a GOM for social competence, defined in this case as an evaluative judgment of how well a person performs social tasks (see Merrell, 1999). They have developed the Initiation Response Assessment (IRA) and documented initial reliability and validity for it. The IRA involves observing students for 10.5 minutes during an analogue cooperative activity and using frequency counts and time sampling to record social interactions (goal-directed or non-goal directed), helping or encouraging behaviors, task engagement, and negative behaviors (e.g., grabbing toys, whining). Across 8 observation periods, the slopes for the presence of negative behaviors was statistically significantly related to scores on the SSBS-2 completed by the students' teachers, providing some initial evidence that the IRA potentially could be developed into a dynamic measurement of social competence. Such a dynamic method of measurement could provide a formative assessment for social competence and SEL, providing an assessment method that would allow educators to measures one's growth and response to SEL-based interventions.

However, the early literature on GOMs and social behaviors must be cautioned with the state of the literature on observing social behavior. The amount of direct observations needed to establish reliability of social behaviors is arduous, requiring

several observations to achieve satisfactory reliability. For example, Hintze and Matthews (2004) examined the reliability of direct observation when observing "on-task" behavior. They found that observing students twice a day for two weeks resulted in reliability of less than $\alpha = .65$. What is more problematic is that they estimated it would take *4 observations a day for 4 school weeks* in order to achieve adequate reliability ($\alpha \geq .90$). Certainly this is a barrier for the current research on GOMs, as any developed GOM must condense 4 weeks worth of observations into one efficient method. In addition, social behavior is extremely sensitive to contextual changes (Merrell, 1999). Hintze and Matthews found that rates of on-task behavior varied across time of day and between settings, providing another consideration when observing students. Although it is encouraging that the research is underway with GOMs and social behaviors, developing an efficient and practical direct observation measure of SEL skills is still a long way off from being well established.

Role of Teacher Behavior

The study also provides further support of the use of the key factors to success outlined in the literature review earlier (fidelity, leadership support, length of the use of the program, and generalization). The study identified teacher treatment fidelity above 75% and had indirect support from the administration, as the school's principal was the first person to agree to the study. The study targeted generalization through the use of praise, pre-correction, and homework, and although the curriculum itself stopped at the booster session, the classrooms continued to displayed the pre-correction posters and discussed the curriculum throughout the year. A more extensive summary of key factors

is outlined by Greenberg, Weissberg, and colleagues (2003). They listed those elements as: (1) a program that is cognitive-behavioral and interactive, (2) a multi-year and multi-domain effort, (3) use of sequenced instruction, (4) a structured curriculum, (5) opportunities to apply the SEL skills in various activities, and (6) changes in a school's ecology, teacher's behavior, and classroom climate. They also add the importance of involving parents and the community, and monitoring the delivery and effectiveness of the program. Although the current study did not employ all of these factors, particularly that the program was not used across years and only with a few classrooms, it is promising that beneficial results can be found on a smaller scale. This raises the question as to what are the *critical* key factors. In other words, are there certain key factors that are more critical for success than others?

It is possible that one of the more crucial factors is the role of the teacher. Teachers were asked to both praise and pre-correct students on the SEL skills and lessons that were being taught to them. This seems like it would be crucial because of its support in other areas in which students are being taught new skills. Based on previous research, we know that the use of praise and pre-correction can affect numerous behaviors, including increasing time on-task (Chalk & Bizo, 2004), decreasing negative behaviors, such as yelling or non-compliance (Stormont, Smith, & Lewis, 2007), managing disruptive behaviors during transition times (Colvin, Sugai, Good, & Lee, 1997), and improving reading performance (Miao, Darch, & Rabren, 2002). The influence of the teacher's behavior also makes sense in light of the relation between fidelity, a factor largely controlled by teachers, and the positive effects of a curriculum (Kam et al., 2003).

As research moves forward with SEL programming, investigating the role of teacher praise and pre-correction is a worthy goal because of the flexibility and "free" cost of using those components. It would be important to determine the threshold needed for change, that is, what amount of pre-correction and praise is needed to influence students' SEL. The present study had approximately 3 to 4 praises and pre-corrects each week, so ascertaining how much (or how little) of these components is needed would be helpful in determining how intense a program should be. This would have particular impact at the secondary and tertiary levels of support, as this would be an easy way to intensify the SEL programming.

Limitations

The results must be interpreted within the context of the study. The small and fairly homogenous sample makes it difficult to generalize the findings too much, particularly given that students from different backgrounds may experience different stressors (Kilmer, Cowen, Wyman, Work, & Magnus, 1998). This would have an impact on the protective nature of any SEL program and raises questions as to how beneficial the curriculum would be with more diverse students and with those that experience more severe risk factors.

Another limitation is the lack of the booster session from one of the classrooms in the treatment condition. Although both classrooms received all 12 lessons, it is unclear the difference not having a booster session had. However, there is reasonable confidence that the booster session did not have a significant impact, as there were not statistical differences between the treatment condition classrooms. A series of *t*-tests conducted

comparing both classrooms at pre-testing, post-testing, and follow-up revealed that the only significant difference was on the SEARS-C at post-testing [t = -2.74; F(51, 46.01) = 3.05, p =.01], a score that would not have been affected by the lack of a booster session.

The findings are also limited by the nesting effects found within the sample. Although attempts were made to understand the nature of the effects (see Appendix L), it is hard to determine if the pre-existing effects were due to random chance, the limited sample size, or some confounding classroom effect. Nesting and clustering are considerable barriers for research in applied settings (Killip et al., 2004), but as research moves forward within this field, larger sample sizes and analyses at the level of the cluster (e.g., classroom or school) can help facilitate meaningful results given these barriers.

Practical Implications

The most apparent practical implication from the present study is that *Strong Kids* may be recommended as a universal-level SEL curriculum. Although the evidence is preliminary, it is encouraging that positive gains were documented from the study two months after post-testing. If future research can document that the gains maintain over longer time periods, then schools could add another option to the list of possible programs from which to build a school-wide model of SEL support. Other universal SEL programs include the PATHS curriculum (CPPRG, 1999) and the Raising Healthy Children program (RHC; Catalano et al., 2003). Although these are great SEL programs that have demonstrated similar potential, *Strong Kids* has an advantage because of its relatively low adaptation cost. For example, the PATHS curriculum has over 50 lessons

and the RHC calls for teacher training in instruction and classroom management, monthly "booster" sessions for teachers, and a component for parent training. With only 12 lessons that can be implemented by the teacher, the time commitment and "set up" of *SK* is less compared to other programs. This variety in universal programs, while sticking to the factors outlined by Greenberg, Weissberg, and colleagues (2003), meets the diversity in needs of schools, as some may choose one program over another based on preference, need, and resources.

Future Research

The current study provides some direction for future research within the field. One of the first implications is that research examining longer follow-up periods for *SK* is needed. Although the current study had ethical obligations (i.e., time was needed to allow the wait-list to run the 12 weekly lessons) and could not push the follow-up period beyond two months, studies in which the effects of *SK* are measured over the course of a year are needed to provide stronger support for use of *SK* as a universal SEL support.

In terms of the field of SEL, questions about the critical factors that lead to positive results were raised. Is there a minimal number of factors that SEL programs need to be effective? Can a single factor that provides maximum change be identified? The factors outlined by Greenberg, Weissberg and colleagues (2003) and by the literature review are fairly extensive, so knowing if a smaller set can produce equally powerful results could go a long way in terms of providing schools with realistic means of SEL implementation. The present study used praise and pre-correction to boost its impact, but did not determine how much of an influence this factor had. Other literature suggests that

classrooms in which positive praise to redirects of 4:1 are more effective in terms of on-task behavior and academic outcomes (Brophy & Good, 1986), but would this ratio translate to the SEL literature? Perhaps future research could ascertain the amount of praise and pre-correction needed to make positive changes, as no study to date has considered praise, pre-correction, and SEL conjunctively.

Another interesting research question is what leads to positive changes more, teacher behavior, as measured by implementation of parts of the curriculum, or student engagement in the curriculum? If students are being provided a protective quality from the curriculum and are learning more SEL skills, then on some level the amount of student "buy-in" is important. A study that examines which is related more to outcomes would be important. Is teacher behavior or student engagement more critical? Fortunately, the data currently being collected on this study's wait-list group may provide an answer to that very question. The author is measuring both teacher fidelity of components of the curriculum and study engagement (measured by on-task behavior and homework completion) to determine if one or the other is a better predictor of positive student outcomes. The results could provide more answers for schools as to which facet of SEL implementation is more critical to focus on.

Finally, as mentioned earlier, future research should consider the "buffering" effect of *SK,* or more generally, of SEL programs. Past studies have examined the gains made from use of a curriculum, but none appear to measure the current state of risk factors students are exposed to in relation to the changes of the curriculum. If a student makes gains in SEL competencies, but also has increases in risk factors, what is the

practical benefit of the curriculum? As research moves forward with SEL programs, measurement of the risk factors and stress students experience over time should be taken into account in order to get a complete picture of the protection factor students may be provided.

Conclusion

The present study has provided initial evidence of the protective nature of the *Strong Kids* curriculum with a sample of 3rd and 4th graders. Students exposed to the treatment group had gains on SEL competencies that maintained over time, relative to a wait-list that had little to no change in competencies, as well as an overall *loss* of scores on some SEL measures. Such findings provide hope for the practicality of school-based SEL programming within a tiered-model. As schools work to integrate academic, behavioral, and social-emotional programs, it appears possible that brief and efficient SEL programs such as *Strong Kids* can provide a universal foundation from which to build more intensive supports.

APPENDIX A:

INFORMED CONSENT FORM

FOR TREATMENT GROUP

UNIVERSITY OF OREGON

Professor Kenneth W. Merrell
School Psychology Department
5208 University of Oregon
Eugene, OR 97403-5208
Phone: (541) 346-2414
Email: kmerrell@uoregon.edu

Office for Protection of Human Subjects
Riverfront Research Park
1600 Millrace Drive, Suite 105
5237 University of Oregon
Eugene, OR 97403-5237
(541) 346-2510

Dear Parent/Guardian:
Your student's school will be using a program called *Strong Kids*. This program teaches *resiliency skills*. Resiliency skills are the skills that help children handle everyday stress and social situations in a positive manner.

Yolanda Elementary has volunteered to be a part of a research study to evaluate the effectiveness of this program. The study is being conducted by Jason E. Harlacher, a doctoral student at the University of Oregon, and supervised by Dr. Ken Merrell, a professor at the University of Oregon. Your child was selected as a possible volunteer because your child's teacher is going to use the program. The program consists of one 45-minute lesson per week for 12 weeks beginning the week of September 24th. One more lesson is held in February.

To check if students are learning resiliency skills from this program, your child be given 3 short questionnaires right before *Strong Kids* starts and at the end of 12 weeks. Your child will also be given the questionnaires in February. Participation is voluntary. Each questionnaire takes about 15 minutes to complete. All of the questionnaires ask what your child has learned from the program and about good behaviors, such as asking for help from others. Your child's teacher will also complete a questionnaire on how your child plays with other students.

Your child will choose a small prize, like a pencil or eraser, each time they are given a questionnaire. Your child's classroom will also get a lunch party with a 6-foot Subway sandwich as a thank you for being in the study (your child does not have to actually fill out the questionnaires to get the free lunch).

Your decision whether or not to participate will not affect your relationship with your child's district, school, teacher, or with the University of Oregon. If you decide that your child will not participate in the study, a supervised and structured activity will be provided for your child when the other students are given the questionnaires. If you decide to participate, you may still withdraw your consent and stop your child's participation at any time without penalty. Your child may also choose to stop filling out the questionnaires at any time.

Are the questionnaires confidential?
Yes! Your child's answers are confidential. Your child will be given a code to use to fill out the questionnaires (the teacher will also use this number when filling out questionnaires). Your child's name will not appear on the questionnaire at any time. A list with each child's name and code number will be kept by the researchers. To keep anyone from finding out a child's number, the list will kept on a computer with a password. Your child's teacher will only see the form when he/she fills out his/her questionnaires and will not see it after that.

Questions?
If you have any questions about the questionnaires, please contact me, Jason Harlacher, the main researcher of this project. I can be reached at (801) 403-9494 or at jharlach@uoregon.edu. Dr. Kenneth Merrell is also available to answer your question. He can be reached at (541) 346-2414 or kmerrell@uoregon.edu.

UNIVERSITY OF OREGON

Professor Kenneth W. Merrell
School Psychology Department
5208 University of Oregon
Eugene, OR 97403-5208
Phone: (541) 346-2414
Email: kmerrell@uoregon.edu

Office for Protection of Human Subjects
Riverfront Research Park
1600 Millrace Drive, Suite 105
5237 University of Oregon
Eugene, OR 97403-5237
(541) 346-2510

This study has been approved by the school district and the school's principal. It is also approved by your child's teacher and by the Office for the Protection of Human Subjects (OPHS) at the University of Oregon. If you have questions about your child's rights as a research volunteer, contact the OPHS at (541) 346-2510. This Office looks over research to protect your rights and is not involved with this study.

Thank you for your time,

Jason Harlacher, MS, NCSP
University of Oregon Doctoral Student
(801) 403-9494

Kenneth W. Merrell, PhD
University of Oregon Professor
(541) 346-2414

IF YOU GIVE PERMISSION FOR THE QUESTIONNAIRES:
I understand that by **NOT** responding to this letter, I am allowing my child to participate in the study.

IF YOU DO NOT GIVE PERMISSION FOR THE QUESTOINNAIRES:
If you do not give your consent for your child's participation in this study, please sign the bottom portion of this form and return it to your student's teacher by **September 19th, 2007.**

I DO NOT give consent for my child (name)_____ to participate in this study

Print Parent/Legal Guardian name: _____

Parent/Legal Guardian Signature: _____ Date:_____

APPENDIX B:

INFORMED CONSENT FORM

FOR WAIT-LIST GROUP

UNIVERSITY OF OREGON

Professor Kenneth W. Merrell
School Psychology Department
5208 University of Oregon
Eugene, OR 97403-5208
Phone: (541) 346-2414
Email: kmerrell@uoregon.edu

Office for Protection of Human Subjects
Riverfront Research Park
1600 Millrace Drive, Suite 105
5237 University of Oregon
Eugene, OR 97403-5237
(541) 346-2510

Dear Parent/Guardian:
Your student's school will be using a program called *Strong Kids*. This program teaches *resiliency skills*. Resiliency skills are the skills that help children handle everyday stress and social situations in a positive manner.

Yolanda Elementary has volunteered to be a part of a research study to evaluate the effectiveness of this program. The study is being conducted by Jason E. Harlacher, a doctoral student at the University of Oregon, and supervised by Dr. Ken Merrell, a professor at the University of Oregon. Your child was selected as a possible volunteer because your child's teacher is going to use the program. The program consists of one or two 45-minute lessons per week for 12 weeks beginning in February.

To check if students are learning resiliency skills from this program, your child be given 3 short questionnaires in September, in December, and in February. Participation is voluntary. Each questionnaire takes about 15 minutes to complete. All of the questionnaires ask what your child has learned from the program and about good behaviors, such as asking for help from others. Your child's teacher will also complete a questionnaire on how your child plays with other students.

Your child will choose a small prize, like a pencil or eraser, each time they are given a questionnaire. Your child's classroom will also get a lunch party with a 6-foot Subway sandwich as a thank you for being in the study (your child does not have to actually fill out the questionnaires to get the free lunch).

Your decision whether or not to participate will not affect your relationship with your child's district, school, teacher, or with the University of Oregon. If you decide that your child will not participate in the study, a supervised and structured activity will be provided for your child when the other students are given the questionnaires. If you decide to participate, you may still withdraw your consent and stop your child's participation at any time without penalty. Your child may also choose to stop filling out the questionnaires at any time.

Are the questionnaires confidential?
Yes! Your child's answers are confidential. Your child will be given a code to use to fill out the questionnaires (the teacher will also use this number when filling out questionnaires). Your child's name will not appear on the questionnaire at any time. A list with each child's name and code number will be kept by the researchers. To keep anyone from finding out a child's number, the list will kept on a computer with a password. Your child's teacher will only see the form when he/she fills out his/her questionnaires and will not see it after that.

Questions?
If you have any questions about the questionnaires, please contact me, Jason Harlacher, the primary researcher of this project. I can be reached at (801) 403-9494 or at jharlach@uoregon.edu. Dr. Kenneth Merrell is also available to answer your question. He can be reached at (541) 346-2414 or kmerrell@uoregon.edu.

UNIVERSITY OF OREGON

Professor Kenneth W. Merrell
School Psychology Department
5208 University of Oregon
Eugene, OR 97403-5208
Phone: (541) 346-2414
Email: kmerrell@uoregon.edu

Office for Protection of Human Subjects
Riverfront Research Park
1600 Millrace Drive, Suite 105
5237 University of Oregon
Eugene, OR 97403-5237
(541) 346-2510

This study has been approved by the school district and the school's principal. It is also approved by your child's teacher and by the Office for the Protection of Human Subjects (OPHS) at the University of Oregon. If you have questions about your child's rights as a research volunteer, contact the OPHS at (541) 346-2510. This Office looks over research to protect your rights and is not involved with this study.

Thank you for your time,

Jason Harlacher, MS, NCSP
University of Oregon Doctoral Student
(801) 403-9494

Kenneth W. Merrell, PhD
University of Oregon Professor
(541) 346-2414

IF YOU GIVE PERMISSION FOR THE QUESTIONNAIRES:
I understand that by **NOT** responding to this letter, I am allowing my child to participate in the study.

IF YOU DO NOT GIVE PERMISSION FOR THE QUESTOINNAIRES:
If you do not give your consent for your child's participation in this study, please sign the bottom portion of this form and return it to your student's teacher by **September 19th, 2007.**

I DO NOT give consent for my child (name)_____ to participate in this study

Print Parent/Legal Guardian name: _____

Parent/Legal Guardian Signature: _____ Date:_____

APPENDIX C:

STUDENT ASSENT FORM FOR

TREATMENT GROUP

Dear Student:

I am a student at the University of Oregon. Your teacher is using a program called *Strong Kids* that teaches students how to stay strong. Being strong means being friends with others. Your teacher has agreed to help me learn how to help students stay strong. Your class will begin using *Strong Kids* in September.

To find out how students stay strong, your teacher will give you three tests that take about 10 – 15 minutes. After 12 weeks, your teacher will give you three more tests. You will also take these tests in February. Filling out these tests should help us learn what you need to know to stay strong.

These tests ask about the things you do with other students and about things you may learn from *Strong Kids.* Your parents have told us it is okay if you take these tests. You do not have to fill out the tests and if you decide not to, you will not get into any trouble. If you choose to fill them out, you can stop at any time. You won't get in any trouble for stopping.

The scores you get on them are not counted on your report cards and will not count towards your grade. You will use a code number instead of your name on the tests, so no one will now whose work it is. The code number will only tell us if you are a girl or a boy, your grade, your age, and what race you are. Your answers will be kept a secret.

If you have questions, you can ask your parents. You can also ask your teacher or you can call me. My name is Jason Harlacher. My number is: 541-346-2414. You can also ask my teacher, Kenneth Merrell at 541-346-2414.

Sincerely,

Jason Harlacher, MS

I, _____, have decided to be in this project.
Signature

APPENDIX D:

STUDENT ASSENT FORM FOR

WAIT-LIST GROUP

Dear Student:

I am a student at the University of Oregon. Your teacher, _____, is using a program called *Strong Kids* that teaches students how to stay strong. Being strong means being friends with others. Your teacher has agreed to help me learn how to help students stay strong.

To find out how students stay strong, your teacher will give you three tests that take about 10 – 15 minutes. After 12 weeks, your teacher will give you three more tests. You will also take these tests in February. Filling out these tests should help us learn what you need to know to stay strong.

These tests ask about the things you do with other students and about things you may learn from *Strong Kids*. Even though your class is not using *Strong Kids* right now, these tests will help us learn how to help children your age. Your parents have told us it is okay if you take these tests. You do not have to fill out the tests and if you decide not to, you will not get into any trouble. If you choose to fill them out, you can stop at any time. You won't get in any trouble for stopping.

The scores you get on them are not counted on your report cards and will not count towards your grade. You will use a code number instead of your name on the tests, so no one will now whose work it is. The code number will only tell us if you are a girl or a boy, your grade, your age, and what race you are. Your answers will be kept a secret.

If you have questions, you can ask your parents. You can also ask your teacher or you can call me. My name is Jason Harlacher. My number is: 541-346-2414. You can also ask my teacher, Kenneth Merrell at 541-346-2414.

Sincerely,

Jason Harlacher, MS

I, _____, have decided to take part in this project.
Signature

APPENDIX E:

STRONG KIDS KNOWLEDGE TEST

Part Two: Strong Kids Knowledge Test

Directions: This test has 20 questions about healthy and unhealthy ways to express feelings, thoughts, and behavior. Read each question carefully and pick what you think is the best answer.

TRUE-FALSE. Read each sentence. If you think it is true or mostly true, circle the T, which means "true." If you think it is false or mostly false, circle the F, which means "false."

1. T F Self-esteem is your feelings of worth for yourself.

2. T F When identifying a problem, it is important to describe how you feel and then listen to how the other person says they feel.

3. T F When people feel embarrassed, they are likely to stand tall, smile, and talk to others.

4. T F Clenched fists and trembling or shaking hands are often signs of stress.

5. T F Your friend took the last ice cream bar at the class party and you hadn't gotten one yet. A good way to deal with this is to first identify how you feel, figure out if you feel comfortable or uncomfortable, and then choose 3 positive ways to express your feeling.

MULTIPLE CHOICE. Circle the letter that goes along with the best answer for each question.

6. Devin's gym teacher tells him to try out for the basketball team. Devin thinks that he is too short and won't make it, so he decides to not try out for the team. What thinking error is Devin making?
   a. Binocular vision
   b. Black and white thinking
   c. Making it personal
   d. Fortune telling

7. An example of an emotion that is uncomfortable for most people is
   a. Excited
   b. Frustrated
   c. Curious
   d. Content

8. What is an emotion?
   a. A thought you have about a situation
   b. Your inner voice inside your head
   c. A memory you have about something that happened to you
   d. A feeling that tells you something about a situation you are in

9. Self-talk is a way to calm down after you get angry. Self-talk includes telling yourself
   a. I don't deserve this
   b. I should get angry when something like this happens
   c. I can work through this
   d. I need to stop getting angry so often

10. Which of the following statements best describes empathy?
    a. Knowing how you are feeling
    b. Not knowing why another person is feeling sad
    c. Understanding another person's feelings
    d. Thinking about another person

11. What is the meaning of the thinking error dark glasses?
    a. Looking at the whole picture
    b. Seeing only the part of a situation that makes you sad
    c. Trying to see things in a different way
    d. Thinking about only the negative or bad parts of things

12. Thinking errors occur when
    a. You see things differently than what really happened
    b. You see both the good and bad of each situation
    c. You think something different than your friend
    d. You tell yourself you shouldn't try to do something

13. Reframing is a way to
    a. See the whole picture
    b. Think about the things that make you smile
    c. Think about the situation more realistically
    d. Think about what you will do next

14. Why would you want to know how someone else is feeling?
    a. So you can leave them alone when they're angry
    b. To better understand that person's feelings
    c. To tell other people about that person
    d. To act the same when you are together

15. What does the ABCDE plan for optimism help you to do?
    a. Look at both sites of a situation
    b. View situations more positively
    c. Control your positive and negative thoughts
    d. Realize that you sometimes have no control over things

16. Conflict resolution is best described as
    a. Discussing a problem until there is a winner and a loser
    b. Arguing with another person until they see your point and give in
    c. Problem-solving so you can reach an agreement
    d. Talking about the problem until something changes the other person's mind

17. Which of the following is a positive way to express how scared you are for your parents to get your report card?
    a. Tell them why you are scared
    b. Hide your report card
    c. Tell your parents they are expecting too much from you
    d. Say that your grades were bad because other kids at school distracted you

18. Why is it important to make an agreement when you are trying to solve a problem?
    a. To understand what the other person is feeling
    b. To let the other person know what you think about the problem
    c. To make sure both people accept the solution to the problem
    d. To solve the problem more quickly

19. Which of the following is one of the best ways to deal with a problem with you are feeling stressed?
    a. Crying
    b. Talking about the problem with a friend
    c. Complaining to your mom
    d. Ignoring the problem

20. Which of the following is the better way to deal with feeling very angry when the person next to you in class keeps talking and annoying you?
    a. Yell at them and tell the to stop
    b. Call out to the teacher about the student
    c. Take their backpack to get even
    d. Stop, count to ten, and try to relax

APPENDIX F:

CAUSEY AND DUBOW'S COPING SCALE

The Coping Scale

Here is a list of behaviors or things to do after a situation occurs. Read the situation below and, thinking of what you normally do in that situation, rate how often you do the behavior. There are no right or wrong answers; just mark how often you do the behavior.

Circle "1" if you *Never* do that behavior, "2" for *Almost Never*, "3" for *Sometimes*, "4" for *Almost Always*, and "5" for *Always*.

Think of this situation and respond to the following items.
"When I have an argument or fight with a friend, I..."

	Never	Almost Never	Sometimes	Almost Always	Always
1. try to think of different ways to solve it.	1	2	3	4	5
2. tell a friend or family member what happened.	1	2	3	4	5
3. talk to somebody about how it made me feel.	1	2	3	4	5
4. change something so things will work out.	1	2	3	4	5
5. get help from a friend.	1	2	3	4	5
6. decide on one way to deal with the problem and do it.	1	2	3	4	5
7. do something to make up for it.	1	2	3	4	5
8. ask a friend for advice.	1	2	3	4	5
9. ask someone who has had this problem what he or she would do.	1	2	3	4	5
10. know there are things I can do to make it better.	1	2	3	4	5
11. ask a family member for advice.	1	2	3	4	5
12. talk to the teacher about it.	1	2	3	4	5
13. go over in my mind what to do or say.	1	2	3	4	5
14. try extra hard to keep this from happening again.	1	2	3	4	5
15. try to understand why this happened to me.	1	2	3	4	5
16. get help from a family member.	1	2	3	4	5

APPENDIX G:

SOCIAL EMOTIONAL ASSETS AND RESILIENCY SCALE FOR CHILDREN

SEARS-C

SOCIAL-EMOTIONAL ASSETS AND RESILIENCE SCALE
Item Tryout Research Version-Child Form (for Grades 3-6)

PART 1: ABOUT ME

Your Name (or ID#) _____ Sex: ☐ Girl ☐ Boy

School _____ Grade _____ Age _____

Race/Ethnic Group: ☐ White ☐ African American ☐ Hispanic or Latino ☐ Asian
☐ Native American ☐ Other _____

PART 2: DIRECTIONS

Here is a list of sentences that tell how kids sometimes feel, think, or act. Read each sentence, and circle the letter that tells about you the best. Circle **N** if the sentence is **NEVER** true for you. Circle **S** if the sentence is **SOMETIMES** true for you. Circle **O** if the sentence is **OFTEN** true for you. Circle **A** of the sentence is **ALWAYS** (or ALMOST ALWAYS) true for you. There are no right or wrong answers. Please answer every sentence, and do your best.

Remember: NEVER SOMETIMES OFTEN ALWAYS

		0	1	2	3
1.	I like to do my best in school.	N	S	O	A
2.	I feel sorry for others when bad things happen to them.	N	S	O	A
3.	I am good at understanding what other people think.	N	S	O	A
4.	I can do many things without help.	N	S	O	A
5.	I try to help others when they need help.	N	S	O	A
6.	People like to be with me.	N	S	O	A
7.	I like to talk to lots of different people.	N	S	O	A
8.	I make friends easily.	N	S	O	A
9.	I try to understand how my friends feel when they are upset, or sad.	N	S	O	A
10.	I am a good listener when other people have something to say.	N	S	O	A
11.	Other kids ask me to hang out with them.	N	S	O	A
12.	I have a best friend.	N	S	O	A
13.	People think I am fun to be with.	N	S	O	A
14.	Even when things don't go well for me, I am okay.	N	S	O	A
15.	Other people like me.	N	S	O	A
16.	My friends come to me for help.	N	S	O	A
17.	I like doing things for others.	N	S	O	A
18.	I am good at solving problems.	N	S	O	A
19.	I understand how other people feel.	N	S	O	A

20.	I feel okay with the way I am	N	S	O	A
21.	It is easy for me to tell people how I feel	N	S	O	A
22.	I ask for help when I need it	N	S	O	A
23.	I could teach someone else how to calm down when they are angry	N	S	O	A
24.	I know the difference between anger and aggression	N	S	O	A
25.	I stay in control when I am angry	N	S	O	A
26.	I care what happens to other people	N	S	O	A
27.	I think before I act	N	S	O	A
28.	I am comfortable when I am with a group of people	N	S	O	A
29.	Other people see me as a leader	N	S	O	A
30.	I like who I am	N	S	O	A
31.	Other kids respect me	N	S	O	A
32.	I make good decisions	N	S	O	A
33.	I think about my problems in ways that help	N	S	O	A
34.	I like to go to school	N	S	O	A
35.	I am happy to be me	N	S	O	A
36.	I feel good about myself	N	S	O	A
37.	I can handle most things on my own	N	S	O	A
38.	I ignore other kids when they tease me or call me names	N	S	O	A
39.	I stand up for myself when I need to	N	S	O	A
40.	I can name lots of different feelings	N	S	O	A
41.	I know how to calm down when I am upset	N	S	O	A
42.	I know how to change my negative thoughts	N	S	O	A
43.	When life is hard, I don't let things get to me	N	S	O	A
44.	I know how to set goals for what I want in life	N	S	O	A
45.	I am able to handle problems that really bother other kids	N	S	O	A
46.	I ask my teacher for help when I do not understand my school work	N	S	O	A
47.	My parents trust me	N	S	O	A
48.	I work well with other kids on school projects	N	S	O	A
49.	I like being at school	N	S	O	A
50.	I can identify errors in the way I think	N	S	O	A
51.	I know when people are upset, even when they do not talk about it	N	S	O	A
52.	I stay calm when there is a problem or argument	N	S	O	A
		0	1	2	3

Remember: **N**EVER **S**OMETIMES **O**FTEN **A**LWAYS

_____
TOTAL

APPENDIX H:

SCHOOL SOCIAL BEHAVIOR SCALE- SECOND EDITION

School Social Behavior Scale

Date: _____

Please rate this student's behavior using all of the items. Ratings should be based on your observations of this student's behavior **during the past three months.** The rating points after each item are based on the following format:

Never If the student does not exhibit a particular behavior, or if you have not had an opportunity to observe a particular behavior, circle 1, which indicates *Never*.

Frequently If the student often exhibits a particular behavior, circle 5, which indicates *Frequently*.

Sometimes Circle the numbers 2, 3, or 4 (which indicate *Sometimes*) if the student exhibits the behavior somewhere in between the two extreme rating points, based on your judgment of how frequently it occurs. The rating points after each item appear in the following format:

Please complete all items, and do not circle between numbers.

		Never		Sometimes		Frequently
1.	Offers help to other students when needed	1	2	3	4	5
2.	Participates effectively in group discussions and activities	1	2	3	4	5
3.	Understands problems and needs of other students.	1	2	3	4	5
4.	Invites other students to participate in activities	1	2	3	4	5
5.	Has skills or abilities that are admired by peers	1	2	3	4	5
6.	Interacts with a wide variety of peers	1	2	3	4	5
7.	Is good at initiating or joining conversations with peers	1	2	3	4	5
8.	Is sensitive to feelings of other students	1	2	3	4	5
9.	Enters appropriately into ongoing activities with peers	1	2	3	4	5
10.	Has good leadership skills	1	2	3	4	5
11.	Notices and compliments accomplishments of others	1	2	3	4	5
12.	Is assertive in an appropriate way when he/she needs to be	1	2	3	4	5
13.	Is invited by peers to join in activities	1	2	3	4	5
14.	Is "looked up to" or respected by peers	1	2	3	4	5

APPENDIX I:

TRAINING AGENDA AND EXAMPLES

OF PRE-CORRECTION AND PRAISE

Agenda:

1) Overview of Strong Kids
    - Teaches social and emotional learning skills (see table)
    - Designed as a starting point/universal level of support
    - 12 lessons plus booster
2) Goal of study
    - Determine if program improves social functioning and SEL skills over course of "year"
    - Incorporate into classroom (generalization feature, poster)
3) Lesson overview
    - Lesson 6: Clear Thinking, part 1

Course of Study:

4) Fidelity checks
    - Will observe first two lessons, possibly 1 or 2 after that
5) Generalization feature
    - Pre-correction is main part I'm interested in
    - Is tracking praise too much?
    - Can track over the months, or can email me weekly (periodic check-ins at the least)
6) Consultation/Check-in
    - Touch base with me over the implementation, use me as a resource for questions

Data collection

7) Data collection (pre-test)
    - Should take roughly 30 mins, but may go over a bit
    - Assent form, SEARS-C, Coping Scale
    - small groups for Knowledge Test (discuss private place to use; use clipboards to keep answers private?)
    - Easiest way to let students know their code number?
        1. Use note cards? Tell them beforehand? Have numbers written on measure already?

Generalization feature:

1. Pre-correct students on the skills they are learning prior to certain tasks.
2. Provide social praise when students are "caught" using them.
    - Record number of times daily for *pre-correction* and *praise* and will email numbers to Jason on a weekly basis.

Mark a tally for each occurrence

	Mon	Tue	Wed	Thu	Fri
Pre-correct					
Praise					

Pre-correction:

Example: "Before we take this quiz, let's remember to *let go of stress*. Remember to take deep breaths and think of a relaxing place while I pass out the quiz."

"It's almost recess time. I want you all to think about how you feel while you're playing and be sure to use "OK" instead of "not OK" ways to deal with your emotions, just like we talked about with *Strong Kids*."

Non-example: "Before we go to PE, let's remember to be a strong kid." *(not tied to a specific skill, vague)*

Praise:

Example: (*Student is upset over a grade, but says "tomorrow is another day"*) "Johnny, that was a great job of using clear thinking and not using any of the thinking errors that we talked about in *Strong Kids*."

(*Students resolve a conflict while working together in class*). "I'm proud of you two for working things out. You used *compromise*, which we talked about this week."

Non-example: "Good job saying how you feel." (*not tied to a specific skill from Strong Kids)*

APPENDIX J:

FIDELITY SHEETS

Fidelity Check
Lesson 1: Emotional Strength Training

Lesson Components	Circle One
I. Introduction (to Strong Kids): *(2-5 minutes)*	
1. Introduces general idea of curriculum outlined in script	Yes No
2. Identifies how often class will use curriculum (once/week)	Yes No
II. Introduction to Topics Covered: *(2-5 minutes)*	
3. Points out at least 2 topics the curriculum will cover	Yes No
4. Uses Supplement 1.1 (overhead)	Yes No
III. Awareness or Disclaimer: *(1-2 minutes)*	
5. Discusses that students with more severe issues may need to talk with other people in private (teacher, counselor...etc)	Yes No
IV. Defining Behavior Expectations: *(2-5 minutes)*	
6. Points out that students may share feelings and stories on a voluntary basis	Yes No
7. Identifies a rule/expectation for group with an example	Yes No
8. Identifies a second rule/expectation with an example	Yes No
9. Identifies a third rule/expectation with an example	Yes No
10. Uses Supplement 1.2 (overhead)	Yes No
V. Closure: *(2-5 minutes)*	
11. Summarizes that they will be using Strong Kids for the next few months	Yes No
12. Reminds students of rules/expectations	Yes No
VI. Homework Handout: (2-5 minutes)	
13. Passes out homework (Supplement 1.3)	Yes No
Totals:	___/13___/13

General Notes/Observations (please note briefly the quality of implementation and any components that may have been "partially" implemented):

Fidelity Check
Lesson 2: Understanding Your Feelings 1

Lesson Components	Circle One	
VII. Review:		
1. Reviews at least 2 rules for group	Yes	No
VIII.Introduction:		
2. Identifies purpose of lesson is to better identify our feelings	Yes	No
IX. Name and Define Skills:		
3. Defines "emotion"	Yes	No
4. Defines "uncomfortable feelings"	Yes	No
5. Defines "comfortable feelings"	Yes	No
6. Uses Supplement 2.1 (overhead)	Yes	No
7. Asks at least 2 questions to facilitate discussion around comfortable vs. uncomfortable feelings	Yes	No
X. Feelings Identification:		
8. Discusses or identifies at least 4 emotions and identifies them as comfortable vs. uncomfortable	Yes	No
9. Passes out Supplement 2.2 (handout)	Yes	No
XI. How Do You Feel?:		
10. Generates list of at least 3 emotions and situations	Yes	No
11. Passes out Supplement 2.3 (handout)	Yes	No
12. Has at least 2 students share responses from Supplement 2.3	Yes	No
XII. Closure/Homework Handout:		
13. Reviews at least 2 ideas/topics from the lesson	Yes	No
14. Passes out homework (Supplement 2.4)	Yes	No
Totals:	____/14	___/14

General Notes/Observations

Fidelity Check
Lesson 3: Understanding Your Feelings 2

Lesson Components	Circle One
I. Review:	
1. Reviews at least 1 topic from last week's lessons (e.g., defines emotion, describes comfortable vs. uncomfortable feelings, points out that emotions are tied to certain situations)	Yes No
II. Introduction:	
2. Identifies purpose of lesson is to learn appropriate ways to express feelings.	Yes No
Activity A:	
3. Conveys that emotions communicate how people feel.	Yes No
4. Points out that there are different ways to show feelings.	Yes No
Activity B:	
5. Identifies at least 2 emotions and common actions associated with them (e.g., happy = smile more, feel good inside…etc).	Yes No
6. Points out that inappropriate expression can hurt yourself or others.	Yes No
7. Points out that appropriate expression can be respectful and/or safe (not hurt others).	Yes No
III. Positive and Negative Examples of Showing Feelings	
8. Places Supplement 3.1 (overhead) on overhead.	Yes No
9. Reads through at least 4 of the examples.	Yes No
10. Places Supplement 3.2 (overhead) on overhead.	Yes No
11. Has students identify an emotion and then appropriate and inappropriate ways to express that emotion using Supplement 3.2.	Yes No
IV. Practice Situations and Examples	
12. Passes out Supplement 3.3 (handout) to students.	Yes No
13. Places Supplement 3.4 (overhead) on overhead.	Yes No
14. Has students get into small-groups and work on activity.	Yes No
V. Closure/Homework Handout	
15. Reviews at least 2 ideas/topics from the lesson.	Yes No
16. Passes out Supplement 3.5 (homework).	Yes No
17. Goes over directions with students on Supplement 3.5	Yes No
Totals:	___/17 __/17
General Notes/Observations	

Fidelity Check
Lesson 4: Dealing with Anger

Lesson Components	Circle One	
I. Review:		
1. Reviews at least 1 topic from a previous lesson.	Yes	No
II. Introduction/Name and Define Anger and Aggression:		
2. Identifies purpose of lesson is to discuss anger and how to deal with anger appropriately.	Yes	No
3. Uses Supplement 4.1 as an overhead.	Yes	No
4. Reviews all 5 definitions on Supplement 4.1.	Yes	No
5. Conveys idea that aggression is an inappropriate/negative way to deal with or express anger.	Yes	No
6. Conveys idea that anger does not have to lead to aggression and that there are better ways to express/manage your anger.	Yes	No
III. Introduce the Anger Model and Definitions/Integrate and Illustrate Anger Model		
7. Uses Supplement 4.2 as an overhead.	Yes	No
8. Defines/discusses all 6 components of the Anger Model.	Yes	No
9. Uses Supplement 4.3 as an overhead.	Yes	No
10. Reads/discusses each component of the Anger Model from Supplement 4.3.	Yes	No
IV. Introduce Anger Control Skills		
11. Uses Supplement 4.4 as an overhead.	Yes	No
12. Discusses 4 ways of coping with anger from Supplement 4.4.	Yes	No
V. Application of Anger Control Skills		
13. Passes out Supplement 4.5 to students.	Yes	No
14. Reads the negative example from Supplement 4.5 and asks at least 1 question to generate some discussion about the example.	Yes	No
15. Reads the positive example from Supplement 4.5 and asks at least 1 question to generate some discussion about the example.	Yes	No
VI. Closure/Homework Handout		
16. Passes out Supplement 4.6.	Yes	No
17. Reads/explains directions for homework to students.	Yes	No
Totals:	___/17	___/17

General Notes/Observations (please note briefly the quality of implementation and any components that may have been "partially" implemented):

Fidelity Check
Lesson 7: Clear Thinking 2

Lesson Components	Circle One	
I. Review:		
1. Reviews at least 1 topic from a previous lesson.	Yes	No
2. Places Supplement 7.1 on overhead ("Common Thinking Errors") and reviews/reads the 6 Thinking Errors.	Yes	No
II. Introduction/Looking for Evidence/How to Reframe Negative Thoughts:		
3. Communicates purpose of lesson is to learn how to change thinking errors.	Yes	No
4. Places Supplement 7.2 on overhead ("Evidence For or Against").	Yes	No
5. Reads examples on Supp 7.2.	Yes	No
6. Places Supplement 7.3 on overhead ("Reframing Negative Thoughts").	Yes	No
7. Reads examples Supp 7.3.	Yes	No
8. Places Supplement 7.4 on overhead ("Changing Thinking Errors").	Yes	No
9. Students get out homework from Lesson 6 to use with Supplement 7.4.	Yes	No
10. Has at least two students provide a negative thought from their homework and works through Supp 7.4 to show students how to reframe a negative thought.	Yes	No
III. Closure		
11. Reviews at least 2 ideas or topics from lesson.	Yes	No
12. Places Supplement 7.5 on overhead ("Feelings Thermometer").	Yes	No
13. Communicates general idea to students that the thermometer can be used to gauge how intense or strong our negative thoughts are.	Yes	No
14. Passes out homework (Supplement 7.6).	Yes	No
Totals:	___/14	___/14

General Notes/Observations (please note briefly the quality of implementation and any components that may have been "partially" implemented):

Fidelity Check
Lesson 12: Finishing Up

Lesson Components	Circle One
I. Introduction and Define Skills:	
1. Communicates purpose of lesson is to review things they've learned in Strong Kids.	Yes No
2. Teacher mentions at least two things that they have learned (e.g., identifying feelings, setting goals, identifying thinking errors...etc).	Yes No
II. Review of Strong Kids	
3. Uses Supplement 12.1 as an overhead.	Yes No
4. Discusses/reviews differences between comfortable and uncomfortable feelings.	Yes No
5. Discusses/reviews at least 2 components of anger model (model includes trigger, interpretation, emotional reaction, decision, behavior, and consequence).	Yes No
6. Discusses/reviews at least 2 thinking errors (e.g., binocular vision, black-and-white thinking, dark glasses, fortune telling, making it personal, blame game).	Yes No
7. Discusses/reviews ABCDE model of positive thinking and/or discusses optimism vs. pessimism.	Yes No
8. Discusses/reviews problem-solving model (identify problem, brainstorm, choose a solution, make an agreement).	Yes No
9. Discusses/reviews at least 1 way they learned to deal with stress	Yes No
10. Discusses/reviews goals and goal setting by either defining the terms or have students define them.	Yes No
11. Generates a list of people the students can talk to for help (e.g., parent, neighbors, family members...etc).	Yes No
Totals:	___/11 ___/11

General Notes/Observations (please note briefly the quality of implementation and any components that may have been "partially" implemented):

APPENDIX K:

MEANS AND STANDARD DEVIATIONS

OF DEPENDENT MEASURES BETWEEN CLASSROOMS

Table K1

Pre-testing Means for each Dependent Measure between Teachers

	Knowledge Test		Coping Scale		SEARS-C		SSBS-2	
Condition	M	SD	M	SD	M	SD	M	SD
Treatment								
Teacher 1	10.89	2.15	40.38	8.74	123.18	15.68	38.74	13.09
Teacher 2	9.89	3.21	40.01	12.05	109.17	21.20	36.37	11.81
Wait-list								
Teacher 3	9.43	3.02	40.73	10.61	117.29	19.58	27.61	12.22
Teacher 4	10.83	3.40	36.78	7.61	102.70	17.62	26.71	11.66

Table K2

Post-testing Means for each Dependent Measure between Teachers

	Knowledge Test		Coping Scale		SEARS-C		SSBS-2	
Condition	M	SD	M	SD	M	SD	M	SD
Treatment								
Teacher 1	13.48	4.01	41.88	7.98	122.10	20.83	45.74	13.33
Teacher 2	13.04	2.10	42.64	10.14	119.49	24.64	39.97	10.69
Wait-list								
Teacher 3	10.63	2.99	39.51	10.61	110.30	23.37	34.25	11.05
Teacher 4	11.42	2.99	31.04	8.57	90.26	25.30	31.92	11.96

Table K3

Follow-up Means for each Dependent Measure between Teachers

	Knowledge Test		Coping Scale		SEARS-C		SSBS-2	
Condition	M	SD	M	SD	M	SD	M	SD
Treatment								
Teacher 1	13.84	3.88	41.34	12.15	124.12	26.31	44.44	14.53
Teacher 2	13.08	2.62	39.77	11.75	112.86	23.55	44.64	8.24
Wait-list								
Teacher 3	11.04	2.66	37.55	10.93	106.08	24.41	33.46	12.05
Teacher 4	11.48	3.66	29.48	11.96	92.28	24.53	28.16	12.32

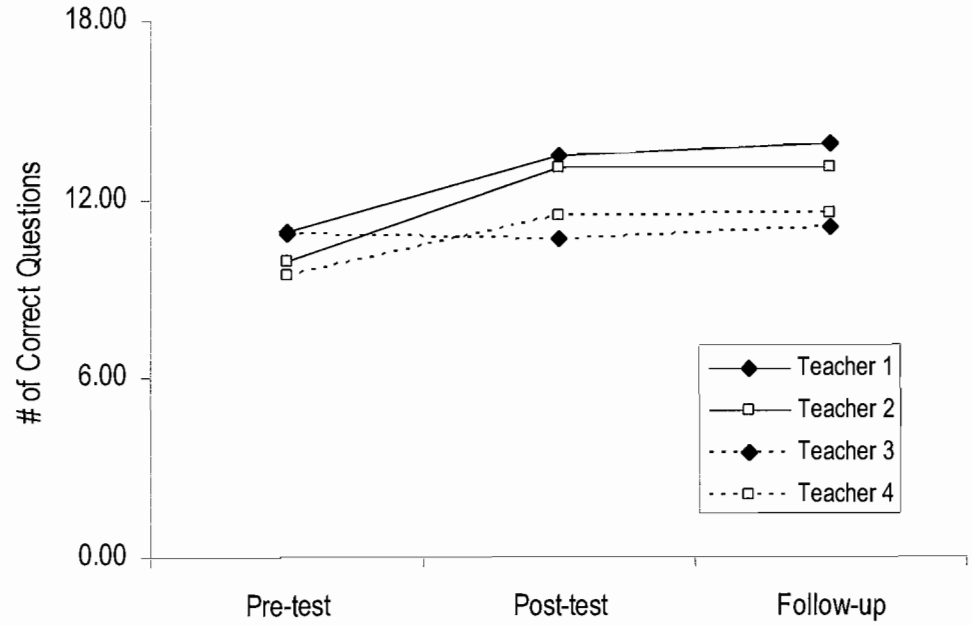


Figure K1. Visual display of means for *SK* Knowledge Test between classrooms and assessment period.

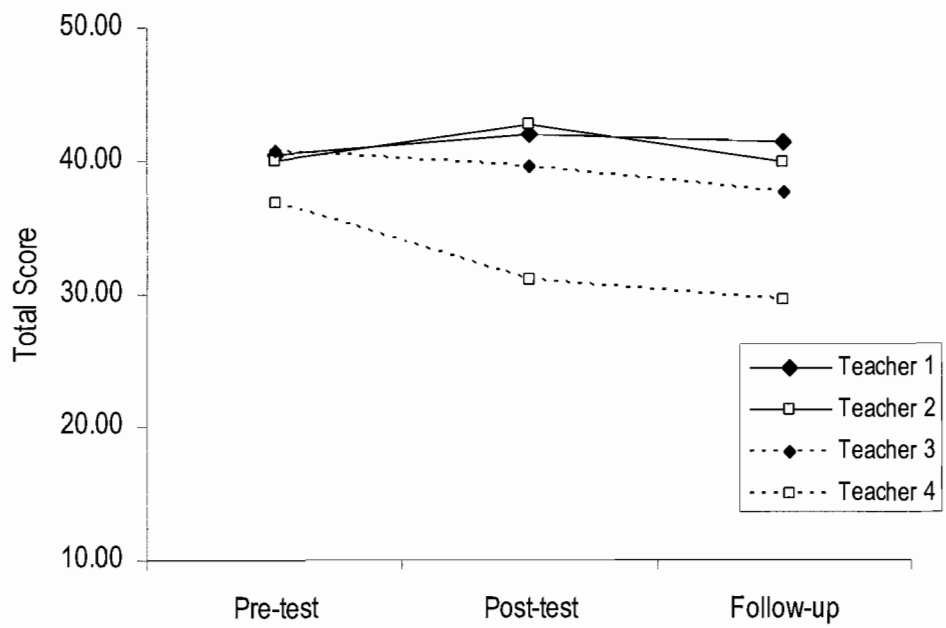


Figure K2. Visual display of means for the Coping Scale between classrooms and assessment period.

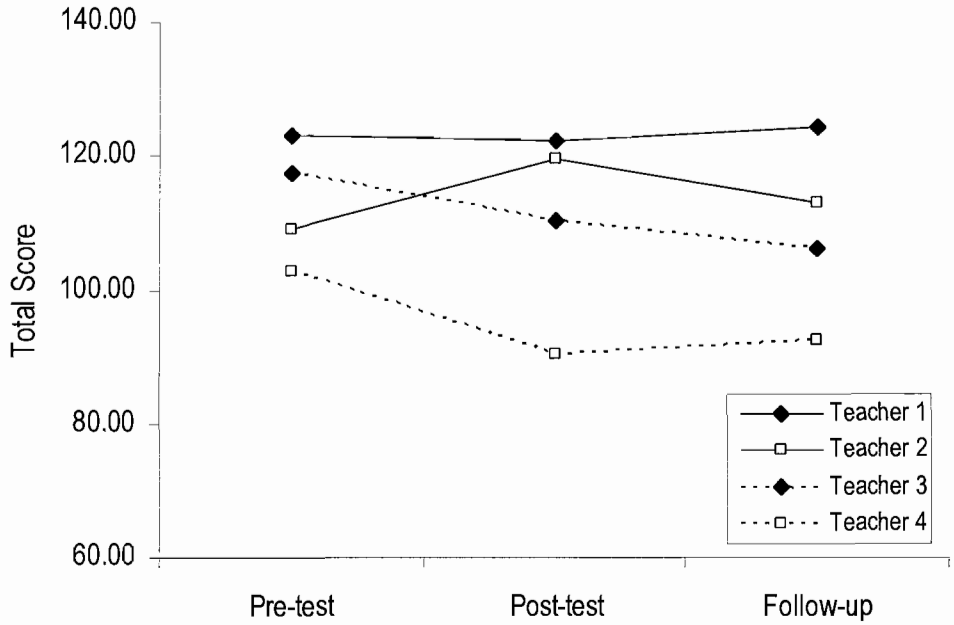


Figure K3. Visual display of means for the SEARS-C between classrooms and assessment period.

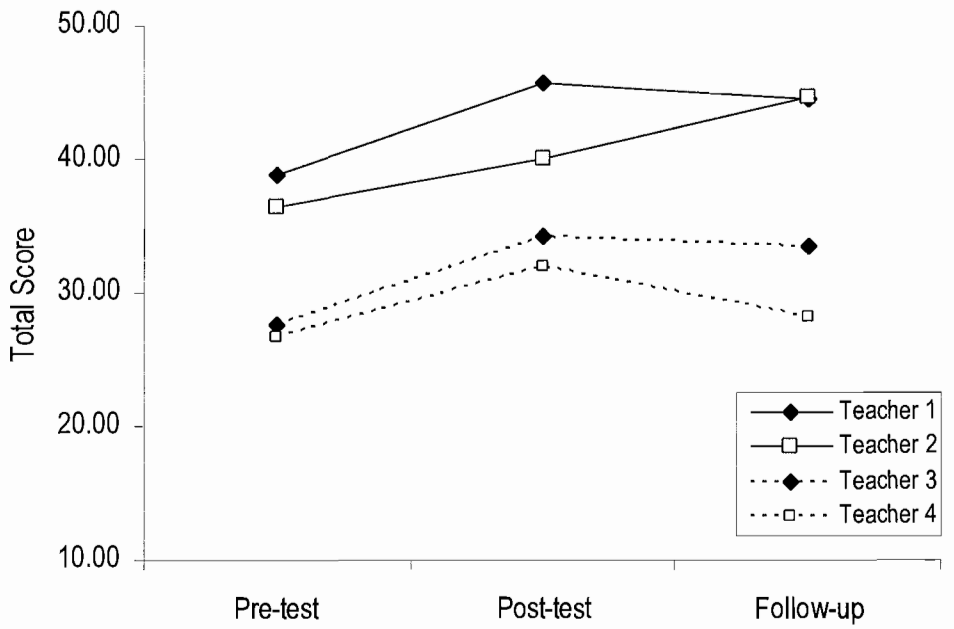


Figure K4. Visual display of means for the SSBS-2 between classrooms and assessment period.

APPENDIX L:

POST-HOC ANALYSES OF SEARS-C & SSBS-2

WITH GRADE AND CLASSROOM

A MANOVA was conducted with the constructed interaction variable of "teacher and grade" as a between-subjects factor and time as a within-subjects factor. Total scores on the SEARS-C and the SSBS-2 were entered as the DVs. A traditional alpha level of $p \leq .05$ was used as the significance level. The omnibus test was significant [$\Lambda = 0.55$, $F(16, 278.65) = 3.83$; $p \leq .01$], indicating that the groups differed over time. Pairwise comparisons were analyzed to determine the exact nature among the groups. Because these were post-hoc analyses and alpha slippage was not a concern, a traditional alpha level of $p \leq .05$ was used with no error correction.

SEARS-C Analyses

Changes between groups. Examining the SEARS-C at pre-testing, the only significant differences between the groups were that Teacher 4's third graders scored statistically significantly lower than Teacher 3's third graders and Teacher 1's fourth graders. Thus, it appears that the differences for the SEARS-C are a result of a classroom effect, as one group of third graders in one classroom scored lower than only two other classrooms (had it been a grade effect, all of the third graders would have scored below the fourth graders).

At post-testing, Teacher 4's third graders scored statistically significantly lower than both Teacher 1 and Teacher 2's classrooms. Teacher 4's fourth graders also scored statistically significantly lower than Teacher 1 and Teacher 2's classrooms. This is a favorable result, as Teacher 4 was within the control group and both Teacher 1 and Teacher 2 were in the treatment group. Thus, it was expected to see a difference between these classrooms at post-testing. However, the other control teacher's classroom (Teacher

2) did not score statistically significantly lower than the treatment group at either post-testing.

At follow-up, Teacher 4's third and fourth graders scored below only Teacher 1, so any differences at post-testing were not present between Teacher 4's classroom and Teacher 3's classroom.

Changes within groups. Examining Teacher 4's third graders, there were not statistically significant changes from pre to post, or from post to follow-up. For Teacher 4's fourth graders, there was a loss of scores from pre to post, as scores were statistically significantly lower at post-testing. The rest of the groups had not statistically significant changes in scores. These analyses indicate that the drop in scores or theorized loss of SEL skills for the wait-list group were only found within one group of fourth graders (Teacher 4's fourth graders).

Table L1

Descriptive Statistics between Grade/Teacher Groups and Assessment Periods for the SEARS-C

	Pre-testing		Post-testing		Post-testing	
Grade and Teacher	M	SD	M	SD	M	SD
Third						
Teacher 2	110.60	21.38	118.05	24.07	111.73	23.36
Teacher 3	117.29	19.58	111.87	22.26	107.64	23.13
Teacher 4	96.63	17.85	93.16	18.04	97.00	22.59
Fourth						
Teacher 1	122.04	15.68	121.61	21.60	123.80	26.80
Teacher 4	108.26	16.15	91.37	29.08	89.70	27.54

Note. Bolded teachers are treatment group.

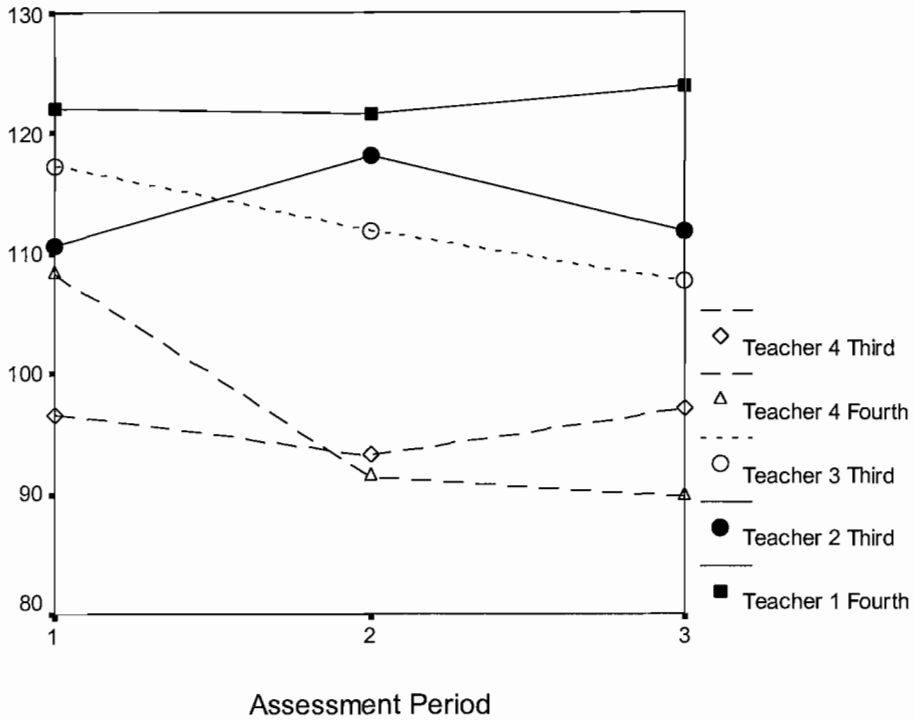


Figure L1. Visual display of means for SEARS-C between Grade and Classroom.

SSBS-2 Analyses

Changes between groups. Examining the scores for the SSBS-2 at pre-testing, Teacher 4's third graders scored statistically significantly lower than the treatment group (Teacher 1 and Teacher 2). Teacher 3's classroom also scored statistically significantly lower than Teacher 4's classroom as well. Coincidentally, differences at pre-testing were also aligned by treatment conditions.

At post-testing, Teacher 3's classroom and Teacher 4's third graders scored statistically significantly lower than Teacher 1's fourth graders. At follow-up testing, Teacher 3's third graders, Teacher 4's third graders and her fourth graders scored statistically significantly lower than Teacher 1 and Teacher 2's classrooms, respectively.

Changes within groups. Teacher 4's third graders and the fourth graders both increased scores over time, as scores on the SSBS-2 were statistically significantly higher from pre to post and from post to follow-up. Teacher 2's classroom followed a similar pattern, as their scores increased from pre to post and from post to follow-up.

Teacher 3's third graders increased statistically significantly from pre to post, but not from post to follow-up. This was also true for Teacher 1's classroom; they increased from pre to post, but not from post to follow-up.

Table L2

Descriptive Statistics between Grade/Teacher Groups and Assessment Periods for the SSBS-2

	Pre-testing		Post-testing		Follow-up	
Grade and Teacher	M	SD	M	M	SD	M
Third						
Teacher 2	37.42	11.59	40.26	10.81	45.00	8.21
Teacher 3	27.19	12.24	34.00	11.18	33.00	12.01
Teacher 4	21.46	8.51	28.91	3.53	27.00	3.60
Fourth						
Teacher 1	38.16	13.39	44.92	13.53	44.44	14.53
Teacher 4	31.50	12.87	36.33	10.68	31.75	11.79

Note. Bolded teachers are treatment group.

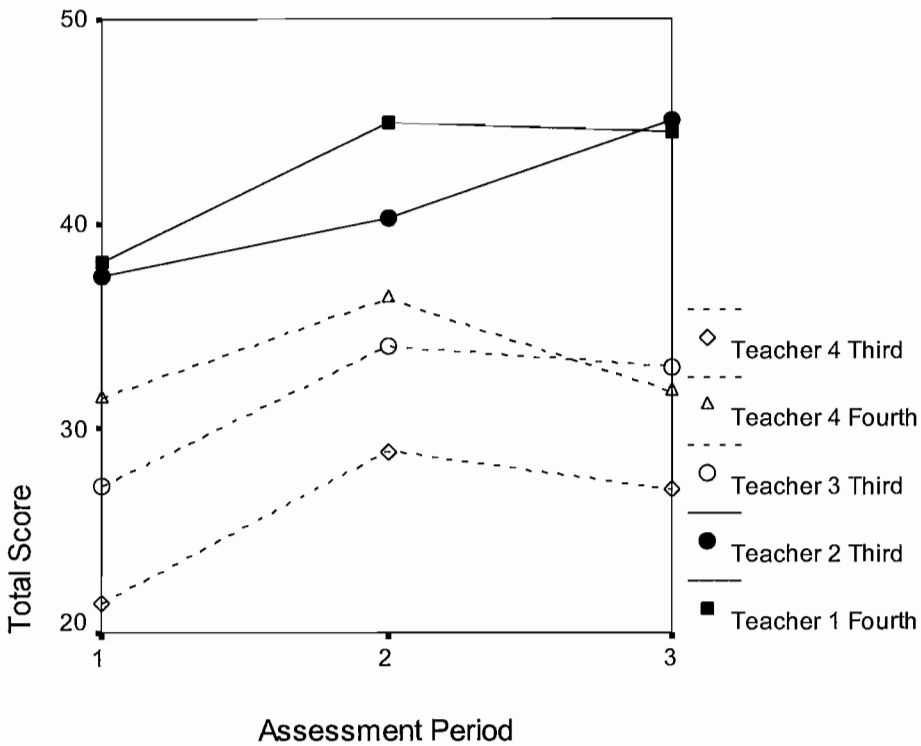


Figure L2. Visual display of means for SSBS-2 between Grade and Classroom.

APPENDIX M:

OFFICE DISCIPLINE REFERRAL INFORMATION

Below is a graph depicting the number of office disciplinary referrals (ODRs) from each classroom from September 2007 to February 2008. Teachers 1 and 2 were in the treatment group, and teachers 3 and 4 were in the wait-list. Teacher 1 had an average of 1.7 ODRs per month, teacher 2 had 1.2 per month, Teacher 3 had 1.3 per month, and teacher 4 had 3.2 per month.

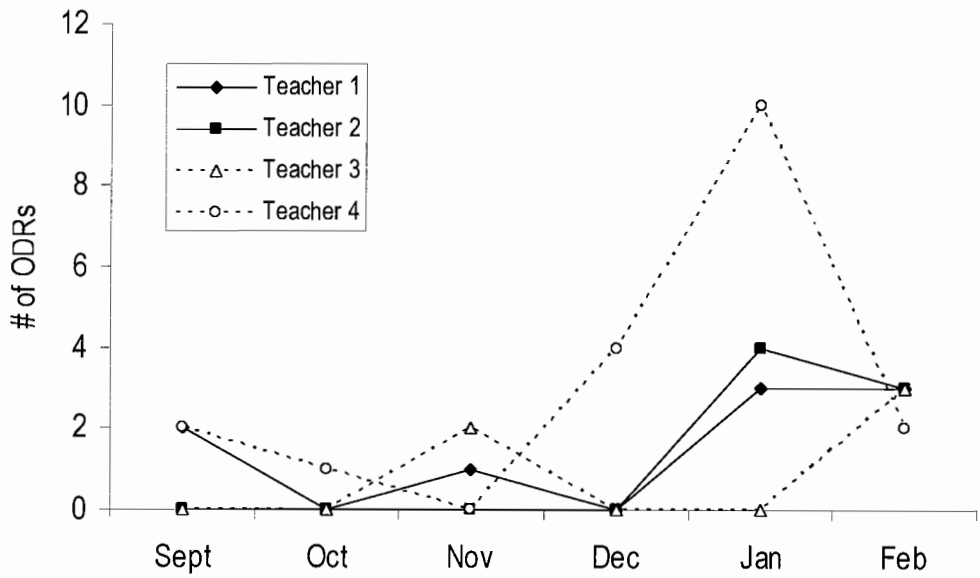


Figure M1. Number of office disciplinary referrals for each classroom.

BIBLIOGRAPHY

Abela, J. R. Z. (2001). The hopelessness theory of depression: A test of the diathesis–stress and causal Mediation components in third and seventh grade children. *Journal of Abnormal Child Psychology, 29*, 241-254.

Alberto, P. A., & Troutman, A. C. (2006). *Applied behavior analysis for teachers.* Upper Saddle River, NJ: Prentice Hall.

American Psychiatric Association, (2004). *Diagnostic and statistical manual of mental disorders- fourth edition.* Washinton, DC : APA.

Bauwens, J., & Houracde, J. J. (1992). School-based sources of stress among elementary and secondary at-risk students. *School Counselor, 40*, 97-102.

Beaver, B. R. (1997). The role of emotion in children's selection of strategies for coping with daily stresses. *Merrill-Palmer Quarterly, 43*, 129-146.

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching (3rd ed..,* pp 328-375). NY: MacMillan Publishing.

Brown, W. H., Odom, S. L., & Holcombe, A. (1996). Observational assessment of young children's social behavior with peers. *Early Childhood Research Quarterly, 11,* 19-40.

Caldarella, P., & Merrell, K. W. (1997). Common dimensions of social skills of children and adolescents: A taxonomy of positive behaviors. *School Psychology Review, 26*, 264-278.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundation of children's academic achievement. *Psychological Science, 11*, 302-306.

Casico, W. F., & Zedeck, S. (1983). Open a new window in rational research planning: Adjust alpha to maximize power. *Personnel Psychology, 36*, 517-526.

Castro-Olivo, S. M. (2006). *The effects of a culturally-adapted social-emotional learning curriculum on social-emotional and academic outcomes of Latino immigrant high school students.* Unpublished Doctoral Dissertation, University of Oregon, Eugene.

Catalano, R. F., Mazza, J. J., Harachi, T. W., Abbott, R. D., Haggerty, K. P., Fleming, C. B. (2003). Raising healthy children through enhancing social development in elementary school: Results after 1.5 years. *Journal of School Psychology, 41*, 143-164.

Causey, D. L., & Dubow, E. F. (1992). Developing a self-report coping measure for elementary school children. *Journal of Clinical Child Psychology, 21*, 47-59.

Chalk, K., & Bizo, L. A. (2004). Specific praise improves on-task behaviour and numeracy enjoyment: A study of year four pupils engaged in the numeracy hour. *Educational Psychology in Practice, 20*, 336-350.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J (2001). *Caring classrooms/intelligent schools: The social emotional education of young children*. New York: Teachers College Press.

Collaborative for Academic, Social, and Emotional Learning. (n.d.). www.casel.org. Viewed on January 15, 2006.

Colvin, G., Sugai, G., Good, R. H., & Lee, Y. (1997). Using active supervision and precorrection to improve transition behaviors in an elementary school. *School Psychology Quarterly, 12*, 344-363.

Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology, 67*, 648-657.

Cowen, E. L. (1994). The enhancement of psychological wellness: Challenges and opportunities. *American Journal of Community Psychology, 22*, 149-179.

Cummings, K. D., Kaminski, R. A., & Merrell, K. W. (in press). Advances in the assessment of social competence: Individual growth and development indicators. *Psychology in the Schools*.

Denham, S. A., & Weissberg, R. P. (2004). Social-emotional learning in early childhood: What we know and where to go from here. In E. Chesebrough, P. King, T. P. Gullotta, & M. Bloom (Eds.), *A blueprint for the promotion of prosocial behavior in early childhood* (pp. 13-50). New York: Klumer Academic/Plenum Publishers.

Doll, B., & Lyon, M. A. (1998). Risk and resilience: Implications for the delivery of educational and mental health services in school. *School Psychology Review, 27*, 348-363.

Doll, B., Zucker, S., & Brehm, K. (2004). *Resilient classrooms: Creating healthy environments for learning.* New York: The Guilford Press.

Dubow, E. F., & Tisak, J. (1989). The relation between stressful life events and adjustment in elementary school children: The role of social support and social problem-solving skills. *Child Development, 60*, 1412-1423.

Durlak, J. A., & Weissberg, R. P. (2005, August) *A Major Meta-analysis of Positive Youth Development Programs.* Presentation at the Annual Meeting of the American Psychological Association. Washington, DC

Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25*, 115-152.

Elias, M. J. (2004). Strategies to infuse social and emotional learning into academics. In J. Zins, R. Weissberg, M. Wang, & H. Walberg (Eds.), *Building academic success on social and emotional learning* (pp. 113-134). New York: Teachers College Press.

Elias, M. J., Zins, J. E., Graczyk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling up of social-emotional and academic innovations in public schools. *School Psychology Review, 32*, 303-319.

Elliott, S.N., Witt, J.C., Kratochwill, T.R., & Callahan-Stoiber, K. (2002). Selecting and evaluating classroom interventions. In M. Shinn, G. Stoner, & H.M. Walker (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 243-294). Bethesda, MD: National Association of School Psychologists

Erdfelder, E. (1984). Zur bedeutung und kontrolle des ß-fehlers bei der inferenzstatistischen prüfung log-linearer modelle. [On significance and control of the ß error in statistical tests of log-linear models]. *Zeitschrift für Sozialpsychologie*, 15, 18-32.

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer Program]. Bonn, FRG: Bonn University, Department of Psychology.

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ.* New York: Bantam.

Good, R.H., Gruba, J., & Kaminski, R. (2002). Best practice in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model. In A. Thomas & Grimes (Eds.), *Best practices in school psychology IV* (pp.679-699). Bethesda, MD: National Association of School Psychologists.

Greenberg, M. T., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention & Treatment, 4*, Article 1. Retrieved October 15, 2007 from http://journals.apa.org/prevention/volume4/pre0040001a.html.

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*, 466-474.

Gresham, F. (2002) Teaching social skills to high risk youth. In M. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 403-432). Bethesda, MD: National Association of School Psychologists.

Gueldner, B. A. (2006). *An investigation of the effectiveness of a social-emotional learning program with middle school students in a general education setting and the impact of consultation support using performance feedback.* Unpublished Doctoral Dissertation, University of Oregon, Eugene.

Hallahan, M. (1999). The Hazards of Mechanical Hypothesis Testing. Retrieved December 12, 2007 from http://psycprints.ecs.soton.ac.uk/archive/00000636.

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*, 60-87.

Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent-child interaction therapy.* New York and London: Plenum Press.

Hilsman, R., & Garber, J. (1995). A test of the cognitive-diathesis stress model of depression in children: Academic stressors, attributional style, perceived competence, and control. *Journal of Personality and Social Psychology, 69*, 370-380.

Horner, R.H., Sugai, G., Todd, A.W., & Lewis-Palmer, T. (2005). School-wide positive behavior support: An alternative approach to discipline in schools. In L. Bambara & L. Kern (Eds.). *Individualized supports for students with problem behavior: Designing positive behavior plans* (pp.359-390). New York: Guilford Press.

Howell, D. C. (2002). *Statistical methods for psychology- 5th edition.* Pacific Grove, CA: Duxbury, Thomson Learning.

Hoagwood, K., & Johnson, J. (2003). School psychology: A public health framework I. From evidence-based practices to evidence-based policies. *Journal of School Psychology, 41*, 3-21.

Hunter, L. (2003). School psychology: A public health framework III. Managing disruptive behavior in schools: The value of a public health and evidence-based perspective, *41,* 39-59.

Jamieson, J. (2004). Analysis of covariance (ANCOVA) with difference scores. *International Journal of Psychophysiology, 52*, 277-283.

Johnson, D. W., & Johnson, R. T. (2004). The three Cs of promoting social and emotional learning. In J. Zins, R. Weissberg, M. Wang, & H. Walberg (Eds.), *Building academic success on social and emotional learning* (pp. 40-58). New York: Teachers College Press.

Kam, C., Greenberg, M., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum, *Prevention Science, 4*, 55-63.

Kame'enui, E. J., Good, R., III, & Harn, B. A. (2005). Beginning reading failure and the quantification of risk: Reading behavior as the supreme index. In *Focus on behavior analysis in education: Achievements, challenges, and opportunities.* (pp. 69-89). Upper Saddle River, NJ: Prentice Hall.

Kazdin, A. E. (2004). Psychotherapy for children and adolescents. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 543-589). New York: Wiley.

Kennedy, C. H. (2005). *Single-case designs.* Boston: Allyn and Bacon.

Keppel, G., & Zedek, S. (1989). *Data analysis for research design*. New York: Freeman.

Killip, S., Mahfoud, Z., & Pearce, K. (2004). What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Annals of Family Medicine, 2,* 204-207.

Kilmer, R. P., Cowen, E. L., Wyman, P. A., Work, W. C., & Magnus, K. B. (1998). Differences in stressors experienced by urban African American, White, and Hispanic children. *Journal of Community Psychology, 26*, 415-428.

Kochenderfer-Ladd, B. J., & Skinner, K. (2002). Children's coping strategies: Moderators of the effects of peer victimization? *Developmental Psychology, 38*, 267-278.

Leff, S. S., Angelucci, J., Goldstein, A. B., Cardaciotto, L., Paskewich, B., & Grossman, M. B. (2007). Using a participatory action research model to create a school-based intervention program for relationally aggressive girls- The Friend to Friend program. In J. Zins, M. Elias, & C. Maher, (Eds.) *Bullying, victimization, and peer harassment: A handbook of prevention and intervention* (pp. 199-218). New York: The Haworth Press.

Mash, E. J., & Barkley, R. A. (1998). *Treatment of childhood disorders.* New York: The Guilford Press.

McCombs, B. L. (2004). The learner-centered psychological principles: A framework for balancing academic achievement and social-emotional learning outcomes. In J. Zins, R. Weissberg, M. Wang, & H. Walberg (Eds.), *Building academic success on social and emotional learning* (pp. 23-39). New York: Teachers College Press.

McConnell, S. R., McEvoy, M. A., & Priest, J. S. (2002). "Growing" measures for monitoring progress in early childhood education: A research and development process for individual growth and development indicators. *Assessment for Effective Instruction, 27*, 3-14.

Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children & adolescents.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Merrell, K. W. (2002). *School social behavior scales- 2nd edition.* Iowa City, IA: Assessment-Intervention Resources.

Merrell, K. W. (2007). *Social emotional assets and resiliency scale- child form.* Retrieved March 15, 2007 from http://strongkids.uoregon.edu/SEARS.html.

Merrell, K. W., & Buchanan, R. (2006). Intervention selection in school-based practice: Using public health models to enhance systems capacity of schools. *School Psychology Review, 35*, 167-180.

Merrell, K. W., Carrizales, D., Feuerborn, L., Gueldner, B. A., & Tran, O. K. (2007). Strong Kids: A Social and Emotional Learning Curriculum for Students in Grades 3-5. Baltimore: Brookes.

Merrell, K. W., Irvin, R. A., Gimpel, G. A. (2006). *School psychology for the 21st century.* New York: The Guilford Press.

Merrell, K. W., Juskelis, M. P., Tran, O. K., & Buchanan, R. (2006). *Social and emotional learning in the classroom: Impact of Strong Kids and Strong Teens on students' social-emotional knowledge and symptoms.* Manuscript submitted for publication.

Merrell, K. W., Levitt, V., & Gueldner, B. (in press). Proactive strategies for promoting social competence and resilience. In R. Ervin, G. Gimpel, E. Daly, and K. Merrell (Eds.), *The practical handbook of school psychology.* New York: The Guilford Press.

Miao, Y., Darch, C., & Rabren, K. (2002). Use of precorrection strategies to enhance reading performance of students with learning and behavior problems. *Journal of Instructional Psychology, 29*, 162-174.

Moroz, K. B., & Jones, K. M. (2002). The effects of positive peer reporting on children's social involvement. *School Psychology Review, 31*, 235-245.

National Center for Injury Prevention and Control (2004). http://webappa.cdc.gov/cgi-bin/broker.exe. Retrieved December 10, 2006.

No Child Left Behind. (n.d.). http://www.ed.gov/nclb/landing.jhtml.

Olweus, D. (2001). Peer harassment: A critical analysis and some important issues. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 3-23). New York: Guilford.

Payton, J. W., Wardlaw, D. M., Gracyzk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of School Health, 70*, 179-185.

Riggs, N. R., Greenberg, M. T., Kusche, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science, 7*, 91-102.

Romano, J. L. (1997). Stress and coping: A qualitative study of 4th and 5th graders. *Elementary School Guidance & Counseling, 31*, 273-282.

Salvia, J., & Ysseldyke, J. E. (2004). *Assessment in special and inclusive education.* Boston: Houghton Mifflin Company.

Seifer, R., Gouley, K., Miller, A. L., & Zakriski, A. (2004). Implementation of the PATHS curriculum in an urban elementary school. *Early Education & Development, 15*, 471-485.

Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Shinn, M.R. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp. 671-693). Bethesda, MD: National Association of School Psychologists.

Sprague, J. R. & Horner, R. H. (in press). Schoolwide Positive Behavior Supports. In S. R. Jimerson & M. J. Furlong (Eds.), *The Handbook of School Violence and School Safety: From Research to Practice.* Mahway, NJ: Lawrence Erlbaum Associates, Inc.

Stormont, M. A., Smith, S. C., & Lewis, T. J. (2007). Teacher implementation of precorrection and praise statements in Head Start classrooms as a component of a program-wide system of positive behavior support. *Journal of Behavioral Education, 16*, 280-290.

Strein, W., Hoagwood, K., Cohn, A. (2003). School psychology : A public health perspective I. Prevention, populations, and systems change. *Journal of School Psychology, 41*, 23-38.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245-249.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics.* Boston: Allyn and Bacon.

Taylor-Greene, S., Brown, D., Nelson, L., Longton, J., Gassman, T., Cohen, J, et al. (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education, 7*, 99-112.

Turner, H. A., Finkelhor, D., & Ormrod, R. (2006). The effect of lifetime victimization on the mental health of children and adolescents. *Social Science & Medicine, 62*, 13-27.

United States Census Bureau (2000). http://www.census.gov/population/www/cen2000/briefs.html. Retrieved December 9, 2006.

United States Department Health & Human Services: Administration on Children, Youth, and Families. (2006). *Child maltreatment 2004.* Washington, DC: Government Printing Office.

Walker, H. M., Todi, B., Block, A. & Severson, H. (1988). *Standardized screening and identification of behavior disordered pupils in the elementary age range: Peer social behavior observer trainer manual.* Eugene, OR: University of Oregon, Center on Human Development.

Weissberg, R. P., & O'Brien, M. U. (2004). What works in school-based social and emotional learning program for positive youth development. *The ANNALS of the American Academy of Political and Social Science, 591*, 86-97.

Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology, 17*, 247-272.

Wright, D. B. (2006). Comparing groups in a before-after design: When *t* test and ANCOVA produce different results. *British Journal of Educational Psychology, 76*, 663-675.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In J. Zins, R. Weissberg, M. Wang, & H. Walberg (Eds.), *Building academic success on social and emotional learning* (pp. 2-22). New York: Teachers College Press.

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building academic success on social and emotional learning.* New York: Teachers College Press.