School Psychology Review,
2012, Volume 41, No. 1, pp. 82–99

The Interaction Effects of Program Training, Dosage, and Implementation Quality on Targeted Student Outcomes for The RULER Approach to Social and Emotional Learning

Maria Regina Reyes, Marc A. Brackett, Susan E. Rivers,
Nicole A. Elbertson, and Peter Salovey
Yale University

Abstract. This study examined how training, dosage, and implementation quality of a social and emotional learning program, The RULER Approach, were related to students' social and emotional competencies. There were no main effects for any of the variables on student outcomes, but students had more positive outcomes when their teachers (a) attended more trainings and taught more lessons, and (b) were classified as either moderate- or high-quality program implementers. Student outcomes were more negative when their teachers were classified as low-quality implementers who also attended more trainings and taught more lessons. Post hoc analyses revealed that low-quality implementers felt less efficacious about their overall teaching than high-quality implementers. The discussion focuses on the importance of assessing the interaction of training and implementation variables when examining the effect of social and emotional learning programs.

School programs that aim either to prevent maladaptive behaviors (August, Bloomquist, Lee, Realmuto, & Hektner, 2006; Conduct Problems Research Group, 2011) or to promote positive development among youth (Domitrovich, Cortes, & Greenberg, 2007; Jones, Brown, & Aber, 2011) have been flourishing across the United States. These programs generally fall under the umbrella term, *social and emotional learning* (SEL), which refers to the process of acquiring the skills of self- and social awareness, emotion regulation, responsible decision making, problem solving, and relationship management (Zins, Weissberg, Wang, & Walberg, 2004). Accordingly, SEL programs are designed both to enhance these skills and create an emotionally supportive climate to increase the likelihood of school engagement, attendance, and academic success. The effects of these programs on youth outcomes have been positive (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011), but most evaluations did not include training or implementation data (Gottfredson & Gottfredson, 2002; Lewis, Battistich, & Schaps, 1990; Tanyu, 2007). The variables surrounding implementation need to be assessed both in research and in practice to better understand the effectiveness of programs in achieving their intended goals (Dane & Schneider, 1998;

Correspondence regarding this article should be addressed to Marc A. Brackett, Yale University, 2, Hillhouse Avenue, New Haven, CT 06511; e-mail: marc.brackett@yale.edu

Durlak & DuPre, 2008; Sanetti & Kratochwill, 2009). The relative importance of the quantity of teacher training, the dosage, or number of lessons students receive, and the quality of implementation, including teacher attitudes toward programming and their delivery style, are of particular interest in SEL programming.

In this study, we examined the extent to which these training and implementation variables for an SEL program, The RULER Approach (www.therulerapproach.org), were related to targeted social and emotional outcomes for students during the program's first year of implementation. We begin with a short overview of literature on program implementation followed by a description of SEL programs. We then highlight the important role that teachers play as implementers of SEL programs.

Implementing SEL Programs

Programs introduced into social settings like schools are not always implemented with fidelity (Fixsen, Blase, Naoom, & Wallace, 2009). This makes formative evaluations or the study of the processes underlying program implementation critical. Implementation occurs in six stages (Fixsen, Naoom, Blase, & Wallace, 2007). In the *exploration stage*, schools consider which program to adopt by examining feasibility and fit. In the *installation stage*, key stakeholders decide that the program will be implemented and plan for its proper execution. In the *initial implementation stage*, staff members are hired; participants are recruited; organizational supports are in place; and because all stakeholders are new to the program, problem solving and troubleshooting are frequent. In the *full implementation stage*, the program is fully integrated, with program processes and procedures part of the regular routine. Once the program has been implemented effectively, improvements are generally tested in the *innovation stage*. Sustaining the program both through continuous staff development and funding support comprise the *sustainability stage*.

Program implementation is rarely a perfect process, and a growing body of research shows that the effectiveness of school-based prevention programs is limited by the extent that they are implemented as intended (Dusenbury, Brannigan, Falco, & Hansen, 2003). Schools have wide-ranging priorities, policies, and politics that may interfere with how a program is delivered (e.g., Fagan & Mihalic, 2003; Gager & Elias, 1997). The evidence of SEL program effectiveness is growing; for example, a meta-analysis of over 200 studies shows that SEL programs have the intended positive effect on students' academic performance and their social and emotional skills (Durlak et al., 2011). Thus, schools will be implementing SEL programs in increasing numbers with varying levels of fidelity. Central to the understanding of how these programs are implemented is the role of teachers, who are the primary deliverers or "intervention drivers" (cf. Fixsen et al., 2009) of SEL programs.

Assessing Training and Implementation

Training

Training is the knowledge acquisition component of an SEL program and is the main avenue by which programs are introduced and implemented in schools. Training, which may include both workshops and coaching, is the vehicle by which teachers acquire background information, theory, and philosophy of the SEL program. Program information generally is introduced in initial trainings; then, follow-up coaching develops teacher's implementation skills more fully (Fixsen et al., 2009; Sanetti & Kratochwill, 2009; Strother, 1989). In their review, Joyce and Showers (2002) revealed that when training was combined with coaching, 95% of teachers acquired knowledge and developed skills for applying that knowledge in the classroom. In the absence of coaching, only 5% of teachers applied the skills in the classroom.

Dosage

Dosage refers to the number of lessons that teachers implement for students to receive in the classroom. There is some evidence that

higher doses of program instruction produce more optimal results in certain intervention contexts (e.g., Connell, Turner, & Mason, 1985). For example, the number of lessons taught significantly affected students' healthy eating in one intervention (Story et al., 2000) and students' perceptions of healthy sexual behaviors in another (James, Reddy, Ruiter, McCauley, & van den Borne, 2006). However, an investigation of one school-based alcohol abuse prevention program revealed that dosage (e.g., teacher reports of the number of class periods used to teach program materials) was not systematically related to reductions in drinking behavior (Hopkins, Mauss, Kearney, & Weisheit, 1988). Among SEL programs, where a primary goal is to improve students' social and emotional skills and engagement in learning, the number of SEL lessons delivered was related to slower growth in negative student outcomes (Aber, Jones, Brown, Chaudry, & Samples, 1998) and fewer unexcused absences (an indicator of engagement) among girls but not among boys (Moskowitz, Schaps, & Malvin, 1982). In sum, these findings suggest that higher dosage may lead to better outcomes.

Implementation Quality

Implementation quality refers to the manner in which a program is being executed (Dane & Schneider, 1998). As the deliverers of SEL programs, teachers' style of delivery is as important as the content (Jennings & Greenberg, 2009). Teachers' delivery styles and attitudes toward the program need to be congruent with the program. For example, SEL lessons often involve sharing personal experiences and being sensitive to students' needs. If teachers lack buy-in and motivation to engage with students openly, there may be dissonance between them and the SEL lesson. In this section, we discuss two components of quality that are critical to SEL programming, in particular: (a) delivery, which refers to quality of program execution or teaching effectiveness, and (b) attitudes, which refer to program buy-in or openness to programming.

Delivery style is vital to SEL programs because they require teachers to deliver the lessons in an effective manner, consistent with the program's philosophy and goals (see Fixsen et al., 2009; Waltz, Addis, Koerner, & Jacobson, 1993). For example, the teacher's display of certain emotions is important for many SEL lessons (Brackett et al., 2009; Elbertson, Brackett, & Weissberg, 2009). If a teacher cannot model the social and emotional skills a program is designed to target, that teacher will likely be less effective in imparting these skills to students. In general, teachers' beliefs about their teaching efficacy also influence their delivery of instructional programming (Han & Weiss, 2005).

Related to delivery style are teacher attitudes toward SEL programming, which also are critical to a program's success (see August et al., 2006). One study showed that within the context of a smoking prevention program, classrooms with teachers who had higher ratings on both positive attitudes (toward the program and their students) and preparedness had students with greater knowledge of and better decision-making skills about smoking (Botvin, Dusenbury, Baker, & James-Ortiz, 1989). Resistance to adopting SEL programs is common among teachers within the context of SEL. Some teachers are skeptical of the effect of SEL programs (Elias, Bruene-Butler, Blum, & Schuyler, 2000). They may be uncertain about the relative importance of SEL compared to other curricular efforts (Buchanan, Gueldner, Tran, & Merrell, 2009). Issues of accountability, such as those stemming from the No Child Left Behind Act (2001), also place tremendous pressure on teachers and schools to ensure their students perform well academically. As a result, teachers may be conflicted about the time they allocate for teaching core curricula versus SEL, both of which require dedication and constant practice.

Program quality in terms of delivery style alone is incomplete. It is unlikely that teachers will deliver SEL lessons with high quality if they are resistant to the program. To illustrate, teachers have varying levels of comfort with and commitment to incorporating

SEL lessons into academic curricula (Brackett, Reyes, Rivers, Elbertson, & Salovey, 2011), which play into how lessons are taught. Likewise, SEL programs are designed to create emotionally supportive climates for learning (Jennings & Greenberg, 2009); teachers with negative attitudes toward programming may undermine this program objective, rendering the program ineffective (Greenberg, Domitrovich, Graczyk, & Zins, 2005). Even if a teacher is implementing a program according to protocol, as judged by a trained observer, the attitude she or he has is integral to implementation quality.

Gaps in the SEL Literature: The Effect of Training and Implementation

The interaction of training and implementation variables with SEL program outcomes has yet to be studied extensively. For example, a teacher may receive a great amount of training and deliver the recommended number of lessons, but do so with a poor attitude or unsatisfactorily. Moreover, a teacher may be highly competent when delivering the program, yet do so infrequently (cf. Gresham, 2009; Waltz et al., 1993). Most SEL program evaluations have not adequately assessed the relative effect of each of these variables on student outcomes. Past research mostly *describes* how the programs were implemented (Kallestad & Olweus, 2003; Penuel, Fishman, Yamaguchi, & Gallagher, 2007; Ransford, Greenberg, Domitrovich, Small, & Jacobson, 2009; Stead, Stradling, Macneil, Mackintosh, & Minty, 2007; Story et al., 2000), yet few published studies report which variables predict program outcomes, as might be outlined in a theory-of-change model (Rossi, Freeman, & Lipsey, 1999). Moreover, although a few studies examined training and implementation variables simultaneously (for a review see Dusenbury et al., 2003), their interactive effect on outcomes was not analyzed. In one study, the number of program lessons taught and the quality of program delivery independently predicted more positive teacher and observer ratings of student outcomes, but interactive effects were not examined (Conduct Problems Research Group, 1999). The dearth of such studies makes it difficult to determine the critical ingredients of an intervention. For example, which affects student outcomes more: the amount of SEL program training a teacher receives, the number of SEL lessons he or she delivers, the quality with which those lessons are implemented, or some combination of the three?

Assessing Training and Implementation of SEL Programs

One challenge in assessing variables surrounding implementation is in their operationalization. In general, implementation quality is more difficult to operationalize than training or dosage, which can be quantified (Mowbray, Holter, Teague, & Bybee, 2003). To illustrate, training information can be obtained from attendance records or sign-up sheets at trainings, and dosage can be defined as teacher reports of lessons taught. Quality indicators, however, often are more difficult to obtain. Indeed, in a review of over 500 studies from 1976 to 2006 that assessed implementation of prevention and health promotion programs for children and adolescents, assessments of quality rarely were included. When quality was assessed, it was defined and measured in various, often unsystematic ways (Durlak & DuPre, 2008).

How should implementation quality be assessed? Having teachers rate the quality of their delivery of lessons introduces potential biases as teachers tend to overestimate their levels of implementation (Sanetti & Kratochwill, 2009), which often are higher than ratings by trained observers (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008). Similarly, when trained observers rate teacher quality (e.g., Kam, Greenberg, & Walls, 2003), they may lack thorough knowledge of both the program and the teachers to make accurate assessments. According to Waltz and colleagues (1993), raters of quality should be "sufficiently experienced and sophisticated to understand the implications of the contextual variables described in the [program] manual" (p. 628). Program coaches, who are trained as

experts in the program, may be the most knowledgeable judges of implementation quality because their interactions with teachers are more frequent and more personal (e.g., they have discussed with teachers their apprehensions and helped them to devise strategies to overcome them).

The Present Study

The present study extends previous research by examining associations and interaction effects of training, dosage, and implementation quality on intended student outcomes of social and emotional competence during the initial implementation phase (Fixsen et al., 2007), i.e., within the first year of adopting an SEL program. This study focuses on The RULER Approach (Brackett et al., 2011), which is grounded in a theoretical model that posits that acquiring the knowledge and skills associated with recognizing, understanding, labeling, expressing, and regulating emotion (i.e., the RULER skills) is critical to positive youth development (Brackett et al., 2009; Rivers & Brackett, 2011). RULER is an SEL program endorsed by the Collaborative for Academic, Social and Emotional Learning (www.casel.org), an organization comprised of distinguished educators and researchers that provides national leadership on SEL. The positive effects of RULER on both social and emotional competencies and classroom climate are reported elsewhere (Brackett, Rivers, Reyes, & Salovey, 2010; Rivers, Brackett, Reyes, Elbertson, & Salovey, 2011).

In the present investigation, we hypothesized that training, dosage, and implementation quality (i.e., delivery and attitudes), and their interaction, would relate positively to student social and emotional competencies. Training was assessed with attendance records at training sessions; dosage included number of program lessons delivered; and implementation quality was measured by observer (coaches') ratings of both teacher attitudes toward programming and their delivery of the program. Student outcomes were obtained from student self-reports, performance assessments, and report cards. Data were analyzed using a multilevel approach owing to their nested nature (Raudenbush & Bryk, 2002).

Method

Participants

Participants included sixth-grade students (n = 812) and their teachers (n = 28) from 28 elementary schools in a large, urban Catholic school district located in the northeastern United States. The schools were part of a randomized controlled trial (RCT) and the participating students and teachers were in schools assigned to use RULER (i.e., the program group). The full sample participating in the RCT consisted of 64 schools with 32 schools assigned randomly to the program group and 32 assigned randomly to the control group. (*Note:* Neither the individual participants nor the individual classrooms were assigned to groups. Schools were assigned randomly to either the program or control groups. Participating classrooms, teachers, and students were within these schools.) Four schools closed (two control and two program schools) during the course of the project. There were no differences in the demographic characteristics of the schools, teachers, or students between schools assigned to each group, except that the schools in the control group had larger enrollment numbers than those in the program group, $t(62) = 2.82$, $p = .006$. The current study focused exclusively on participants in the program group in the RCT for whom we had baseline data, which yielded 28 teachers and 812 students. We did not include participants in the control group.

On average, schools included 70% (SD = 33%) minority students (range = 5%–100%), and 24% (SD = 33%) of students received free or reduced-price lunch. Schools ranged in size from 178 to 656 students (M = 293.0, SD = 103.3) with a student–teacher ratio ranging from about 11:1 to 25:1 (M = 17.9, SD = 3.4). Participating schools varied in how they structured the school day for their sixth-grade students, such that at some schools, students received instruction from a single teacher for the entire day, and at others, students rotated through two or more

teachers throughout the day. The percentage of students in a school performing below average was based on the percentage of students with Levels 1 or 2 scores on the TerraNova Achievement Test (CTB/McGraw-Hill, 2002), which ranged from 8% to 86% ($M = 32.7\%$, $SD = 17.5\%$) in reading and from 0% to 67% ($M = 22.5\%$, $SD = 16.5\%$) in math.

Teachers were 84.4% female and identified themselves as 81.1% White/Caucasian, 9.1% Hispanic, and 9.1% Black/African American. These demographics resemble the racial and ethnic breakdown provided in 2010 U.S. census data: 72.4% White/Caucasian, 16.3% Hispanic, and 12.6% Black/African American (U.S. Census Bureau, 2011). Most of the teachers had either received their bachelor's degrees and/or were working toward a master's degrees (59.1%), and 31.8% had earned their master's degree or doctorates (9.1% missing these data). On average, teachers had been teaching for 13.1 year ($SD = 10.6$), with an average of 10.3 years ($SD = 9.4$) at their current school.

According to school records, students (48.6% female) were 27.0% White/Caucasian, 30.4% Black/African American, 22.0% Hispanic, 7.5% Asian/Pacific Islander, 3.7% multiracial, and 0.1% other race not mentioned (9.0% missing data). The composition of the student sample in this study was roughly similar to the racial and ethnic composition of the study's locale, although Caucasian students were underrepresented: 47.5% White/Caucasian, 28.4% Black/African American, 27.0% Hispanic, 11.1% Asian/Pacific Islander, and 4.9% multiracial (U.S. Census Bureau, n.d.).

Design and Procedure

RULER targets all students and is designed to be implemented throughout a school district. This study focuses on the training and implementation of RULER within the program group at the end of the first year of programming. This study is embedded into a large RCT in which program schools participated in training and used RULER for 2 years before schools in the control condition received the program.

The present study was divided into three waves of data collection: Wave 1 (March 2008) occurred prior to random assignment to condition and served as a baseline. Wave 2 occurred in the fall (September 2008) of the first programming year, as the program was being introduced; and Wave 3 occurred at the end of the first programming year (April 2009). Each wave of data collection lasted eight weeks. Students completed surveys and a performance test of emotion skills at each wave. Report cards were collected at Wave 3, the end of the first year of implementation, and contained data across all waves.

Curriculum Model and Implementation

RULER is grounded in research showing that a core set of emotion skills, recognizing, understanding, labeling, expressing, and regulating emotion, is essential to positive youth development (Brackett, Rivers et al., 2010; Salovey & Mayer, 1990). First, adult stakeholders (i.e., superintendents, school leaders, teachers, and staff) attend two full-day (6 hr per day) trainings on the role of emotion skills in school success, the theory underlying RULER, and on how to foster an emotionally supportive learning environment through the teaching and personal use of program Anchor tools, including the Charter (a collaborative mission statement for the learning environment) and the Mood Meter (a tool for plotting emotions and mood states), among other tools (Brackett, Caruso, & Patti, 2008; Brackett, Caruso, & Stern, 2008). Teachers then attend a second training, which is one full day focusing on the instruction of the Feeling Words Curriculum (Brackett et al., 2011), a literacy-based SEL program that provides teachers with programmatic units that infuse into and complement existing curriculum, including English language arts. The Feeling Words Curriculum helps children to develop emotion skills through an in-depth exploration of terms like *commitment*, *elation*, and *empathy*. These "feeling word units" are the vehicles by which children learn to identify, evaluate, and under-

stand their own and others' thoughts, feelings, and behavior, understand the emotions and points of view of characters in stories, and develop strategies to manage emotions in real-life situations. In the training, teachers learn how to use the curricular units in alignment with their English language arts teaching. Each unit, which focuses on one feeling word, is comprised of five 10- to 20-min lessons. Teachers teach one unit, with its five lessons, across a 2-week period. For instance, for the unit on alienation, three lessons may be completed during the first week and the remaining two the second week (see Brackett et al., 2011, for a review of the units).

The implementation process involves support through coaching. Each teacher works with a certified coach who visits the classroom, models lessons, reviews lesson plans, provides constructive feedback, and offers solutions and resources to help the teacher deliver quality lessons.

In September of the first year of implementation, English language arts teachers in program schools attended the first 2-day training on using emotional literacy and the Anchor tools to enhance the learning environment. Approximately 1 month later, teachers attended the second full-day training on the Feeling Words Curriculum. Of the two available trainings sessions offered, teachers attended an average of 1.87 sessions ($SD = 0.87$). Teachers in program schools then were paired with a certified RULER coach with whom they met for 45 min after a lesson was observed. Teachers received up to five coaching sessions, with an average of 4.02 sessions ($SD = 0.92$).

In this study, five female coaches each worked with teachers in up to eight schools. Coaches underwent intensive training with the developers of RULER programming before working in schools. A senior RULER trainer supervised all coaches throughout the duration of the project through regular meetings conducted in person and on the phone, as well as through routine reviews of all written documentation about the coaching sessions (e.g., observation checklists and notes). Each week, coaches submitted to the head coach the written documentation completed during and after each coaching session and classroom observation.

Teachers were asked to cover between 10 and 12 word units per year. Throughout the program year, teachers taught, on average, 7.20 word units ($SD = 2.60$, range 0–12 units), which yielded approximately 35 discrete emotional literacy lessons (i.e., 7 units $\times$ 5 lessons).

Measures

Training. Training was measured by the number of training and coaching sessions teachers attended, as obtained from training attendance records. The maximum training value was 7, including two trainings and five coaching sessions.

Dosage. Dosage was assessed by the number of lessons taught (lessons), as obtained from teacher reports, at the end of the first year of programming (Wave 3). The maximum number of lessons a teacher could teach was 60 (12 units with 5 lessons in each).

Implementation quality. To measure implementation quality, each of the five coaches rated (both at the beginning and end of the school year; i.e., Waves 2 and 3) the extent to which teachers (a) demonstrated buy-in or an open attitude toward the program (1 = *very resistant*, 5 = *very open*) and (b) delivered RULER lessons with high quality (1 = *needs a lot of improvement*, 5 = *excellent*). During each coaching session, coaches reviewed forms that teachers completed for each feeling word unit. At Wave 2, coaches had met with teachers for at least two of the five coaching sessions to assess quality delivery. By Wave 3, the remaining coaching sessions (up to three) were completed. The correlations between openness to programming and delivery at the beginning and end of the year (Waves 2 and 3) were r values (26) = 0.63, and 0.62, p values $< .001$, respectively.

Because the measure of implementation quality incorporated two items assessed across two time points, a parsimonious measure of

Table 1
Assessing Implementation Quality: Teacher Quality Clusters at the Beginning and End of the Year (Waves 2 and 3)

	Openness		Delivery	
Cluster	Wave 2	Wave 3	Wave 2	Wave 3
Low	1.79 (0.92)	3.33 (0.75)	1.17 (0.39)	2.58 (0.79)
Moderate	2.67 (0.82)	3.64 (0.70)	2.67 (0.49)	3.89 (0.58)
High	4.07 (0.80)	4.87 (0.23)	3.87 (0.74)	4.67 (0.49)

Notes. Based on the nature of cluster analysis, all clusters are significantly different from each other on all criterion variables.

quality was created by subjecting the indicators (i.e., openness and delivery) to cluster analysis to test whether distinct profiles of program quality existed. To select the optimal number of clusters, we first subjected the variables to an agglomerative hierarchical clustering procedure and then inspected the hierarchical tree diagram (Everitt, Landau, & Leese, 2001). A three-cluster solution proved to be optimal. The centroids from the hierarchical solution were entered as initial cluster centers in the final *k*-means iterative procedure. The three clusters that emerged were labeled: low-quality implementers (i.e., teachers who were initially very resistant to the program and delivered it poorly but became open to the program by the end of the school year; $n = 7$), moderate-quality implementers (i.e., teachers who were moderate in their attitudes toward the program and in their delivery of the program from beginning to end; $n = 12$), and high-quality implementers (i.e., teachers who were consistently open to and delivered the program very well from beginning to end; $n = 9$). There was no evidence to support a profile of teachers who were resistant to programming but high in delivery, nor was there evidence to support a profile of teachers who were open to programming but low in delivery. Table 1 summarizes the means and standard deviations for each cluster.

Social and emotional competence. Multiple methods were used to assess students' social and emotional competence. Table 2 summarizes the means, standard deviations, reliabilities, and intercorrelations among these variables at Wave 3.

First, students' report cards contained three items that reflected social competence (i.e., respects the rights of others, interacts appropriately, and complies with school policies) using a scale where 1 = *unsatisfactory*, 2 = *needs improvement*, 3 = *satisfactory*, 4 = *good*, and 5 = *excellent*. (Grades in these three areas were not necessarily given by the English language arts teachers [those who conducted the RULER lessons], depending on the structure of the students' school day and whether they were instructed by multiple teachers.) A composite score was created for the three items by adding the scores.

Social problem-solving skills were assessed with the Conflict Resolution Skill subscale of the Elementary Student Questionnaire of the Child Development Project (Developmental Studies Center, 2000). This eight-item scale presents students with four peer-conflict scenarios (two items per scenario). For each item, students selected one response from a multiple-choice list. Higher scores reflected the selection of more collaborative and compromise-centered responses to conflict, whereas lower scores reflected more aggressive or evasive responses to conflict. Students receiving a school-based program aimed at promoting their social, ethical, and intellectual

Table 2
Intercorrelations, Means, Standard Deviations, and Reliability Coefficients of Students' End-of-Year (Wave 3) Social and Emotional Competencies (N = 812)

	1	2	3
1. Emotional Literacy	–		
2. Social Problem Solving	.28	–	
3. Social Competence	.24	.32	–
M	105.52	2.71	4.08
SD	12.84	0.97	0.86
Range	56.86–127.26	1.00–4.50	1.00–5.00
Cronbach's α	.87	.79	.96

Note. All variables are significant at $p < .001$.

development had higher scores than a control group of students on this scale (Schaps, Battistich, & Solomon, 2004).

Emotional literacy was measured with the Strategic Emotional Intelligence component of the Mayer-Salovey-Caruso Emotional Intelligence Test—Youth Version (MSCEIT-YV; Mayer, Salovey, & Caruso, in press), which is appropriate for children between 11 and 17 years old. The test assesses the extent to which respondents understand emotional information and use that information for planning and self-management. Scores are calculated by combining two subtest scores: emotion understanding and emotion regulation. There are 23 multiple-choice items on the understanding subtest, which assesses the ability to identify both the definitions and causes of emotions. The regulation branch asks respondents to evaluate the effectiveness of several actions in making an individual feel a certain way. Respondents indicate the extent to which the chosen action would help the target character achieve a specified goal using a 5-point scale (1 = *not at all helpful*, 5 = *very helpful*). This section describes six situations, each of which has three alternatives, for a total of 18 items. Performance on the test is calculated by veridical scoring, which is described extensively in the technical manual (Mayer, Caruso, & Salovey, 2005). To explain briefly: emotion experts consulted the empirical literature to determine independently the best responses to each test item and then agreed on the best responses. Scores on the MSCEIT-YV are interpreted similarly to IQ scores with a mean of 100 and standard deviation of 15. Higher performance scores on understanding and regulation correlate positively with psychosocial functioning (Rivers, Brackett, & Salovey, 2008) and with standardized achievement test scores in reading (Peters, Kranzler, & Rossen, 2009).

Teaching efficacy. Teaching efficacy was assessed with the five-item Adaptive Efficacy Scale (Search Institute, 2006), which measures teachers' beliefs in their ability to modify their teaching methods, when needed, to have a positive effect on students. Teachers rated the extent to which they agreed or disagreed with each statement (e.g., "When a student has trouble learning something new, I try a new strategy"; "I am certain that I am making a positive difference in the lives of students") using a 5-point Likert scale (1 = *strongly disagree*; 5 = *strongly agree*). Cronbach's α values were .75 and.78 for beginning- and end-of-year teaching efficacy, respectively.

Analytic Strategy

The main and interaction effects of training, dosage, and implementation quality

on students' year-end social and emotional competencies were examined, controlling for student demographics and baseline scores.

Missing data. Of the 812 students, 173 had missing data, leaving 639 students with any data on the social problem-solving skills and social competence indicators. Missing data were treated with multiple imputation procedures in NORM (Schafer, 2000), which created five complete data files. Multilevel analyses were conducted for each of the five imputed data files and coefficients. Standard errors resulting from each analysis were averaged to provide estimates of the associations among our variables of interest (Schafer, 1999). Furthermore, return rates were lower for the MSCEIT-YV than the other assessments. Of 812 students, 425 had no MSCEIT data at either Waves 2 or 3, leaving only 387 students with MSCEIT data from Wave 1 and either Wave 2 or 3. Our imputations were based on data from these waves for these 387 students. The lower return rates for the MSCEIT probably could be attributed to the fact that teachers (and not the research team) administered this test. Separate imputations were conducted for emotional literacy scores because of low return rates. Comparable results were obtained from both complete and imputed data sets.

Primary analyses. Because of the nested design, we analyzed data using hierarchical linear modeling with full-information maximum-likelihood estimation with separate models for each student-level outcome. We nested students (Level 1) within teachers (Level 2) because we were interested in teachers' implementation of RULER. A three-level hierarchical model (students nested in teachers nested in schools) was unnecessary because there was a 1:1 correspondence between teachers and schools. To analyze the effect of training and implementation variables on our target outcomes, we ran two models: a main effects model and an interaction effects (Training × Dosage × Implementation Quality) model. The first model examined the direct relationships between training, dosage, and quality with student outcomes (Model 1). The second model tested interaction or moderation effects, crossing training, dosage, and implementation quality indicators (Model 2). To determine whether Model 2 contributed incrementally to the explanation of the outcome variable, we examined the change in R^2 by testing the change in χ^2 ($\Delta\chi^2$).

Finally, we calculated effect sizes using the formula:

$$\delta = \frac{\gamma}{\sqrt{\tau_{00} + \sigma^2}}$$

where γ is the association between the predictor and outcome variables, and the denominator is the *SD* of the outcome variable, where τ_{00} and σ^2 are the between- and within-groups variances, respectively, from the unconditional model. Interpretation of δ is similar to Cohen's (1988) *d*: 0.2 is small, 0.5 is moderate, and 0.8 is large.

Results

There were no main effects of training, dosage, or implementation quality on the student outcome variables at the end of the year, after controlling for baseline status (Model 1); however, numerous interaction effects were detected (Model 2), as Table 3 shows. Because quality indicators were coded as dummy variables, we chose the reference variable to be low-quality implementers. All analyses, therefore, are in comparison to this group. Moreover, all student outcomes pertain to year-end status (Wave 3) after controlling for baseline (Wave 1).

Among high-quality implementers, those who taught more feeling word units had students with higher scores on all three student outcomes: social competence ($t = 3.83$, effect size [ES] = 0.23), social problem solving ($t = 5.96$, ES = 0.19), and emotional literacy ($t = 5.47$, ES = 0.16). High-quality implementers who attended more training also had students who scored higher on the measures of social problem solving ($t = 2.58$, ES = 0.28), emotional literacy ($t = 1.82$, ES = 0.34), and social competence ($t = 1.78$, ES = 0.24);

Table 3
Training, Dosage, and Implementation Quality: Main and Interaction Effects on Year-End Student Outcomes (Wave 3)

	Students' Social and Emotional Competence Year-End Scores		
	Emotional Literacy (n = 387)	Social Problem-Solving Skills (n = 812)	Social Competence (n = 812)
Model 1: Main Effects			
ICC[a]%	8.62	11.25	35.86
Intercept	107.19 (2.06)***	2.81 (0.18)***	4.60 (0.19)***
Level 1 (Student)			
Black	0.42 (1.19)	−0.24 (0.14)	−0.11 (0.09)
Hispanic	0.36 (1.45)	−0.19 (0.13)	−0.12 (0.06)
Asian	3.27 (2.27)	0.04 (0.16)	−0.04 (0.09)
Other race	−3.21 (5.83)	0.17 (0.28)	0.32 (0.11)**
Male	−2.04 (1.06)	−0.08 (0.08)	−0.23 (0.07)**
Baseline score[b]	0.62 (0.05)***	0.52 (0.04)***	0.46 (0.06)***
Level 2 (Teachers)			
Training	−0.65 (1.28)	0.03 (0.06)	0.09 (0.07)
Dosage	−0.54 (0.48)	−0.01 (0.02)	0.01 (0.03)
Implementation Quality[c]			
Moderate	0.69 (2.50)	0.05 (0.18)	−0.36 (0.23)
High	0.78 (2.70)	0.14 (0.16)	−0.34 (0.20)
Model 2: Interaction Effects[d]			
Intercept	102.71 (1.62)***	2.69 (0.07)***	4.48 (0.10)***
Training × Low	−7.01 (1.47)***	−0.25 (0.07)**	−0.07 (0.11)
Training × Moderate	8.35 (2.50)**	0.18 (0.15)	−0.04 (0.25)
Training × High	4.24 (2.33)	0.27 (0.15)*	0.21 (0.12)
Dosage × Low	−1.37 (0.25)***	−0.13 (0.02)***	−0.18 (0.04)***
Dosage × Moderate	−0.27 (0.41)	0.11 (0.04)**	0.26 (0.05)***
Dosage × High	2.03 (0.37)***	0.19 (0.03)***	0.20 (0.05)**
Model 1 R^2	39.83	50.66	46.72
Model 2 R^2	87.00	86.01	69.75
$\Delta\chi^2(4)$	16.20**	14.93**	12.91*

Note. Estimated means (standard errors) reported.
[a] ICC = Intraclass correlation coefficient; [b] Baseline (Wave 1) score of corresponding outcome variable assessed; [c] Low is the reference group; [d] Truncated output.
* $p < .05$, ** $p < .01$, *** $p < .001$.

however, the latter two findings did not reach conventional levels of statistical significance ($p < .10$).

Among moderate-quality implementers, those who attended more training had students with higher emotional literacy scores (t = 3.34, ES = 0.68). Moderate-quality implementers who taught more feeling word units also had students with higher scores on both the social competence (t = 4.86, ES = 0.29) and social problem-solving (t = 3.11, ES = 0.12) assessments.

A different pattern was found for teachers classified as low-quality implementers. Teachers in this cluster who attended more training had students with lower scores on

both the social problem-solving assessment ($t = -3.47$, ES = 0.25) and emotional literacy test ($t = -4.78$, ES = 0.57). Moreover, low-quality implementers who taught more feeling word units had students with lower scores on all outcomes: social competence ($t = -4.65$, ES = 0.20), social problem solving ($t = -6.03$, ES = 0.13), and emotional literacy ($t = -5.46$, ES = 0.11).

To investigate possible explanations for the disparate findings among low-, moderate-, and high-quality implementers, we ran post hoc analyses to examine whether differences in teaching efficacy existed among teachers in each cluster. The means for low-, moderate-, and high-quality implementers in teaching efficacy at Wave 3 were as follows: 3.84 ($SD = 0.22$), 4.38 ($SD = 0.34$), and 4.49 ($SD = 0.54$), respectively. Differences among the teacher clusters were significant, $F(2, 20) = 4.13$, $p = .034$. Bonferroni-corrected post hoc analyses revealed low-quality implementers scored lower in teaching efficacy than high-quality implementers ($p = .037$).

In summary, there were no main effects of training, dosage, or implementation quality on student outcomes. However, several interaction effects emerged, such that student outcomes were affected by a combination of the number of trainings teachers attended and of lessons they taught and the quality with which these teachers implemented the program.

Discussion

Although SEL programs have positively affected key developmental outcomes among youth (Durlak et al., 2011), the majority of past investigations did not address the relative importance of training and implementation variables on targeted program outcomes. In this study, we examined whether the amount of training teachers received, the number of lessons students received, and the quality of delivery for one SEL program, RULER, were associated with students' social and emotional competencies. Similar to others' investigations (Hopkins et al., 1988; Kam et al., 2003), we found no main effects for our indicators of training and implementation on expected outcomes. However, we did find numerous significant interactions. Higher attendance at trainings and coaching sessions for moderate- and high-quality implementers, but not low-quality implementers, resulted in students with higher scores on indices of social problem-solving skills and emotional literacy. For moderate- and high-quality implementers but not for low-quality implementers, teaching more lessons also resulted in better student outcomes.

The unfavorable effects of more training among low-quality implementers may be partly explained by teaching efficacy. Post hoc analyses revealed that low-quality implementers were less efficacious about their general teaching practices than high-quality implementers. Low-quality implementers may not have been prepared to deliver SEL lessons without first becoming more confident in their general teaching practices (cf. Buchanan et al., 2009). These findings add to the growing research base on factors that may contribute to effective SEL programming (Collaborative for Academic, Social, and Emotional Learning, 2003; Gager & Elias, 1997; Lewis et al., 1990).

Analyzing training as the number of training and coaching sessions attended *and* dosage as the number of program lessons taught (i.e., feeling word units) was highly informative. For example, we found that among moderate- and high-quality implementers, but not low-quality implementers, the number of feeling word units taught had more significant and positive associations with student outcomes than the number of trainings attended, suggesting that active implementation may be more important than mere attendance at training sessions. Certainly, professional development is critical to learning the instructional strategies of RULER or any SEL program, but it may not be sufficient for affecting outcomes. What appeared to matter more was how training and coaching sessions were actualized in the classroom (i.e., through quality instruction). Assessing quality in terms of both attitudes and delivery, which have been associated positively in other investigations (Botvin et al., 1989), sheds light on how

teachers implement the program with varying levels of openness and skill.

Implications for Teacher Training and Professional Development

When new programs are introduced in schools during the installation and initial implementation stages, there usually exists a high degree of variability in terms of buy-in or openness to programming (Fixsen et al., 2007). Implementing SEL programs can be difficult for teachers who are balancing their time between meeting traditional academic requirements and the new demands of SEL programs. Indeed, asking teachers to integrate SEL into their already busy schedules can be physically, mentally, and emotionally taxing (Ransford et al., 2009). Our findings revealed that having teachers with low levels of openness (program buy-in) and delivery, but who either attended more trainings (including coaching sessions) or conducted more program lessons, resulted in lower levels of positive social and emotional outcomes among students. One strategy for addressing this may be for schools and SEL program providers to focus training efforts during initial implementation on teachers with an open attitude toward programming. Once these teachers have been trained and the program is moving toward full implementation, teachers who report high resistance to programming can begin their training, as concerted efforts are made by program providers and school administrators to increase their buy-in to the program.

There are various reasons that teachers may be resistant and lack buy-in to SEL programs. Effective programming approaches will acknowledge these attitudes, devote attention toward addressing them, and incorporate critical feedback from resistant teachers into program content and instructional strategies (Greenberg et al., 2005). Moreover, additional program-related information, support, and resources could be offered to target resistant teachers. For instance, these teachers could be provided with: (1) more empirical rationale for and real-life examples of the program's positive effect on students; (2) emphasis on the match between program goals and the schools' or districts' goals, values, policies, and philosophies; (3) additional instructional support from their principals or from program coaches to improve their program-specific or general teaching efficacy, if necessary; and (4) connections with teachers who have experienced success with the program, in particular those who were resistant at first themselves and whose attitudes toward programming were transformed. Until initially resistant teachers are more supportive of the program, they should be advised to conduct fewer lessons, with close monitoring and support from a coach.

Although RULER, like many SEL programs, is designed to integrate into existing school curricula, without quality training and ongoing support, its sustainability will likely be at risk (Fagan & Mihalic, 2003; Gager & Elias, 1997; Gottfredson & Gottfredson, 2002). In the past, many schools have applied the "train-and-hope" model (Stokes & Baer, 1977) to teacher professional development; some schools rely solely on the purchase of "kits" that require no additional training. Teaching SEL effectively requires ongoing training, coaching, and monitoring, each of which is critical to successful implementation (Fagan & Mihalic, 2003; Fixsen et al., 2009; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Lewis et al., 1990). Coaching, for instance, provides the opportunity to give teachers immediate feedback on all aspects of program delivery (Strother, 1989). Because many schools employ school psychologists, counselors, and social workers who often are asked to coordinate SEL initiatives or cofacilitate the teaching of SEL, our findings have many implications for these stakeholders who play a key consultative role to SEL program providers, school administrators, and teachers.

Strengths, Limitations, and Future Directions

A primary strength of this study was the multimethod assessment of constructs. Training and implementation variables were assessed with self-reports, attendance records,

and ratings from coaches. Student outcomes were assessed with self-ratings, teacher ratings, and a performance assessment tool. The differential interaction effects found between training, dosage, and implementation quality on student outcomes highlight the intricacies of identifying the key ingredients of effective SEL programming.

One area of future research is how to balance capitalizing on available, existing school data with the need to collect additional data. For instance, in the current study, the social competence items from the report card were selected because they were ratings with which teachers were already familiar and which could be gathered for all students across schools without missing data. However, we acknowledge that this measure is not ideal. For one, we do not know the factors that teachers used to assign scores to each student. In the case of preexisting implementation data that schools have on file, missing data often are an issue. The problem here is determining whether implementation data are missing systematically or at random. To illustrate: (1) are program noncompliers more likely to have missing data than program compliers, or (2) are program noncompliers just as likely to have missing data as compliers? How then can researchers obtain the most essential data available from teachers (or even schools) who may be resistant to programming, data collection, or both? Archival records such as attendance sheets, lesson plans, report cards, and classroom observations are important in order to obtain as much complete data as possible. The drawback with working with these types of archival data are that they usually are not standardized and likely are influenced by the perceptions and biases of the staff recording the information. Implementation data are particularly difficult to assess as implementation processes vary considerably. Different schools implement programs at different rates and in different ways. Future research could compare the use of various forms of archival data with that of more standardized assessments in order to identify best practices for collecting data related to implementation and related outcomes.

Another area ripe for investigation is the assessment of coaching quality and style. Although this study employed coaches' ratings of teacher implementation quality, it did not employ systematic assessments of the quality or style of each coach or the potential biases of their observational ratings, which are not unlikely, given they are invested in the positive outcomes of their efforts and have frequent personal interactions with the teachers they rated. Even though coaches received extensive training and were monitored closely, assessing their implementation of the coaching protocol and the objectiveness of their observational assessments is important for future research. The quality of coaching that a teacher receives could affect that teacher's attitudes and approaches to implementation. Similarly, the biases in the coach's observations could influence how the teacher is categorized with regard to implementation quality. Although an investigation of these phenomena was beyond the scope of the current project, it would be a valuable contribution to future implementation research.

Examining teacher learning outcomes achieved during training and coaching sessions also may be important for determining the key ingredients to effective interventions. It is likely that the quality of teacher trainings as well as the differential effect of the same training on individual teacher learning would influence student outcomes. Thus, future research and practice should include some measure of what skills and knowledge teachers gleaned from training and coaching.

The role of teachers' social and emotional competencies in the successful delivery of SEL lessons also was not studied, but offers another area for future investigation. It is likely that these competencies are associated with multiple facets of program implementation, including attitudes and delivery (Brackett et al., 2009; Durlak & DuPre, 2008). For example, once specific competencies are identified to be associated with high-quality implementation, the teaching of such competencies could be integrated into teacher training. Such competencies also may serve as moderators of implementation quality on

student outcomes, or as mediators such that an SEL program may shift the skill set of teachers, making them more effective in the classroom.

Finally, this study focused exclusively on participants assigned to the program group in the RCT; we did not include participants from the control group. Ideally, implementation is analyzed systematically in both program and control groups. For this particular program, the inclusion of a control group would facilitate the building of an evidence base for establishing the effectiveness of the RULER intervention. In general, the inclusion of a control group would allow for a more advanced understanding of the true effect of SEL training and its implementation on student outcomes (see Cordray, 2000). One way to account for this variation is to create implementation measures that capture the essential elements of both SEL programs and related, standard teaching practices, to administer them to both conditions, and then to use these data as potential moderating variables in analyses (O'Donnell & Lynch, 2008). This approach, however, would require careful monitoring of both the program and control conditions, which is an added research cost.

Conclusion

Teachers play an important role in SEL programming, as they are the intermediaries between students and the program. The adoption of SEL programs can be met with either enthusiasm or resistance among teachers. The components of SEL programming framework used in this study, which was composed of training, dosage, and implementation quality (attitudes and delivery), proved useful in evaluating the success of RULER, one of many promising SEL programs. Our findings suggest that mere delivery of SEL lessons is not sufficient for cultivating benefits for students. Lessons must be taught frequently and delivered with quality. Further research is warranted on the many facets of program implementation and their associations with the effectiveness of SEL programs.

References

Aber, J. L., Jones, S. M., Brown, J. L., Chaudry, N., & Samples, F. (1998). Resolving conflict creatively: Evaluating the developmental effects of a school-based violence prevention program in neighborhood and classroom context. *Development and Psychopathology, 10,* 187–213. doi:10.1017/s0954579498001576

August, G. J., Bloomquist, M. L., Lee, S. S., Realmuto, G. M., & Hektner, J. M. (2006). Can evidence-based prevention programs be sustained in community practice settings? The Early Risers' advanced-stage effectiveness trial. *Prevention Science, 7,* 151–165. doi: 10.1007/s11121–005-0024-z

Botvin, G. J., Dusenbury, L., Baker, E., & James-Ortiz, S. (1989). A skills training approach to smoking prevention among Hispanic youth. *Journal of Behavioral Medicine, 12,* 279–296. doi:10.1007/bf00844872

Brackett, M. A., Caruso, D. R., & Patti, J. (2008). *Emotional literacy for school leaders.* New Haven, CT: RULER Group.

Brackett, M. A., Caruso, D. R., & Stern, R. (2008). *Emotional literacy for educators.* New Haven, CT: RULER Group.

Brackett, M. A., Kremenitzer, J. P., Maurer, M., Carpenter, M., Rivers, S. E., & Elbertson, N. (2011). *Creating emotionally literate classrooms: An introduction to The RULER Approach to social and emotional learning.* Port Chester, NY: National Professional Resources.

Brackett, M. A., Patti, J., Stern, R., Rivers, S. E., Elbertson, N. A., Chisholm, C., et al. (2009). A sustainable, skill-based approach to building emotionally literate schools. In M. Hughes, H. L. Thompson, & J. B. Terrell (Eds.), *Handbook for developing emotional and social intelligence: Best practices, case studies, and strategies* (pp. 329–358). San Francisco, CA: Pfeiffer/John Wiley & Sons.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2011). Assessing teachers' beliefs about social and emotional learning. *Journal of Psychoeducational Assessment.* doi: 10.1177/0734282911424879

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2010). Enhancing academic performance and social and emotional competence with the RULER Feeling Words Curriculum. *Learning and Individual Differences.* doi:10.1016/j.lindif.2010.10.002

Buchanan, R., Gueldner, B. A., Tran, O. K., & Merrell, K. W. (2009). Social and emotional learning in classrooms: A survey of teachers' knowledge, perceptions, and practices. *Journal of Applied School Psychology, 25*(2), 187–203. doi:10.1080/15377900802487078

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

Collaborative for Academic, Social, & Emotional Learning. (2003). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning (SEL) programs.* Chicago, IL: Author.

Conduct Problems Research Group. (1999). Initial impact of the fast track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clin-*

ical Psychology, 67(5), 648–657. doi:10.1037/0022–006x.67.5.648

Conduct Problems Research Group. (2011). The effects of the Fast Track preventive intervention on the development of conduct disorder across childhood. *Child Development, 82*(1), 331–345. doi:10.1111/j. 1467–8624.2010.01558.x

Connell, D. B., Turner, R. R., & Mason, E. F. (1985). Summary of findings of the school health education evaluation: Health promotion effectiveness, implementation, and costs. *Journal of School Health, 55,* 316–321.

Cordray, D. S. (2000). Enhancing the scope of experimental inquiry in intervention studies. *Crime & Delinquency, 46,* 401–424.

CTB/McGraw-Hill. (2002). *TerraNova* (2nd ed.). Monterey, CA: Author.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18,* 23–45. doi:10.1016/s0272–7358(97)00043–3

Developmental Studies Center. (1988–2005). *Student questionnaire, Child Development Project (Elementary school version*). Oakland, CA: Developmental Studies Center.

Domitrovich, C., Cortes, R., & Greenberg, M. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention, 28*(2), 67–91.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41,* 327–350. doi:10.1007/s10464–008-9165–0

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82,* 405–432. doi:10.1111/j.1467–8624.2010.01564.x

Dusenbury, M. L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18,* 237–256. doi:10.1093/her/18.2.237

Elbertson, N. A., Brackett, M. A., & Weissberg, R. P. (2009). School-based social and emotional learning (SEL) programming: Current perspectives. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (Vol. 23, pp. 1017–1032). Dordrecht: Springer Netherlands.

Elias, M. J., Bruene-Butler, L., Blum, L., & Schuyler, T. (2000). Voices from the field: Identifying and overcoming roadblocks to carrying out programs in social and emotional learning/emotional intelligence. *Journal of Educational & Psychological Consultation, 11,* 253–272. doi:10.1207/s1532768xjepc1102_06

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis.* (4th ed.). London: Arnold.

Fagan, A. A., & Mihalic, S. (2003). Strategies for enhancing the adoption of school-based prevention programs: Lessons learned from the blueprints for violence prevention replications of the Life Skills Training program. *Journal of Community Psychology, 31,* 235–253. doi:10.1002/jcop.10045

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19,* 531–540. doi:10.1177/1049731509335549

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Wallace, F. (2007). Implementation: The missing link between research and practice. *APSAC Advisor, 19*(1 & 2), 4–11.

Gager, P. J., & Elias, M. J. (1997). Implementing prevention programs in high-risk environments: Application of the resiliency paradigm. *American Journal of Orthopsychiatry, 67,* 363–373. doi:10.1037/h0080239

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency, 39,* 3–35.

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). *The study of implementation in school-based preventive interventions: Theory, research, and practice (Volume 3).* Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Gresham, F. M. (2009). Evolution of the treatment integrity concept: Current status and future directions. *School Psychology Review, 38,* 533–540.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33,* 665–679.

Hopkins, R. H., Mauss, A. L., Kearney, K. A., & Weisheit, R. A. (1988). Comprehensive evaluation of a model alcohol education curriculum. *Journal of Studies on Alcohol, 49,* 38–50.

James, S., Reddy, P., Ruiter, R. A. C., McCauley, A., & van den Borne, B. (2006). The impact of an HIV and AIDS life skills program on secondary school students in Kwazulu-natal, south Africa. *AIDS Education and Prevention, 18,* 281–294. doi:10.1521/aeap.2006.18.4.281

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research, 79,* 491–525. doi:10.3102/0034654308325693

Jones, S. M., Brown, J. L., & Aber, J. L. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development, 82,* 533–554.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kallestad, J. H., & Olweus, D. (2003). Predicting teachers' and schools' implementation of the Olweus Bullying Prevention Program: A multilevel study. *Prevention & Treatment, 6,* article 21. doi:10.1037/1522–3736.6.1.621a

Kam, C.-M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science, 4,* 55–63. doi:10.1023/a:1021786811186

Lane, K. L., Kalberg, J. R., Bruhn, A. L., Mahoney, M. E., & Driscoll, S. A. (2008). Primary prevention programs at the elementary level: Issues of treatment integrity, systematic screening, and reinforcement. *Education & Treatment of Children, 31,* 465–494. doi:10.1353/etc.0.0033

Lewis, C., Battistich, V., & Schaps, E. (1990). School-based primary prevention: What is an effective program? *New Directions for Child and Adolescent Development, Winter, 50,* 35–59.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2005). *The Mayer-Salovey-Caruso Emotional Intelligence Test—Youth Version (MSCEIT-YV), Research Version.* Toronto, Canada: Multi Health Systems.

Mayer, J. D., Salovey, P., & Caruso, D. R. (in press). *Mayer-Salovey-Caruso Emotional Intelligence Test: Youth Version (MSCEIT: YV*): *Item booklet.* Toronto, Canada: Multi-Health Systems.

Moskowitz, J. M., Schaps, E., & Malvin, J. H. (1982). Process and outcome evaluation in primary prevention: The Magic Circle program. *Evaluation Review, 6,* 775–788. doi:10.1177/0193841x8200600604

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *The American Journal of Evaluation, 24,* 315–340. doi:10.1016/s1098–2140(03)00057–2

No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2001).

O'Donnell, C. L., & Lynch, S. J. (2008). *Fidelity of implementation to instructional strategies as a moderator of science curriculum effectiveness.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44,* 921–958.

Peters, C., Kranzler, J. H., & Rossen, E. (2009). Validity of the Mayer-Salovey-Caruso Emotional Intelligence Test: Youth Version—Research Edition. *Canadian Journal of School Psychology, 24*(1), 76–81. doi:10.1177/0829573508329822

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review, 38,* 510–532.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* (2nd ed.). Thousand Oaks, CA: Sage.

Rivers, S. E., & Brackett, M. A. (2011). Achieving standards in the English language arts (and more) using The RULER Approach to social and emotional learning. *Reading & Writing Quarterly, 27,* 75–100.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Elbertson, N., & Salovey, P. (2011). *Improving the social and emotional climate of classrooms with emotional literacy skill building: A clustered randomized controlled trial testing The RULER Approach.* Manuscript submitted for publication.

Rivers, S. E., Brackett, M. A., & Salovey, P. (2008). Measuring emotional intelligence as a mental ability in adults and children. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 2*, *Personality measurement and testing* (pp. 440–460). Thousand Oaks, CA: Sage.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185–211.

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review, 38,* 445–459.

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research, 8,* 3–15.

Schafer, J. L. (2000). *NORM Version 2.03.* Retrieved June 27, 2008, from http://www.stat.psu.edu/~jls/misoftwa.html

Schaps, E., Battistich, V., & Solomon, D. (2004). Community in school as key to student growth: Findings from the Child Development Project. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 189–205). New York: Teachers College Press.

Search Institute. (2006). *Creating a great place to learn survey: A survey of school climate. Technical manual.* Minneapolis: MN: Search Institute.

Stead, M., Stradling, R., Macneil, M., Mackintosh, A. M., & Minty, S. (2007). Implementation evaluation of the Blueprint multi-component drug prevention programme: Fidelity of school component delivery. *Drug & Alcohol Review, 26,* 653–664. doi:10.1080/09595230701613809

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349–367. doi:10.1901/jaba. 1977.10–349

Story, M., Mays, R. W., Bishop, D. B., Perry, C. L., Taylor, G., Smyth, M., et al. (2000). 5-a-Day Power Plus: Process evaluation of a multicomponent elementary school program to increase fruit and vegetable consumption. *Health Education & Behavior, 27,* 187–200.

Strother, D. B. (1989). Peer coaching for teachers: Opening classroom doors. *Phi Delta Kappa International, 70,* 824–827.

Tanyu, M. (2007). Implementation of prevention programs: Lessons for future research and practice: A commentary on social and emotional learning: Promoting the development of all students, a chapter by Joseph E. Zins and Maurice J. Elias. *Journal of Educational & Psychological Consultation, 17,* 257–262.

U.S. Census Bureau. (2011). *Overview of race and Hispanic origin: 2010.* Retrieved April 12, 2011, from http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf

U.S. Census Bureau. (n.d.). *Profile of general demographic characteristics 2000*, New York. Retrieved April 12, 2011, from http://factfinder.census.gov/

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620–630. doi:10.1037/0022–006x.61.4.620

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

Date Received: December 14, 2010
Date Accepted: October 09, 2011
Action Editor: Shannon Suldo ■

Maria Regina Reyes, PhD, is a postdoctoral associate at the Child Study Center in the School of Medicine at Yale University. Her training is in applied developmental psychology, which she studied at Fordham University where she earned her doctorate. She currently is applying sophisticated analytic techniques in the evaluation of social programs in early childhood education and examining the effects of these programs on children. She is involved in numerous policy projects on early child development with an international focus.

Marc A. Brackett, PhD, is a research scientist in the Department of Psychology at Yale University; deputy director of the Health, Emotion, and Behavior Laboratory; and head of the Emotional Intelligence Unit in the Edward Zigler Center in Child Development and Social Policy. He is the lead developer of The RULER Approach to Social and Emotional Learning, and the author of more than 80 scholarly publications, including seven social and emotional learning curricula. His grant-funded research focuses on examining the effect of emotion-based skills training on child and adult outcomes related to psychological health, social competence, and both academic and work performance. He works in school systems and corporations around the world in the areas of assessment and training.

Susan E. Rivers, PhD, is an associate research scientist in the Department of Psychology at Yale University, where she earned her doctorate. She also is associate director of the Health, Emotion, and Behavior Laboratory at Yale. She is a co-developer of The RULER Approach, as well as several other curricula designed to teach emotion skills. In her grant-funded research, she investigates how emotional skills training affects positive youth development and creates supportive learning environments. She is the co-author of many scholarly articles and papers, a consultant to businesses and schools, and works as an educational advisor on children's television programs, providing expertise on both resources to teach children emotional literacy and best practices for evaluating such programs.

Nicole A. Elbertson is a research associate in the Department of Psychology at Yale University, and the Manager of the Health, Emotion, and Behavior Laboratory. Her work focuses on the development and implementation of emotional literacy programs and the dissemination of information and research related to social and emotional learning. She is the co-author of many scholarly articles, papers, and book chapters on the development, application, and evaluation of The RULER Approach.

Peter Salovey, PhD, Provost of Yale University, is the Chris Argyris Professor of Psychology. He joined the Yale faculty in 1986 after receiving an AB and AM from Stanford University and a PhD from Yale. He has authored or edited 13 books translated into 11 languages and has published more than 350 journal articles and essays, focused primarily on human emotion and health behavior. With John D. Mayer, he developed the broad framework called emotional intelligence and for decades has studied the profound effect that measurable emotional skills have on thinking and action. In his research on health behavior, he investigates the effectiveness of health promotion messages in persuading people to change risky behaviors and adapt healthy ones.