Prev Sci (2013) 14:77–87
DOI 10.1007/s11121-012-0305-2

# Improving the Social and Emotional Climate of Classrooms: A Clustered Randomized Controlled Trial Testing the RULER Approach

**Susan E. Rivers · Marc A. Brackett · Maria R. Reyes · Nicole A. Elbertson · Peter Salovey**

Published online: 28 November 2012


**Abstract** The RULER Approach ("RULER") is a setting-level, social and emotional learning program that is grounded in theory and evidence. RULER is designed to modify the quality of classroom social interactions so that the climate becomes more supportive, empowering, and engaging. This is accomplished by integrating skill-building lessons and tools so that teachers and students develop their emotional literacy. In a clustered randomized control trial, we tested the hypothesis that RULER improves the social and emotional climate of classrooms. Depending upon condition assignment, 62 schools either integrated RULER into fifth- and sixth-grade English language arts (ELA) classrooms or served as comparison schools, using their standard ELA curriculum only. Multi-level modeling analyses showed that compared to classrooms in comparison schools, classrooms in RULER schools were rated as having higher degrees of warmth and connectedness between teachers and students, more autonomy and leadership among students, and teachers who focused more on students' interests and motivations. These findings suggest that RULER enhances classrooms in ways that can promote positive youth development.

This research was supported by a grant from the William T. Grant Foundation (Grant #8364).

S. E. Rivers (✉) · M. A. Brackett · N. A. Elbertson · P. Salovey
Department of Psychology, Yale University,
2 Hillhouse Avenue, PO Box 208205,
New Haven, CT 06520-8205, USA
e-mail: susan.rivers@yale.edu

M. R. Reyes
Yale Child Study Center, School of Medicine, Yale University,
230 S. Frontage Road,
New Haven, CT 06520-7900, USA

**Keywords** Emotional literacy · Social and emotional learning · Positive youth development · The RULER Approach

The field of positive youth development holds that the needs of youth must be addressed by creating environments or settings that provide mutually supportive relationships with adults and peers and that encourage school achievement, problem solving, and civic engagement (Catalano et al. 2004). Efforts to promote positive youth development differ from those aimed at reducing risk factors in that the former are focused on enhancing skills, building assets, and promoting resilience to achieve positive outcomes (Catalano et al. 2002; Greenberg et al. 2003). Accordingly, intervention efforts typically utilize a skill-building, whole-child approach that is focused on cultivating assets as opposed to preventing problems. As the predominant settings that youth inhabit, schools are compelling intervention targets.

This article presents data from the first year of a clustered randomized control trial (RCT) testing the impact of The RULER Approach ("RULER") on creating classroom settings that promote positive youth development (www.therulerapproach.org). RULER is a social and emotional learning (SEL) program that is applied universally (i.e., with all teachers and all students within a school; Institute of Medicine 1994), and includes professional development for school leaders, teachers, and staff, as well as classroom instruction protocols to enhance skill-building opportunities and characteristics of the learning environment (Brackett et al. 2009; Maurer and Brackett 2004). The theory of change underlying RULER specifies that by changing the setting of the classroom, RULER will improve youth outcomes, including academic engagement and achievement, as well as relationships with peers and adults.

## Changing Settings Through Social and Emotional Learning Programs

Ecological systems theory posits that the settings youth inhabit shape their development (Bronfenbrenner 1979). Environments that ignore the social and emotional needs of youth thwart development (Eccles et al. 1993). According to self-determination theory, to develop optimally, youth need to feel connected to others, competent in their abilities, and that their behavior is self-directed (Deci and Ryan 1985). Features of school settings related to meeting children's social and emotional needs include interactions between and among students and teachers that are empowering, caring and supportive, and safe and orderly (Catalano et al. 2004). When in classrooms that foster these kinds of interactions, youth are more likely to meet their social and emotional needs and, in turn, flourish; e.g., achieve academically, have meaningful friendships (e.g., Battistich et al. 1997; National Research Council/Institute of Medicine [NRC/IOM] 2002).

Best practices for creating positive youth-serving settings in schools are in their infancy, but some evidence-based approaches are emerging as potentially significant change agents (Shinn and Yoshikawa 2008). One such approach is the systematic integration of SEL programming, an educational strategy designed to foster the acquisition of social and emotional skills, such as self and social awareness, emotion regulation, responsible decision making and problem solving, and relationship management (Greenberg et al. 2003). SEL programs vary in their approaches; e.g., some focus on emotion regulation while others focus on coping or social support (Collaborative for Academic, Social, and Emotional Learning [CASEL] 2003).

Two categories of change agents for promoting SEL are those that provide skill-building opportunities for individuals within a setting and those that modify features of the environment itself (Catalano et al. 2004; CASEL 2003). A meta-analysis of 213 studies shows that students in schools using universally-applied SEL programs (compared to those in schools that do not) have improved social, emotional, and academic skills, including an 11-percentile point increase in achievement, as well as more prosocial behavior and positive attitudes toward the self and others, and lower levels of emotional distress (Durlak et al. 2011). The present study demonstrates the impact of an SEL program, RULER, on changing the environment of fifth- and sixth-grade classrooms.

## The RULER Approach ("RULER")

With the goal of promoting positive youth development, RULER combines comprehensive professional development with a literacy-based, skill-building SEL program for students. RULER builds social and emotional skills by focusing on the teaching and learning of emotion-related concepts or "feeling words" and by introducing tools for leveraging emotions in the learning environment. Educators and students analyze the emotional aspects of personal experiences, academic materials, and current events; evaluate how various people, characters, and historical figures feel and manage their feelings; and discuss techniques and use tools for identifying, problem solving about, and regulating their own and others' emotions.

Emotions are dynamic processes ever present in the daily lives of students and their educators and families (Schutz and Pekrun 2007). Among the many processes and behaviors that affect and are affected by emotions are social interactions, learning, self-reflective thinking, and perspective taking. The emotion-related skills individuals have also affect these activities (e.g., Mayer et al. 2008). Accordingly, RULER includes skill-building lessons and activities for *recognizing* emotions in oneself and others, *understanding* the causes and consequences of emotions, *labeling* emotions with an accurate and diverse vocabulary, and *expressing* and *regulating* emotions in socially appropriate ways (i.e., the RULER skills; Rivers and Brackett 2011). The delineation of these skills derives from the ability model of emotional intelligence (Mayer and Salovey 1997; Salovey and Mayer 1990), and accumulating evidence that shows the skills are related to social competence, psychological well-being, and academic performance (e.g., Fine et al. 2003; Rivers et al. 2012).

RULER is a multiyear program available for kindergarten through grade eight, with units that extend across the academic year. Students learn about a wide range of emotions and practice building their skills for leveraging and managing emotions across myriad learning tasks within the context of a standard academic curriculum. Professional development focuses on how to teach the lessons that comprise the student curriculum and on developing the skills and knowledge of the adults who create learning environments. The design of RULER is based on the achievement model of emotional literacy (Rivers and Brackett 2011). Emotional literacy refers to having a mastery of the RULER skills as well as an appreciation for the significance of both emotions and RULER skills in social interactions, personal growth, and learning. The achievement model proposes that emotional literacy is acquired through experience and develops through: the acquisition of emotion-related knowledge and skills; being in environments that are safe and supportive for experiencing a wide range of emotions; consistent opportunities to practice using the RULER skills with feedback on their application so that their use becomes refined and automatic; and frequent exposure to adults who model the RULER skills.

Evaluation research shows support for distal outcomes of RULER. Students in classrooms that integrated RULER had greater academic and social achievements compared to students in comparison classrooms (Brackett et al. 2010). For example, students in fifth- and sixth-grade classrooms using

RULER had better performance in writing and better work habits (e.g., followed directions) than students in comparison classrooms. Research has not yet examined the proximal outcomes of RULER (i.e., its impact on the emotional climate of the classroom) nor has an RCT on RULER been conducted.

## The Present Study

This RCT, conducted in fifth- and sixth-grade classrooms in 62 schools, tests the hypothesis that RULER improves classroom emotional climate (CEC). Guided by theories and empirical evidence on positive youth development, we define CEC as the social processes or interactions between and among teachers and students that reflect warmth and support, consideration of student perspectives in learning, teacher awareness of individual students' abilities, and teachers helping to empower students in their own learning (Hamre and Pianta 2007; Pianta et al. 2002). CEC is critical to both student academic performance and behavior (Brackett et al. 2011; Reyes et al. 2012). Multiple informants, including independent observers, teachers, and students, were used to assess CEC, an approach that may improve accuracy (cf. Moos 1979).

## Method

### Participants

The sample included teachers and students in the Catholic Schools of Brooklyn and Queens, NY. The principal investigator (PI), in collaboration with the Superintendent's Office, presented the project to principals during a regularly scheduled district meeting. Seventy principals (of 109) attended, and 66 volunteered their fifth- and sixth-grade English language arts (ELA) classrooms to participate. After randomization, two program and two comparison schools dropped out, leaving 62 schools, 155 classrooms, 105 teachers (70.5 % female, 20 % missing data), and 3,824 students (42.4 % boys, 15.4 % missing data). School size ranged from 178 to 656 students ($M$=325.92, $SD$=97.06), with a student-to-teacher ratio ranging from 10.96–25.08 to 1 ($M$=24.47, $SD$=3.74). Between 5.05 % and 100 % of students were minorities ($M$= 66.85 %, $SD$=32.30 %), and between 0 % and 94.67 % received free/reduced lunch ($M$=23.34 %, $SD$=32.30 %). At baseline, 54.83 % of schools reported having a program to address social or emotional skills, which was either Olweus Bullying Prevention (Olweus 1997) or Learning for Life (www.learning-for-life.org). On average, teachers had taught for 14.4 years and had worked at their current school for 10.7 years. In terms of education level, 40.0 % of teachers had worked toward or completed a master's degree, 36.2 % earned BAs, and 23.8 % did not specify.

### Procedure

*Timeline* In January 2008, principals heard a presentation by the PI that included a description of the project and the commitment required for participation. Principals who agreed to participate returned a signed agreement form. In March 2008, baseline data collection commenced (Wave 1). Schools were assigned randomly to a condition in July 2008. Training for program schools commenced in October 2008, after which teachers began to use the intervention throughout the academic year. Waves 2 and 3 commenced in both conditions in October 2008 and April 2009, respectively. Each wave of data collection lasted 8 weeks.

*Randomization* To ensure an equal number of schools were represented in each condition, we first divided the 62 schools randomly into pairs and then used a random number generator to assign either a 0 or 1 to each pair. If the number generated was 0, then the first school of that pair was assigned to the comparison condition, and the other to the program condition; if a 1 was generated, the opposite assignment was made.

*Training* Program school teachers attended a day and a half of training on RULER at the beginning of the academic year (fall 2008) and worked with a certified program coach for five sessions across the year. Training included didactic instruction, role-playing exercises, lesson design, feedback, and observations of instruction. Coaching sessions included sharing of successes and challenges, the provision of teaching tips and resources, and classrooms observations of lessons with feedback. Three optional "booster" training sessions also were offered to provide additional training. Attendance records showed that 88.9 % of teachers attended the training sessions, and 75.5 % of teachers attended at least four of the five coaching sessions ($M$=4.02, $SD$=0.92). Most teachers (91.1 %) attended at least one booster session.

*Intervention* Teachers in program schools were provided with 12 RULER units, each with five lessons or steps. Each unit was taught over a 2-week period and focused on one feeling word, such as "elation" or "shame." The steps were designed to last 15–20 min and to be integrated into regular classroom instruction. The five steps followed a basic structure. Briefly, teachers introduced the feeling word using a personalized connection (Step 1); students connected the feeling word to academic material (Step 2); students demonstrated the meaning of the word through a creative arts activity (Step 3); students discussed the feeling word with family members and wrote a short paragraph about the conversation (Step 4); and the class discussed and evaluated the effectiveness of different strategies for managing the feeling (Step 5).

*Data Collection* University-approved informed consent procedures were followed throughout the project. Each wave of data collection included observational assessments, teacher surveys, and student surveys. Observational data were collected using video cameras that ELA teachers set up in their classrooms following a training session with accompanying written instructions. Teachers recorded their ELA class (about 40 min in length) on three different days of "normal" classroom activity (i.e., not on days in which tests were administered or movies were shown). Each day of taping yielded one videotape for a maximum number of three videotapes per classroom per wave. Teachers completed surveys on their own time, and returned surveys and the videotapes using preaddressed, stamped envelopes. Research aides administered student surveys by reading aloud each item to which students filled in responses using a bubble format.

## Measures

*School-Level Variables* Treatment condition was the primary variable of interest (1 = *RULER*, 0 = *comparison*). Other school-level variables included: school size, minority composition, percentage of students receiving free/reduced lunch, student-teacher ratio, and a variable indicating whether the school had an existing SEL program (1 = *yes,* 0 = *no*).

*Classroom Emotional Climate (CEC)* CEC was measured using observational indicators, teacher reports, and student reports. *Observational assessments* of CEC were obtained using the Classroom Assessment Scoring System (CLASS; Pianta et al. 2008), which measures the quality of interactions among teachers and students (La Paro et al. 2004) and was developed from extensive national, federally-funded, observational studies, as well as thorough reviews of relevant literature (Hamre et al. 2006). The CLASS has face, construct, and predictive validity, and predicts academic and social adjustment (Brackett et al. 2011; Howes et al. 2008; NICHD Early Child Care Research Network 2003). The CLASS measures classroom social processes across three broad domains—emotional support, classroom organization, and instructional support—that are divided into subcategories or dimensions (Hamre et al. 2006). The observational assessment of CEC included scores on the emotional support domain, which consists of four dimensions: positive climate (degree of warmth and connection), negative climate (degree of negativity; reverse-coded), teacher sensitivity (teacher's awareness and responsiveness to students' academic and social needs), and regard for student perspectives (degree to which the classroom is focused on students' interests and motivations). Coders assign dimension scores using a 7-point scale (1–2 = *low*, 3–5 = *mid*, 6–7 = *high*) based on the presence or absence, frequency, and quality of specific observable indicators. The technical manual reports acceptable internal consistencies, inter-rater reliabilities, test-retest reliability coefficients, and correlations with academic and social outcomes (Hamre et al. 2006). Table 2 reports the inter-rater reliability coefficients (*r*s) on the emotional support dimensions, which ranged from .68 to .84.

To maintain reliability and prevent coding drift, several steps were taken. Each videotape was converted into two segments of equal length (average segment length=14.8 min), yielding up to six segments per classroom across the requested three days of taping. Segments were discarded when they were shorter than 8 min, they had poor audio quality, or students were not visible. Classrooms returning usable videotapes provided an average of five useable segments (range 1–6 segments). Coders completed a 2-day CLASS certification training, and were required to pass initial and weekly reliability testing. If a coder did not demonstrate reliability (i.e., 80 % of assigned codes within 1 point of the master code), CLASS master trainers worked with the coder until reliability was achieved and maintained. Segments were assigned randomly to one or more coders who were blind to the classroom's condition assignment and to the hypotheses of the study. Each segment was coded between one and three times by a unique coder, yielding an average of 13 sets of CLASS scores per classroom. Classroom dimension scores were computed by, first, averaging scores across all raters for a segment and, then, by aggregating scores across all segments. The average of the dimension scores was computed to create a classroom's emotional support domain score.

*Teacher surveys* included three assessments of CEC. The Classroom Supportiveness Scale had 14 items (4 reverse-scored) that assessed teachers' perceptions of students' behaviors toward each other, the extent to which those behaviors are prosocial (e.g., *Students in this class treat each other with respect;* Developmental Studies Center 2000). Teachers responded using a 5-point agreement scale (1 = *disagree a lot*; 5 = *agree a lot*). A 12-item Emotion-Focused Interaction Scale assessed the extent to which teachers engaged in personalized or emotion-focused interactions with their students. Teachers reported the frequency with which, in the last month, they engaged in a behavior (e.g., *Noticed a change in a student and acknowledged it to the student*). A 6-item Cooperative Learning Strategies Scale assessed the extent to which teachers created opportunities for students to interact. For example, teachers reported the frequency with which, in the last month, they *organized students for group activities so that they could work together.* Five-point Likert-type scales were used for the latter items (1 = *never*, 5 = *very often*). The research team designed the latter two scales. Higher scores reflected greater CEC. Cronbach's alphas were .88, .93, and .78, respectively.

*Student CEC surveys* included both the 8-item Affiliation with Teacher Scale (e.g., *My teacher respects my feelings;* Cook et al. 1995) and the 14-item Classroom Supportiveness

Scale (e.g., *Students in my class treat each other with respect;* Developmental Studies Center 2000). Students used a 5-point Likert-type response scale (1 = *disagree a lot*, 5 = *agree a lot*). Cronbach's alphas were .94 and .92, respectively. Classroom aggregated scores on each scale were used; higher scores reflected greater CEC.

*Treatment Fidelity and Program Enjoyment of RULER* Training logs, teacher reports, student reports, and coach reports were used to assess implementation fidelity. To measure the amount of the program being delivered (i.e., program dosage), teachers and students in the program group reported on which RULER units were taught during the academic year using a checklist format. Teachers also rated their enjoyment of the program, their perceptions of active student participation during program lessons, the extent to which they perceived students to enjoy the program lessons, as well as their perceptions of the quality of and their satisfaction with the coaching, using 5-point agreement scales (5 = *strongly agree*). Finally, coaches assessed the quality of lesson delivery through observations. They completed a best-practices checklist on three separate occasions, approximately 2 months apart, rating adherence to the RULER training protocol. Coaches also completed an overall impression of implementation quality at the end of the year by responding to this question for each teacher with whom they worked, *How would you rate the overall quality of this teacher's implementation of emotional literacy?*, using a 1 (*needs a lot of improvement*) to 5 (*excellent*) response scale.

## Results

*Randomization* To ensure that randomization was effective, baseline data from the two groups were compared. Table 1 reports the descriptive statistics for each CEC outcome variable. Only one difference emerged: Student reports of teacher affiliation were higher in the program group than in the comparison group, $\gamma$=0.25, *SE*=0.10, $t$=2.42, $p$=.01. We also compared the two groups on baseline school demographic variables. Program schools had fewer students (i.e., smaller school size) than comparison schools (*M*s=293.44 and 358.41, *SD*s=81.50 and 101.61, respectively), $t(62)$=2.82, $p$=.006. No other demographic differences emerged. Based on these analyses, we included student-rated teacher affiliation and baseline school size as covariates in subsequent analyses, as described later.

*Treatment Fidelity and Program Enjoyment* On average, teachers reported completing 7.38 (*SD*=2.79) of the 12 units (each consisting of five lessons), and students reported completing 6.97 (*SD*=2.16) units, reflecting high agreement between the two sources. On average, teachers reported that: they enjoyed program (*M*=4.08, *SD*=0.54), their students participated in the lessons (*M*=4.12, *SD*=0.81), and their students enjoyed the program (*M*=3.88, *SD*=0.84). Teachers also reported that the coach quality was high (*M*=4.34, *SD*=0.54) and that they were highly satisfied with the coaching (*M*=4.61, *SD*=0.56). Coaches' observations indicated that overall, teachers adhered to the implementation protocol,

**Table 1** Descriptive statistics classroom emotional climate (CEC) processes, Wave 1 (baseline) and Wave 3 (end-of-intervention year)

| Variable | Wave 1 | | | | Wave 3 | | | |
|---|---|---|---|---|---|---|---|---|
| | Reliability Coefficients | ICC% | *M* (*SD*) | | Reliability Coefficients | ICC% | *M* (*SD*) | |
| | | | RULER | Comparison | | | RULER | Comparison |
| Observations ($n$=66) | | | | | | | | |
| Emotional Support | .83 | 27.61 | 4.39 (0.53) | 4.41 (0.53) | .78 | 32.26 | 4.71 (0.48) | 4.46 (0.49) |
| Positive Climate | .73 | 26.21 | 4.66 (0.67) | 4.50 (0.70) | .75 | 15.51 | 4.52 (0.65) | 4.12 (0.72) |
| Negative Climate | .84 | 28.35 | 1.26 (0.41) | 1.15 (0.19) | .79 | 20.91 | 1.18 (0.27) | 1.19 (0.34) |
| Teacher Sensitivity | .68 | 24.30 | 4.55 (0.72) | 4.62 (0.69) | .59 | 31.86 | 4.17 (0.63) | 4.06 (0.61) |
| Regard for Student Perspectives | .80 | 32.14 | 3.63 (0.67) | 3.66 (0.75) | .69 | 56.85 | 3.31 (0.74) | 2.85 (0.67) |
| Teacher Reports ($n$=117) | | | | | | | | |
| Classroom Supportiveness | .88 | 20.05 | 3.54 (0.59) | 3.40 (0.56) | .88 | 20.66 | 3.53 (0.56) | 3.44 (0.52) |
| Emotion-Focused Interactions | .92 | 41.87 | 3.97 (0.64) | 3.74 (0.67) | .93 | 45.25 | 4.08 (0.67) | 3.79 (0.67) |
| Cooperative Learning Strategies | .80 | 40.82 | 3.63 (0.68) | 3.59 (0.61) | .78 | 36.78 | 3.93 (0.51) | 3.66 (0.59) |
| Student Reports ($n$=155) | | | | | | | | |
| Classroom Supportiveness | .88 | 41.52 | 2.98 (0.47) | 3.03 (0.47) | .92 | 41.45 | 2.75 (0.47) | 2.79 (0.52) |
| Teacher Affiliation | .92 | 27.90 | 3.70 (0.57) | 3.55 (0.59) | .94 | 47.38 | 3.65 (0.68) | 3.49 (0.63) |

Reliability coefficients are based on inter-rater reliability coefficients ($r$s) for observations, and Cronbach's alpha for teacher and student reports. ICC% = intraclass correlation coefficient percentage (amount of variation in the variable that is accounted for by schools)

meeting 79 % of quality indicators (*SD*=21 %); average evaluation quality was 3.29 (*SD*=1.02).

*Sample Attrition* Differences between groups in compliance rates with the data collection protocol were examined. With regards to observational (video) data, at Wave 2, 87 of 155 classrooms (56.1 %; *n*=45 program and 42 comparison) returned videotapes. At Wave 3, video data were available for 66 classrooms (42.6 %; *n*=34 program and 32 comparison) at Wave 3. Video data from both Waves 2 and 3 were available for 60 classrooms (38.7 %) from 29 schools distributed equally between program and comparison conditions, $\chi^2(1)=0.53$, $p=.467$. Teacher survey data from both Waves 2 and 3 were available for 104 classrooms (67 % of the sample) from 48 schools distributed equally between the two conditions, $\chi^2(1)=2.31$, $p=.129$. Wave 2 teacher survey data were available for 118 classrooms (*n*=66 program and 53 comparison), compared to 117 classrooms at Wave 3 (*n*=68 program and 49 comparison). All classrooms returned student surveys.

*Treatment of Missing Data* Hierarchical generalized linear modeling (HGLM), the multilevel version of a binomial logistic regression, was used to examine variables associated with Wave 3 missing data. Because teachers determined whether or not their classrooms were videotaped and whether or not to return teacher surveys, we examined which teacher-level (Level 1; number of classrooms taught; student-rated teacher affiliation and classroom supportiveness) and school-level (Level 2; school characteristics and demographics) variables were associated with returning videotapes and survey data at Wave 3. Teachers from schools that were smaller ($\gamma=-0.01$, $t=-2.24$, $p=.030$, odds=0.99) and that had lower percentages of students performing poorly in math ($\gamma=-0.07$, $t=-2.27$, $p=.027$, odds=0.94) were more likely to return videotapes. Teachers were more likely to return surveys when they were from schools with lower percentages of students performing poorly in ELA ($\gamma=-0.08$, $t=-2.67$, $p=.011$, odds=0.92), but greater percentages of students performing poorly in math ($\gamma=0.09$, $t=2.50$, $p=.016$, odds=1.10). No other significant differences emerged (*t*s<1.00). Return of videotapes at Wave 3 was unrelated to Wave 2 observer-rated *emotional support* ($\gamma=-.07$, *SE*=.12, $p=.569$), and teachers' return of survey data at Wave 3 was unrelated to Wave 2 teacher-rated classroom quality (*cooperative learning*, $\gamma=-.15$, *SE*=.18, $p=.418$ *classroom supportiveness*, $\gamma=-.15$, *SE*=.16, $p=.333$; *emotion-focused interactions,* $\gamma=-.14$, *SE*=.19, $p=.451$).

We also examined missing data patterns among teachers in program schools. Teachers who returned videotapes at Wave 3 attended significantly more booster sessions (*M*=2.38, *SD*=0.82) than teachers who did not return videotapes (*M*=1.48, *SD*=0.98), $t(43)=-3.34$, $p=.002$. Likewise, teachers who returned surveys at Wave 3 attended significantly more booster sessions (*M*=2.11, *SD*=0.94) than those who did not (*M*=1.25, *SD*=1.04), $t(43)=-2.31$, $p=.002$.

With the exception of school size, there were no significant relationships between variables that predicted missingness (e.g., school-averaged math and English performance, attendance at RULER booster sessions), and any of the CEC outcomes (*p*s>.05). Thus, the final models included only school size as a covariate (see Raudenbush et al. 2007).

In sum, the majority of teacher- and school-level variables tested were unrelated to missing data. Most importantly, we found no indication that missing observational data and/or teacher survey data might be a reflection of classroom CEC or influenced by treatment condition. Accordingly, we considered the pattern of missing data as largely "at random" (Graham 2009; Rubin 1987), and applied multiple imputations to maximize power (Graham 2009; Rubin 1987; Schafer 1999). However, it is possible that unobserved variables including characteristics of the school (i.e., size, students with average math and English performance) or of the teacher (i.e., engagement, motivation) influenced teachers' propensity to return data at Wave 3.

Given the large proportion of missing observations (89 out of 155 missing at Wave 3) and the moderate proportion of missing data in teacher surveys (36 out of 155 missing at Wave 3), we followed a conservative approach to missing data imputation. Wave 3 missing values were imputed only when Wave 2 data were available for that classroom. Sample size increased from 66 to 92 for observational data, and from 119 to 132 for teacher survey data. Wave 2 scores of the same measures (i.e., CLASS scores, cooperative learning strategies, classroom supportiveness, emotion-focused interaction, respectively) were used as auxiliary variables to impute Wave 3. Imputing on alternative measures of the dependent variable, when these are not related to patterns of missing data, is considered the best strategy to reduce bias due to differential response rates (Graham 2009). We imputed data using NORM 2.03 (Schafer 2000), and created 10 new data sets, as prescribed by Schafer (1999). Results from both imputed and non-imputed analyses were similar and are available from the authors upon request.

*Correlations Between CEC Measures* Table 2 shows the intercorrelations among CEC measures. There were moderate to high correlations among the observer ratings of CEC (i.e., CLASS emotional support dimensions). Among teacher reports, all but emotion-focused interactions and classroom supportiveness correlated significantly. Student-rated classroom supportiveness and teacher affiliation correlated significantly. Observer ratings were mostly unrelated to teacher and student reports with two exceptions: observer-rated negative climate correlated negatively with teacher-rated classroom supportiveness, and observer-rated

**Table 2** Intercorrelations among observations, teacher, and student reports of classroom emotional climate (CEC) processes at Wave 1 (baseline) and Wave 3 (end-of-intervention year)

| Variable | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Wave 1 (above diagonal) | | | | | | | | | |
| Observations | | | | | | | | | | | |
| 1. Emotional Support | | – | .92*** | −.64*** | .92*** | .89*** | .25* | −.07 | .24* | .31** | .14 |
| 2. Positive Climate | Wave 3 (below diagonal) | .92*** | – | −.44*** | .79*** | .80*** | .26* | −.01 | .20 | .32** | .20 |
| 3. Negative Climate | | −.61*** | −.44*** | – | −.62*** | −.41*** | −.13 | .13 | −.10 | −.07 | .02 |
| 4. Teacher Sensitivity | | .86*** | .80*** | −.48*** | – | .69*** | .23* | −.09 | .22* | .25* | .03 |
| 5. Regard for Student Perspectives | | .84*** | .68*** | −.41*** | .53*** | – | .23* | −.06 | .26* | .33** | .20 |
| Teacher Reports | | | | | | | | | | | |
| 6. Classroom Supportiveness | | .25 | .23 | −.31* | .19 | .17 | – | .19* | .41*** | .45*** | .13 |
| 7. Emotion-Focused Interactions | | .19 | .19 | .02 | .19 | .19 | .15 | – | .46*** | .12 | .27** |
| 8. Cooperative Learning Strategies | | .30 | .23 | −.16 | .21 | .36** | .22* | .61*** | – | .27** | .33*** |
| Student Reports | | | | | | | | | | | |
| 9. Classroom Supportiveness | | .16 | .21 | .03 | .11 | .17 | .45*** | .09 | .26** | – | .32*** |
| 10. Teacher Affiliation | | .22 | .24 | .03 | .21 | .21 | .22* | .36*** | .32** | .44*** | – |

*$p$<.05; **$p$<.01; ***$p$<.001

regard for student perspectives correlated positively with teacher-rated cooperative learning. There were significant positive correlations between teacher and student reports with one exception: student-rated classroom supportiveness did not correlate with teacher-rated emotion-focused interactions. Given the lack of complete convergence, subsequent analyses looked at the measures separately.

*Impacts of RULER* Program impacts were examined using an intent-to-treat analysis and a multilevel framework to account for the nesting of classrooms within schools. To obtain an unbiased estimate of the treatment effect, the intent-to-treat analysis computed the average effect of offering RULER to members of the program and comparison groups (Bloom 2005). Multilevel or hierarchical linear modeling (HLM) allows for the simultaneous estimation of variation associated with both classroom- (level 1) and school-level (level 2) components, while also taking into consideration homogeneity of variance of residuals, yielding less biased parameter estimates (Raudenbush and Bryk 2002). All analyses were conducted using the HLM software package (Version 6.06), employing full information maximum likelihood estimation for the models. The formulas used to estimate program impacts on each of the outcomes were:

$$\text{Level 1} \quad Y_{ij} = \beta_{0j} + r_{ij}$$
$$\text{Level 2} \quad \beta_{0j} = \gamma_{00} + \gamma_{01}(size) + \gamma_{02}(RULER) + \mu_{0j}$$

In Level 1, the CEC outcome variable ($Y$) for a classroom ($i$) who is in school ($j$) is a function of the mean CEC score ($\beta_{0j}$) and the error term associated with this estimated mean ($r_{ij}$). In Level 2, the adjusted mean CEC outcome variable for classrooms in each school ($\beta_{0j}$) is a function of the grand mean CEC ($\gamma_{00}$), school size ($\gamma_{01}$; grand-mean centered), and assignment to condition ($\gamma_{02}$; 1 = program, 0 = comparison). Exploratory analyses, conducted in HLM, identified which school-level predictors to include in the analyses. School-level demographic variables were not associated with the CEC outcome variables; thus, following recommendations from Raudenbush and colleagues (2007), the non-significant predictors were excluded to increase power. School size was included as a covariate to account for biases both in school size between comparison and program schools and in return rates (see Missing Data section).

To obtain the estimated or adjusted mean CEC score of program and comparison schools, the parameter estimate associated with assignment to condition (i.e., $\gamma_{02}$) is examined. Because the intercept is the estimated mean value of CEC when all the values of the predictors in the model are zero—in this case, a school of average size and assignment to condition is comparison—the estimated value of CEC for the comparison group (controlling for school size) is equivalent to the value of the intercept ($\gamma_{00}$). Because $\gamma_{02}$ is interpreted as the discrepancy between the comparison and program conditions, the estimated value of CEC for the program group is $\gamma_{00}+\gamma_{02}$*1. Effect sizes were computed using Hedges' $g$.

Table 3 summarizes the adjusted CEC means and standard errors of both program and comparison schools across all measures, and includes the effect sizes. Consistent with our theoretical model and hypotheses, assignment to the program condition impacted CEC as assessed by both observer and teacher ratings. Assignment to the RULER

**Table 3** End-of-year impacts on classroom emotional climate (CEC) processes

| | Adjusted means | | *SE* | *p* | Effect size |
|---|---|---|---|---|---|
| | RULER | Comparison | | | |
| Observations | | | | | |
| Emotional Support domain | 4.72 | 4.46 | 0.12 | .048 | .50 |
| Positive climate | 4.51 | 4.08 | 0.15 | .007 | .55 |
| Negative climate | 1.17 | 1.23 | 0.08 | .468 | |
| Teacher sensitivity | 4.26 | 4.08 | 0.14 | .212 | |
| Regard for student perspectives | 3.31 | 2.84 | 0.21 | .030 | .60 |
| Teacher reports | | | | | |
| Classroom supportiveness | 3.61 | 3.43 | 0.12 | .148 | |
| Emotion-focused interactions | 4.10 | 3.76 | 0.16 | .038 | .52 |
| Cooperative learning strategies | 3.95 | 3.64 | 0.13 | .021 | .53 |
| Student reports | | | | | |
| Classroom supportiveness | 2.78 | 2.75 | 0.11 | .810 | |
| Teacher affiliation | 3.67 | 3.43 | 0.15 | .109 | |

Computations for observations and teacher reports were based on multiple imputations, classroom *n*s=92 and 132, respectively. School size was used as a covariate in all analyses. For student reports of teacher affiliation, baseline school-level mean scores for this variable was added as a covariate due to significant differences between RULER and comparison groups at baseline, $\gamma$=0.25, *SE*=0.10, *t*=2.42, *p*=.01

group was associated with higher observer ratings of overall emotional climate, and further analyses of the scale's dimensions showed that the program significantly impacted positive climate and teacher regard for student perspectives. Assignment to use RULER also impacted teacher-rated emotion-focused interactions and cooperative learning strategies. Effect sizes were moderate. Student-rated CEC did not demonstrate program impact.

## Discussion

This RCT shows that schools assigned to use RULER had classrooms with higher emotional climate at the end of the school year compared to schools assigned to the comparison group. After the first 8 months of implementation, classrooms randomized to use RULER were rated both by observers and teachers as having a more positive emotional climate in contrast to those randomized to the comparison group. Specifically, observers blind to the hypotheses and condition assignment rated RULER classrooms as having more behavioral markers reflecting positive climate (e.g., the presence of positive interactions and personal connection) and regard for student perspective (e.g., the extent to which teachers incorporate students' ideas and interests in their teaching versus using a teacher-driven approach). Teachers using RULER were more likely to report that they interacted with students in emotion-focused ways and created more opportunities for students to interact through cooperative learning than did teachers in the comparison group. RULER professional development and the classroom curriculum promote these kinds of interactions, which are instrumental to both positive youth development (NRC/IOM 2002) and effective SEL programming (Durlak et al. 2011). The effect sizes reported in Table 3 are somewhat larger (range=0.50–0.60) than those reported in the meta-analysis of SEL programs for more distal outcomes like attitudes and academic performance (range=0.24–0.27), and align well with those for emotion skills (Durlak et al. 2011, Table 5), which are more proximal, like the emotional climate outcomes of this study.

This study provides a rigorous test of the impact of an SEL program on the classroom setting. RCTs control for demographic variation and related socioeconomic factors that may contribute to outcomes, thus offering authoritative evidence of program impact (e.g., Bloom 2005). Such experimental tests are critical for establishing the evidence base that SEL programs can improve learning environments. Further, the CLASS reduces potential biases of self-report assessments while assessing classroom interactions using extensively trained and monitored coders who are blind to condition assignment. The use of the CLASS also allows for comparisons of the impact of other SEL programs that use this measure (e.g., Brown et al. 2010; Raver et al. 2008). This study extends current findings by identifying the classroom climate as a potential mechanism by which student outcomes may be affected (Durlak et al. 2011), and future research will test the hypothesis that changing the emotional climate is the mechanism by which RULER impacts student achievement (Brackett et al. 2010).

The study's findings also are notable because they were gleaned from a theoretically-grounded program. Created with the goal of promoting the social and emotional aspects of

students' lives and the school environment simultaneously, RULER provides opportunities for students and educators to build knowledge and skills. The program components encourage frequent social interactions between peers and teachers, along with activities that focus on creativity, group problem solving, conflict resolution, and empathy (Rivers and Brackett 2011).

## Limitations and Future Directions

There are some important limitations of this study with implications for future research. First, longer-term impact needs to be examined, as many interventions for positive youth development do not maintain positive effects after initial implementation (Catalano et al. 2004). Further, although this RCT had sufficient power to detect an effect on setting-level outcomes, it lacked the power for subgroup analyses to explore the conditions under which RULER may be most effective at impacting the setting (Bloom and Michalopoulos 2011). For example, we could not determine the level of dosage under which the impact of RULER is the greatest, or the characteristics of teachers for which the program has the most (or least) impact.

Second, return rates for video data were low, which may reflect bias in the findings. Although we did not identify correlates of missingness in the data with the exception of school size, we cannot rule out the possibility that other unobserved factors contributed to the observed pattern of missingness (Graham 2009). It is possible that more motivated or conscientious teachers, for example, were more likely to return videos and also may have been more responsive to the intervention. Informal surveys with participating teachers across both conditions about their willingness or lack thereof to videotape their classrooms did not yield additional insight into the low return rates. Some teachers stated that they felt uncomfortable with being on camera, others stated they did not want to spend time taping, and others offered no response. The rates of missing data were consistent between conditions, suggesting that the intervention did not bias attrition. Nevertheless, the amount of missing data and the potential for unobserved variables influencing attrition limits the generalizability of these results and necessitates replication of the findings in other populations. A substantively different research approach would be to determine how to estimate treatment effects based on post-treatment choice (Peck 2003). Determining how to estimate teachers' propensity to buy into the intervention would be a challenge for the present study because randomization was conducted at the level of the school and not at the level of the teacher. Future research should address this important issue.

Third, student reports showed no impact of RULER, nor were student reports correlated with the observational indicators. It is possible that the student measures may not be the best assessment of the outcome of interest. The student surveys focused more on liking for teachers and interactions between students than on overall climate of the classroom. Students may not perceive the climate changes detected by observational instruments.

Finally, a confound within the study may be that RULER professional development was only offered to the program group and may have been the active ingredient that produced positive outcomes. Further, teachers in the program group may have reported paying more attention to the emotional qualities of the classroom as that message was central to RULER professional development. However, the observational data does corroborate the teacher survey data (i.e., their interactions reflected greater emotional support). Additional evidence from an RCT that includes an alternative professional development opportunity for the comparison group may help to elucidate the sources of the intervention effects.

## Implications

To accomplish an educational agenda that addresses the needs of the whole child, school-based programming should strive to improve the quality of the environments in which academic, social, and emotional learning occurs (e.g., Greenberg et al. 2003). SEL programs like RULER provide this opportunity. Evidence-based SEL curricula address social and emotional competencies both directly and indirectly across the curriculum. In addition to boosting academics, SEL programs intend to teach students the broad range of skills needed to be successful in society, cultivate quality relationships, and be psychologically and physically healthy. Yet, their success is dependent upon the extent to which learning occurs in caring, supportive, and empowering settings (e.g., Bronfenbrenner 1979; Catalano et al. 2004). The present study provides evidence that RULER is able to impact the learning environment in these important ways.

## References

Battistich, V., Solomon, D., Watson, M., & Schaps, E. (1997). Caring school communities. *Educational Psychologist, 32*, 137–151. doi:10.1207/s15326985ep3203_1.

Bloom, H. S. (Ed.). (2005). *Learning more from social experiments: Evolving analytic approaches*. New York: Russell Sage.

Bloom, H. S., & Michalopoulos, C. (2011). When is the story in the subgroups? *Prevention Science*. Advance online publication. doi:10.1007/s11121-010-0198-x.

Brackett, M. A., Patti, J., Stern, R., Rivers, S. E., Elbertson, N. A., Chisholm, C., & Salovey, P. (2009). A sustainable, skill-based approach to building emotionally literate schools. In D. Thompson, M. Hughes, & J. Terrell (Eds.), *The handbook of developing emotional and social intelligence: Best practices, case studies, & tools* (pp. 329–358). New York: Pfeiffer.

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*. Advance online publication. doi:10.1016/j.lindif.2010.10.002.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2011). Classroom emotional climate, teacher affiliation, and student conduct. *Journal of Classroom Interaction, 46*, 27–36.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Brown, J. L., Jones, S. M., LaRusso, M., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs program. *Journal of Educational Psychology, 102*, 153–167. doi:10.1037/a0018160.

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

Catalano, R. F., Hawkins, J. D., Berglund, L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*, 230–239. doi:10.1016/S1054-139X(02)00496-2.

Catalano, R. F., Berglund, L., Ryan, J. A. M., Lonczek, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *The Annals of the American Academy of Political and Social Science, 591*, 98–124. doi:10.1177/0002716203260102.

Cook, E. T., Greenberg, M. T., & Kusche, C. A. (1995). *People in my life survey: Attachment relationships in middle childhood*. Paper presented at the meeting of the Society for Research in Child Development, Indianapolis, IN.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Developmental Studies Center. (2000). Student Questionnaire, Child Development Project (Elementary School Version). Department. Oakland, CA. Retrieved from www.devstu.org.

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.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist, 48*, 90–101. doi:10.1037/0003-066X.48.2.90.

Fine, S. E., Izard, C. E., Mostow, A. J., Trentacosta, C. J., & Ackerman, B. P. (2003). First grade emotion knowledge as a predictor of fifth grade self-reported internalizing behaviors in children from economically disadvantaged families. *Development and Psychopathology, 15*, 331–342. doi:10.1017/S095457940300018X.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576. doi:10.1146/annurev.psych.58.110405.085530.

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. doi:10.1037/0003-066X.58.6-7.466.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness & the transition to kindergarten in the era of accountability* (pp. 49–83). Baltimore, MD: Brooks Publishing Co.

Hamre, B. K., Mashburn, A. J., Pianta, R. C., Locasle-Crouch, J., & La Paro, K. M. (2006). CLASS: Classroom Assessment Scoring System, Technical Appendix. Department. Retrieved from http://www.classobservation.com/research/class_tech_manual.pdf.

Howes, C., Burchinal, M., Pianta, R. C., Bryant, D., Early, D. M., Clifford, R., & Barbarin, O. A. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly, 23*, 27–50. doi:10.1016/j.ecresq.2007.05.002.

Institute of Medicine. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal, 104*, 409–426. doi:0013-5984/2004/10405-0004.

Maurer, M., & Brackett, M. A. (2004). *Emotional Literacy in the Middle School: A 6-step program to promote social, emotional, & academic learning*. Port Chester, NY: National Professional Resources.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist, 63*, 503–517. doi:10.1037/0003-066X.63.6.503.

Moos, R. H. (1979). *Evaluating educational environments*. San Francisco: Jossey-Bass.

National Research Council & Institute of Medicine [NRC/IOM]. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

NICHD Early Child Care Research Network. (2003). Social functioning in first grade: Associations with earlier home and child care predictors and with current classroom experiences. *Child Development, 74*, 1639–1662. doi:0009-3920/2003/7406-0007.

Olweus, D. (1997). Bully/victim problems in school: Facts and intervention. *European Journal of Psychology of Education, 12*, 495–510. doi:10.1007/BF03172807.

Peck, L. R. (2003). Subgroup analysis in social experiments: Measuring program impacts based on post-treatment choice. *American Journal of Evaluation, 24*, 157–187. doi:10.1177/109821400302400203.

Pianta, R. C., La Paro, K. M., Payne, C., Cox, M., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal, 102*, 225–238. doi:10.1086/499701.

Pianta, R. C., La Paro, K., & Hamre, B. (2008). *Classroom assessment scoring system: K-3*. Baltimore, MD: Brookes Publishing.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis, 29*, 5–29. doi:10.3102/0162373707299460.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champon, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly, 23*, 10–26. doi:10.1016/j.ecresq.2007.09.001.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). The emotional climate of learning: Links to academic achievement. *Journal of Educational Psychology*. Advance online publication. doi:10.1037/a0027268

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. doi:10.1080/10573569.2011.532715.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Mayer, J. D., Caruso, D. R., & Salovey, P. (2012). Measuring emotional intelligence in early adolescence with the MSCEIT-YV: Psychometric properties and relationship with academic performance and psychosocial functioning. *Journal of Psychoeducational Assessment, 30*(4), 344–366. doi:10.1177/0734282912449443.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185–211. doi:10.2190/DUGG-P24E-52WK-6CDG.

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research, 8*, 3–15. doi:10.1177/096228029900800102.

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

Schutz, P. A., & Pekrun, R. (Eds.). (2007). *Emotion in education*. New York: Academic.

Shinn, M., & Yoshikawa, H. (Eds.). (2008). *Toward positive youth development: Transforming schools and community programs* (pp. 3–17). New York: Oxford University Press.