

A School-Randomized Clinical Trial of an Integrated Social–Emotional Learning and Literacy Intervention: Impacts After 1 School Year

Stephanie M. Jones
Harvard University

Joshua L. Brown
Fordham University

Wendy L. G. Hoglund
University of Alberta, Edmonton

J. Lawrence Aber
New York University

Objective: To report experimental impacts of a universal, integrated school-based intervention in social–emotional learning and literacy development on change over 1 school year in 3rd-grade children’s social–emotional, behavioral, and academic outcomes. **Method:** This study employed a school-randomized, experimental design and included 942 3rd-grade children (49% boys; 45.6% Hispanic/Latino, 41.1% Black/African American, 4.7% non-Hispanic White, and 8.6% other racial/ethnic groups, including Asian, Pacific Islander, Native American) in 18 New York City public elementary schools. Data on children’s social–cognitive processes (e.g., hostile attribution biases), behavioral symptomatology (e.g., conduct problems), and literacy skills and academic achievement (e.g., reading achievement) were collected in the fall and spring of 1 school year. **Results:** There were main effects of the 4Rs Program after 1 year on only 2 of the 13 outcomes examined. These include children’s self-reports of hostile attributional biases (Cohen’s $d = 0.20$) and depression ($d = 0.24$). As expected based on program and developmental theory, there were impacts of the intervention for those children identified by teachers at baseline with the highest levels of aggression ($d = 0.32$ – 0.59) on 4 other outcomes: children’s self-reports of aggressive fantasies, teacher reports of academic skills, reading achievement scaled scores, and children’s attendance. **Conclusions:** This report of effects of the 4Rs intervention on individual children across domains of functioning after 1 school year represents an important first step in establishing a better understanding of what is achievable by a schoolwide intervention such as the 4Rs in its earliest stages of unfolding. The first-year impacts, combined with our knowledge of sustained and expanded effects after a second year, provide evidence that this intervention may be initiating positive developmental cascades both in the general population of students and among those at highest behavioral risk.

Keywords: intervention, social–emotional skills, academic skills, school-randomized, elementary school

Over the last decade, children’s developmental pathways toward or away from social–emotional, mental health, and academic problems have become increasingly well understood (e.g., Tolan & Gorman-Smith, 2002). Fortunately, this period has also witnessed significant advances in prevention science, resulting in more conceptually grounded and empirically validated preventive interventions to reduce mental health and problem behaviors in elementary school (e.g., Hoagwood et al., 2007). In particular, knowledge has

grown about the effectiveness of school-based intervention strategies to reduce children’s risk for future aggressive and violent behavior (e.g., Aber, Brown, & Jones, 2003; Conduct Problems Prevention Research Group [CPPRG], 1999, 2004, 2007).

Although the literature on school-based preventive interventions is rich with studies of interventions targeted at subgroups of high-risk children, it has only recently expanded to studies of whole-school interventions implemented with general populations

Stephanie M. Jones, Graduate School of Education, Harvard University; Joshua L. Brown, Department of Psychology, Fordham University; Wendy L. G. Hoglund, Department of Psychology, University of Alberta, Edmonton; J. Lawrence Aber, Steinhardt School of Education, Culture, and Human Development, New York University.

This research was supported by grants from the Institute of Education Sciences, U.S. Department of Education, in collaboration with the Centers for Disease Control and Prevention, U.S. Department of Health and Human Services (Grant R305L030003), and the William T. Grant Foundation. We thank Tom Roderick (executive director), Audrey Major (program coordinator), and the staff of the Morningside Center for Teaching Social Responsibility for their work in developing and implementing the 4Rs Program and their collaboration in this research.

We thank our research team, particularly our site coordinator, Genevieve Okada; our data manager, Wendy L. G. Hoglund; our dedicated field and data coordination staff, including Juliette Berg, Suzanne Elgendy, Matthew Kim, Vanessa Lyles, Emily Pressler, Catalina Torrente, and Kathleen Zadzora; our exceptional team of data collectors; and our project administrator, Shirley Archer-Fields. Finally, we especially thank the students, teachers, assistant principals, and principals in each of the research schools for their participation in and cooperation with this study.

Correspondence concerning this article should be addressed to Stephanie M. Jones, Graduate School of Education, Harvard University, Larsen 603, 14 Appian Way, Cambridge, MA 02138. E-mail: jonesst@gse.harvard.edu

of students (e.g., Hahn et al., 2007). But even among the best studies of such universal school-based interventions, various methodological challenges have limited the quality and generalizability of the knowledge base (Hundert et al., 1999). For example, with the exception of several to date (e.g., CPPRG, 1999, 2004, 2007; Ialongo, Poduska, Werthamer, & Kellam, 2001; Kellam, Ling, Merisca, Brown, & Ialongo, 1998), few studies have employed school-randomized designs that support definitive causal statements about the impact of the intervention on outcomes. There are also broader challenges facing this field. For example, there remains an unresolved tension in the movement to reform education between standards-based accountability, with its focus on academic achievement, on one hand and social and emotional development, with its focus on social-emotional competence, behavior, and mental health, on the other. The present study has been designed to contribute to ongoing scholarship in the school-based prevention of social-emotional, behavioral, and academic problems by reporting experimental impacts of a universal, integrated school-based intervention in social-emotional learning and literacy development (the 4Rs Program: Reading, Writing, Respect, and Resolution) on change in third-grade children's social-emotional, behavioral, and academic functioning after 1 school year of intervention.

A Theory of Change for the 4Rs Program

Recent evidence suggests the value of effectively integrating academic and social-emotional intervention efforts and of examining the direct effects of such integrated interventions on both domains of development (e.g., Durlak, Weissberg, Dymnicki, Taylor, Schellinger, in press). The value of such integration is supported by a mounting body of research linking successful school adjustment and performance with the development of social-emotional competence (Miles & Stipek, 2006; Rogoff, 1990; Wentzel & Asher, 1995). To date, there has been greater emphasis on developing and testing integrated approaches to building social-emotional and literacy skills in the preschool period (e.g., Bierman et al., 2008; Raver et al., 2009), but integrated approaches in the elementary school years are needed as well (e.g., Domitrovich et al., 2010).

The 4Rs Program is a universal, school-based intervention for Grades K–5 focused on social-emotional learning and literacy development. Pedagogically, 4Rs links the teaching of social-emotional competencies and the teaching of fundamental academic skills, thereby capitalizing on their mutual influence on successful development (Hinshaw, 1992). (See Method section for a full program description.) Building off its programmatic predecessor, the Resolving Conflict Creatively Program (RCCP; Aber et al., 2003), the 4Rs Program can be characterized as a “developmental intervention” (National Research Council, 1993), because it focuses on changing underlying mental and interpersonal processes that lead to aggression and violence when they are still in the formative stage. The theoretical model underlying the social-emotional learning core of the 4Rs Program emphasizes social-cognitive and interpersonal processes that link individual, family, and community risk factors to the development of aggressive behavior and that place children at higher risk for a broader set of mental, emotional, and behavioral problems (e.g., Coie & Dodge, 1998; O’Connell, Boat, & Warner, 2009).

The key mental and interpersonal social-cognitive processes directly targeted by 4Rs include hostile attributional bias, namely, the tendency to attribute hostile intent to an ambiguous or prosocial cue (e.g., Dodge, Bates, & Pettit, 1990); normative beliefs about aggression, or the perceived acceptability of aggressive responses under varying conditions of provocation; aggressive and prosocial fantasies, a form of cognitive script rehearsal about aggressive/antisocial and prosocial interactions and events (Huesmann, Guerra, Miller, & Zelli, 1992); and aggressive interpersonal negotiation strategies, or children’s propensity to select aggressive over prosocial responses in hypothetical problem-solving situations (Leadbeater, Hellner, Allen, & Aber, 1989). Theory and both basic and applied research converge to suggest that these social-cognitive processes (a) are affected by certain types of experiences (e.g., a history of harsh, punitive, or abusive parenting or exposure to community violence; Coie & Dodge, 1998) or a peer environment in which violence is normative (Tremblay, Masse, Vitaro, & Dobkin, 1995). In turn, these social-cognitive processes (b) increase the probability of aggression and violence by children and youths (e.g., Dodge, Laird, Lochman, & Zelli, 2002) and are linked to a broader set of mental, emotional, and behavioral problems (e.g., Domitrovich et al., 2010).

Theory Guiding the Evaluation of the 4Rs Program

Our approach to the evaluation of the 4Rs Program is guided both by a multilevel program theory (Jones, Brown, & Aber, 2008) and by developmental cascades theory (Masten & Cicchetti, 2010; Masten, Long, Kuo, McCormick, & Desjardins, 2009). With regard to multilevel program theory, three major sets of premises guided the design of 4Rs and have direct implications for the design and conduct of its evaluation. First, 4Rs promotes change processes at multiple levels (e.g., individuals, their interactions, and proximal settings including their classroom and school environments; Tseng & Seidman, 2007). Second, 4Rs was designed based on the idea that improving functioning in one domain (e.g., interpersonal interactions) influences functioning in other domains (e.g., academic engagement and attention; Domitrovich et al., 2010; Guerra & Bradshaw, 2008). Finally, 4Rs is designed to promote change at multiple levels in multiple domains over time as dynamic systems (e.g., Cox, Mills-Koonce, Propper, & Garipey, 2010; Jones et al., 2008).

Other whole-school reform strategies in social-emotional learning appear to be founded on these (and other) premises. Targets of the intervention are processes at multiple levels (individual, relational, peer networks, classrooms, whole schools; e.g., Seattle Social Development Project, Fast Track) and in multiple domains (social-cognitive processes, social relational processes, peer and classroom functioning, whole-school climate; e.g., the Good Behavior Game, the Multisite Violence Prevention Project). Further, it is believed that changing nested systems (i.e., children in classrooms, classrooms in schools) requires extended periods of time (Farrell & Vulin-Reynolds, 2007; Hawkins, Guo, Hill, Battin-Pearson, & Abbott, 2001). Some small or modest changes may appear in some domains of functioning at some levels early on in the intervention, but it is likely that larger and more significant changes will occur in more domains and across more levels later on. In multiyear whole-school interventions, this unfolding of intervention effects across domains of functioning and levels of the

system is influenced both by the intensity and quality of intervention implementation and by the rhythm of summer breaks and yearly re-sorting of teachers and children.

Although these premises are shared by many social-emotional learning and whole-school intervention strategies, they are not typically reflected in the design and conduct of their evaluations. Furthermore, until very recently, program thinking about the unfolding of intervention effects across domains and levels and over time was rather nonspecific (e.g., Aber et al., 2003; Aber, Jones, Brown, Chaudry, & Samples, 1998). It is impossible to confidently interpret modest early impacts as promising or discouraging without employing developmental and other theoretical perspectives to guide evaluation design for multidomain, multilevel, dynamic systems intervention (Hawkins et al., 2001). Fortunately, this gap in using theory to guide the analysis and interpretation of unfolding intervention effects has begun to be addressed by a number of scholars and researchers. In this project, we drew heavily and explicitly on developmental cascades theory and research. According to Masten and Cicchetti (2010), "Developmental cascades" refer to the cumulative consequences for development of the many interactions and transactions occurring in developing systems that result in spreading effects across levels, among domains at the same level, and across different systems or generations" (p. 491). As Masten and Cicchetti noted, developmental cascades are referred to in other literatures as chain reaction, snowball, amplification, spillover, or progressive effects.

On the basis of extant theory and empirical research, we hypothesized both short- and longer term effects of 4Rs on children's outcomes. Our expectations for short-term effects after 1 school year were in the social-cognitive domain, particularly in those areas in which we saw change in our evaluation of the RCCP Program after 1 school year (Aber et al., 1998): children's hostile attributional biases, aggressive fantasies, and aggressive interpersonal negotiation strategies. The 4Rs programmatic approach of embedding lessons in social-emotional learning in a balanced literacy delivery strategy supports our expectation for short-term effects on children's basic literacy skills. In addition, we expected longer term effects on children's aggressive and socially competent behavior and academic achievement that might require more than 1 year of intervention to emerge. We also hypothesized that while there might be evidence of small or modest spillover effects early in the intervention (e.g., impacts on other mental health outcomes such as depression or attention problems), if 4Rs is implemented with fidelity over several years, more evidence of spillover would become evident in later years. Finally, given prior results from our evaluation of the RCCP in which intervention effects were robust across key demographic subgroups including gender, race/ethnicity, and socioeconomic status (Aber et al., 2003), we did not expect variation in intervention effects for these subgroups. However, variation in intervention effects by levels of aggression at baseline was expected consistent with growing evidence of stronger intervention effects of school-based social-emotional learning programs for children rated as high on preintervention measures of aggression (e.g., CPPRG, 2007; Farrell, Meyer, Sullivan, & Kung, 2003).

On the basis of these expectations and in order to effectively assess the theory of developmental cascades underlying the 4Rs Program and its evaluation design, we planned to report the results of the 4Rs study in a staged manner over time (Masten & Cicchetti,

2010). Accordingly, we plan to report (a) effects separately by years of intervention (Years 1, 2, and 3), (b) effects separately by level (child, classroom, school), and (c) effects simultaneously across domains (social-emotional and academic learning). Our first set of reports focuses on findings after 1 year of intervention at the individual child level (this report) and the classroom level (Brown, Jones, LaRusso, & Aber, 2010). Our next report focuses on impacts after 2 years of intervention at the child level (Jones et al., in press), and we are planning reports at the teacher and classroom levels after 2 years. Additional reports will follow this strategy through the full 3 years of the intervention.

Method

This study was conducted in one of seven sites that participated in the Social and Character Development (SACD) Research Program funded by the U.S. Department of Education's Institute of Education Sciences, in collaboration with the National Center for Injury Prevention and Control at the Centers for Disease Control (CDC). The Institute of Education Sciences and CDC selected seven grantees to implement and evaluate school-based SACD interventions using a random-assignment, 3-year longitudinal design. Data included as part of the larger SACD research program is presented in the not yet released National Evaluation Report (SACD Research Consortium, 2008).

Procedures

School recruitment and randomization. Building on the long-standing relationships in the school community developed by the 4Rs Program practitioners at the Morningside Center for Teaching Social Responsibility (MCTSR), we originally identified 41 schools as potential participants in the 4Rs evaluation (see Figure 1). After a recruitment visit and detailed explanation of the intervention and study design, 24 of these 41 agreed to the process of matching and randomization. Prior to randomization, a pairwise matching procedure was used to ensure demographic similarity of intervention and control groups. An algorithm developed by Mathematica Policy Research (SACD Research Consortium, 2008) was employed and enabled us to select the most closely matched pairs of schools by minimizing the multivariate distance between school pairs based on 20 measurable demographic and school characteristics. These variables were drawn from the 2001-2002 administrative databases kept by New York City's Department of Education and were selected to represent a number of dimensions related to the outcomes. They include such variables as the number of students, percentage of students receiving a free lunch, racial/ethnic composition, student achievement, average spending per student, and teacher experience.

The nine best matched pairs of schools were initially selected for inclusion in the study with the three remaining pairs reserved as backups. The first backup pair was eliminated from consideration by a local instructional superintendent. Two of our initially selected best matching pairs were dropped (in one case the principal was previously trained in the RCCP—an original exclusion criterion—and in the other, the principal was unwilling to proceed with the study if assigned to the control group) and were replaced with our second and third backup pairs.

To conduct random assignment, we employed a random numbers generator to generate, in sequence, nine random numbers

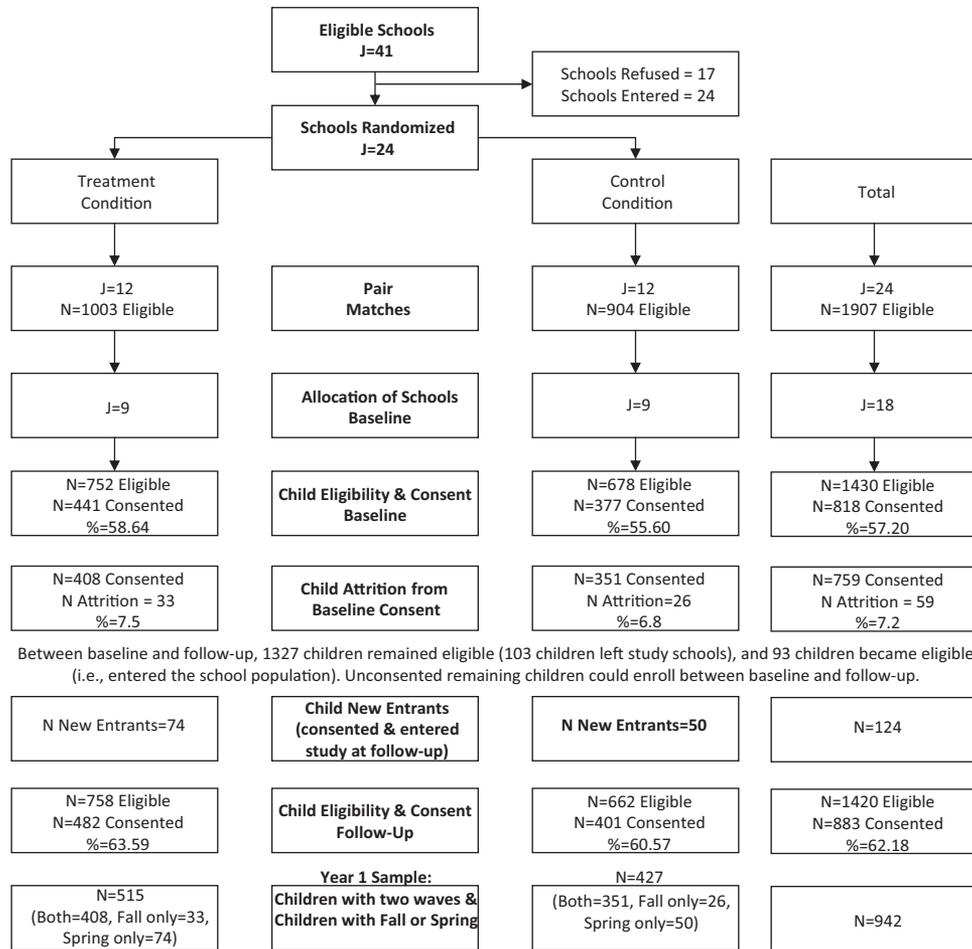


Figure 1. 4Rs evaluation participant flowchart. J = Schools.

ranging from 0 to 1 that were assigned to the first school in each of the nine pairs. Based on the randomly generated number, the first school in each pair was assigned to the intervention or the control group and the second school in the pair was assigned to the other group. After random assignment, the two groups were compared across the 20 demographic characteristics employed in the matching procedure. As expected, the two groups did not differ significantly on any of these characteristics and eta squared values were minimal. The final set of schools was racially and ethnically diverse, composed primarily of students who received a free school lunch, and characterized by attendance rates over 89% and 1-year stability rates that ranged from 86% to 95%. The demographic characteristics of the schools are similar to the demographic characteristics of public elementary schools in New York City at this time.

Consent and participants. Consent packages (in English and Spanish) were sent home to all parents of third-grade children in the 18 schools informing them of the study and seeking consent for their child to participate. Consent rates were calculated at the end of Year 1 and included the children consented by the spring of third grade ($n = 883$) divided by the total eligible population at spring of third grade ($n = 1,420$; see Figure 1), resulting in a rate

of 62.18%. These rates did not differ between intervention schools (63.59%) and control schools (60.57%). Nonparticipants included children whose parents did not speak English or Spanish well enough to consent to participate and children with special needs who could not be interviewed on an individual basis. Comparisons of demographic and achievement characteristics between consented and nonconsented children were conducted with school records data. Although significant differences were found for gender (fewer male participants, Cohen's $d = 0.15$) and school absences (higher absences for nonparticipants, Cohen's $d = 0.21$), they were of small size, and no differences were found for children's race/ethnic status, receipt of free lunch, number of suspensions, and reading and math test scale scores. On the basis of these results, we are confident that we obtained a representative sample of children from the 18 schools.

Parents who granted consent completed a demographic questionnaire at home, rated their children's social competence and externalizing problems, and returned the questionnaires in sealed envelopes to the school, where they were collected by our research team. Parents who consented to participate but did not return a questionnaire were contacted via telephone to complete the survey. Parents were paid \$15 for completion of the surveys at each

assessment. Teachers completed questionnaires rating the language and literacy skills, social competence, and externalizing problems of each child in their class with consent. Teachers were paid at the union wage of \$36.50 per hour for their time. Children completed questionnaires rating their social cognitions and internalizing symptoms. Data were collected from the children in small class groups ($n = 5$ to 20). All questions were read out loud by a research assistant while a second research assistant circulated to monitor children's responses and answer questions. Children who did not have consent or who did not assent worked on an alternative activity.

The total sample includes 942 third-grade children (49% boys; average age = 8.07 years, $SD = 0.7$), 799 parents (85% mothers; average age = 35.5 years), and 85 teachers (94% female; average age = 36.1 years). Children and teachers were nested in nine intervention schools ($n = 515$; 54.7%) and nine control schools ($n = 427$; 45.3%). Data for this report come from Year 1 of the 3-year evaluation. Baseline data were gathered in the fall of third grade (2004, Wave 1), and follow-up data were collected in the spring of third grade (2005, Wave 2). Because randomization occurred at the school level, children who moved out of a participating school were not followed ($n = 59$), and consent was requested for children who either moved into a participating school or decided to join the study in the spring of third grade ($n = 124$) (Vuchinich et al., 2010). Attrition from fall to spring was minimal

(7.2%) and primarily due to student mobility out of the participating schools. Refusals from parents to continue participation were rare ($n = 1$; 0.1%). As shown in Figure 1, children who were present in both waves of the study ($n = 759$) and children who were present in either wave but not both ($n = 59$ fall only, $n = 124$ spring only) were included in the analyses presented below ($N = 942$). There is no evidence that children with complete data on the outcomes of focus in this paper (i.e., at both Waves 1 and 2) differed significantly from those with data only at Wave 1 or only at Wave 2 (SACD Research Consortium, 2008).

According to parent reports at baseline, 53.4% ($n = 503$) of children lived in a single-parent household, 15.1% ($n = 142$) of parents were unemployed, 31% ($n = 292$) of parents had less than a high school diploma or GED, and 61.8% ($n = 582$) of households were at or below 100% of the federal poverty level. Children represented diverse racial/ethnic groups; 45.6% ($n = 430$) were Hispanic/Latino, 41.1% ($n = 387$) were Black/African American, 4.7% ($n = 44$) were non-Hispanic White, and 8.6% ($n = 81$) were other racial/ethnic groups (e.g., Asian, Pacific Islander, Native American). Table 1 presents baseline demographic characteristics.

Intervention. The 4Rs Program is a universal, school-based intervention in social-emotional learning and literacy development that integrates a focus on broad social and emotional development into the language arts curriculum for children in Grades K–5. Developed and run by a community-based nonprofit organi-

Table 1
Sample Baseline Demographic Characteristics by Intervention and Control Schools

Demographic characteristic	α (where applicable)	Intervention ($n = 515$)	Control ($n = 427$)	Total ($N = 942$)
Child age				
Age in years, M (SD)		8.08 (0.70)	8.06 (0.65)	8.07 (0.68)
Child gender				
Boys		250 (48.5%)	210 (49.2%)	460 (48.8%)
Girls		265 (51.5%)	217 (50.8%)	482 (51.2%)
Child race/ethnicity				
Non-Hispanic White		24 (4.7%)	20 (4.7%)	44 (4.7%)
Hispanic/Latino		231 (44.9%)	199 (46.6%)	430 (45.6%)
Black/African American		219 (42.5%)	18 (39.3%)	387 (41.1%)
Other		41 (8.0%)	40 (9.4%)	81 (8.6%)
Household SES risks				
Low parental education		154 (29.9%)	138 (32.3%)	292 (31%)
Unemployment		68 (13.2%)	74 (17.3%)	142 (15.1%)
Single-parent household		289 (56.1%)	214 (50.1%)	503 (53.4%)
Poverty		313 (60.8%)	269 (63.0%)	582 (61.8%)
Household SES risk index				
No risks		102 (19.8%)	76 (17.8%)	178 (18.9%)
One risk		153 (29.7%)	120 (28.1%)	273 (29.0%)
Two risks		133 (25.8%)	141 (33.0%)	274 (29.1%)
Three risks		103 (20.0%)	67 (15.7%)	170 (18.0%)
Four risks		24 (4.7%)	23 (5.4%)	47 (5.0%)
SES risk, M (SD)		1.60 (1.15)	1.63 (1.11)	1.61 (1.13)
Community risk	Year 1 = .88, Year 2 = .88			
Community risk, M (SD)		2.06 (0.77) ^a	1.96 (0.68)	2.02 (0.73)
Child baseline aggression index	A ^b : Year 1 = .95, Year 2 = .96 C ^b : Year 1 = .75, Year 2 = .83			
No risks		431 (83.7%)	379 (88.8%)	810 (86.0%)
One risk		51 (9.9%)	33 (7.7%)	84 (8.9%)
Two risks		33 (6.4%)	15 (3.5%)	48 (5.1%)
Behavioral risk, M (SD)		0.23 (0.55) ^a	0.15 (0.44)	0.19 (0.51)

Note. SD = standard deviation; SES = socioeconomic status.

^a Baseline characteristic significantly higher for intervention than control group: Community risk, $t(940) = 2.20$, $p < .05$; baseline aggression, $t(940) = 2.41$, $p < .05$. ^b A = Aggression subscale, C = Conduct subscale.

zation (MCTSR; www.morningsidecenter.org), the 4Rs Program uses high-quality children's literature as a springboard for helping students gain skills and understanding in the areas of handling anger, listening, assertiveness, cooperation, negotiation, mediation, building community, celebrating differences, and countering bias. Broadly speaking, the social-emotional learning core of 4Rs is designed to change how children think, feel, and behave in situations of potential interpersonal conflict (see examples below). Our evaluation's emphasis on the set of social-cognitive mental processes that underlie aggressive and other problem behaviors is consistent with this program theory.

The program has two primary components: (a) a comprehensive 7-unit, 21–35 lesson, literacy-based curriculum in social-emotional learning and (b) 25 hrs of training followed by ongoing coaching of teachers to support them in teaching the 4Rs curriculum (with a minimum of 12 coaching contacts per year). Each curriculum unit focuses on a specific grade-appropriate children's book and begins with a book reading and discussion, ensuring that students understand the primary themes of the story. This is followed by three to five social-emotional learning skill lessons in which children practice specific skills in the context of a discussion of the book. For example, the lessons for the third-grade unit "Understanding and Dealing with Feelings" focus first on identifying feeling words from the story, then on practicing "reading" feelings other students act out, and finally on identifying and practicing strategies for "cooling down" through role play. Consistent with this unit's emphasis on the social-cognitive processes of detecting situational and emotional cues, the "Listening" unit emphasizes basic skills of effective listening including making direct eye contact, paraphrasing, and acknowledging comprehension.

Intensive professional development for teachers consists of a 25-hr introductory training course and ongoing classroom coaching by a 4Rs staff developer from MCTSR. Teachers are provided with learning kits with a full set of materials needed to implement the program. The introductory training is designed to (a) introduce the teachers to the curriculum, (b) give them an opportunity to practice social-emotional skills at the adult level through role playing and experiential learning, and (c) inspire them to employ the ideas and skills embodied in the curriculum in their own lives both professionally and personally. Ongoing classroom coaching encompasses class lesson modeling, workshops, and group meetings led by the staff developer, co-planning and teaching of lessons by the teacher and staff developer, and lesson observations and feedback. A full cost study for this evaluation is under way but not complete; however, MCTSR reports the cost of the 4Rs Program to be approximately \$90 per child per year (this amount excludes all costs associated with the evaluation).

4Rs implementation. Schoolwide implementation of the two primary components of the 4Rs Program was systematically tracked and monitored. During the first year of implementation, teachers on average delivered three quarters of a lesson in the 4Rs curriculum per week ($SD = 0.70$), with the majority closer to the benchmark of one lesson per week. Further, the majority of teachers appear to have spent on average between 20 and 25 total hours (~ 40 min per week) implementing the 4Rs curriculum throughout the year ($SD = \sim 32$ min per week). Teachers in the nine intervention schools received on average 2.4 ($SD = 0.33$) days of training in the delivery of the 4Rs curriculum, and schools received

an average of 38 ($SD = 9.6$) days per school of coaching by 4Rs staff developers. There is more variability in 4Rs implementation between teachers than between schools. There were no significant differences in average levels of implementation (lessons delivered per week, time spent per week, and the amount of training and coaching received) between teachers in different grades.

Information on the implementation of various social and character development activities was also gathered from all third-, fourth-, and fifth-grade teachers in both intervention and control schools with a measure developed for the broader SADC research program. The percentage of control teachers reporting using any SADC activities in their classroom ranged from 75.6% to 85.6% over the 3 years of the study. A greater percentage of intervention than control teachers reported using any SADC activity in Year 1 (Cohen's $d = 0.31$), as well as specific activities targeting violence prevention/peace promotion (Cohen's $d = 0.50$) and social and emotional development (Cohen's $d = 0.52$; SADC Research Consortium, 2008).

Measures

Scale scores, computed as the mean across the items for each construct, were employed in the analyses presented below. Basic psychometrics, including alpha coefficients for each construct in the fall and spring of third grade, are presented in Table 2.

Social-cognitive processes. *Hostile attribution biases* were measured with child self-reports on an adaptation of the Home Interview Questionnaire (Dalhberg, Toal, & Behrens, 1998). Six vignettes that depict ambiguous but provocative social scenarios are read aloud while pictorial representations of the scenarios are presented (e.g., a student's milk carton is spilled on another child's back). Following presentation of each vignette, children are asked to select one of four possible causal attributions regarding the intent of the provocateur. Two attributions refer to the provocateur's intent as benign or accidental (rated 0; e.g., the milk was spilled accidentally) and two refer to it as hostile or purposeful (rated 1; e.g., the student was being mean).

Normative beliefs about aggression were measured with the Normative Beliefs About Aggression scale (Huesmann & Guerra, 1997), a measure designed to assess children's beliefs about the acceptability of the use of aggression in certain situations. The total normative beliefs scale ranges from 1 (low normative beliefs; the use of aggression is "perfectly OK") to 4 (high normative beliefs; the use of aggression is "really wrong") and includes 12 items.

Aggressive and prosocial fantasies were measured with self-reports on the What I Think instrument (Rosenfeld, Huesmann, Eron, & Torney-Purta, 1982). The Aggressive Fantasies subscale contains six items that ask children about aggressive thoughts that just "pop into your head" or "daydreams" (e.g., about hitting or hurting someone; pretending to fight with someone). The Prosocial subscale contains six items that ask children about prosocial thoughts (e.g., about helping other kids, about doing nice things for other kids). Items are rated on a 3-point scale (0 = no, 1 = a little, 2 = a lot).

Aggressive interpersonal negotiation strategies were also self-reported with an adaptation of the Home Interview Questionnaire (Dalhberg et al., 1998). Following the presentation of each vignette and causal attributions, children are asked what they would do next

Table 2
Psychometric and Descriptive Characteristics of Dependent Variables at Wave 1 (Baseline) and Wave 2 by Intervention and Control Schools

Variable	α	Scale range	Intervention ($n = 515$)	Control ($n = 427$)	Total ($n = 942$)
Hostile attribution biases (CR)					
W1	.76	0–1	0.35 (.32)	0.38 (.32)	0.36 (.32)
W2	.79	0–1	0.36 (.32)	0.42 (.35)	0.39 (.34)
Aggressive interpersonal negotiation strategies (CR)					
W1	.89	0–1	0.17 (.30)	0.21 (.33)	0.19 (.31)
W2	.90	0–1	0.23 (.33)	0.26 (.37)	0.24 (.35)
Normative beliefs about aggression (CR)					
W1	.85	1–4	1.29 (0.46)	1.35 (0.56)	1.32 (0.51)
W2	.90	1–4	1.42 (0.63)	1.46 (0.67)	1.44 (0.64)
Aggressive fantasies (CR)					
W1	.59	0–2	0.64 (.44)	0.70 (.46)	0.67 (.45)
W2	.65	0–2	0.72 (.46)	0.76 (.46)	0.74 (.46)
Prosocial fantasies (CR)					
W1	.71	0–2	1.23 (.50)	1.22 (.51)	1.23 (.50)
W2	.71	0–2	1.19 (.49)	1.17 (.50)	1.18 (.50)
Depressive symptoms (CR)					
W1	.51	0–1	0.49 (.26)	0.48 (.27)	0.49 (.26)
W2	.49	0–1	0.47 (.26)	0.52 (.26)	0.49 (.26)
Attention/hyperactivity symptoms (TR)					
W1	.91	1–4	1.80 (0.70)	1.77 (0.65)	1.78 (0.68)
W2	.92	1–4	1.86 (0.73)	1.85 (0.68)	1.85 (0.71)
Aggression (TR)					
W1	.95	1–4	1.55 (0.64)	1.47 (0.54)	1.51 (0.59)
W2	.96	1–4	1.68 (0.70)	1.61 (0.62)	1.65 (0.66)
Social competence (TR)					
W1	.97	1–4	2.78 (0.73)	2.82 (0.73)	2.80 (0.73)
W2	.97	1–4	2.74 (0.74)	2.74 (0.74)	2.74 (0.74)
Academic skills (TR)					
W1	.97	1–5	3.07 (1.05)	3.04 (1.08)	3.06 (1.07)
W2	.97	1–5	3.33 (1.06)	3.43 (1.07)	3.37 (1.07)
Math achievement (SR)					
W2		385–740	608.03 (42.79)	613.69 (40.55)	610.58 (41.86)
Reading achievement (SR)					
W2		480–750	616.13 (38.48)	624.45 (33.27)	619.88 (36.44)
Attendance rate % (SR)					
Prior year		0–100	90.96 (8.86)	91.58 (7.13)	91.22 (8.15)
W2		39–100	90.99 (7.99)	91.54 (7.39)	91.24 (7.72)

Note. CR = child report; W1 = Wave 1, fall third grade; W2 = Wave 2, spring third grade; TR = teacher report; SR = school records.

in the scenario and are asked to select one of four response strategies. Three response strategies are nonaggressive or benign (rated 0; e.g., act like nothing happened), and one strategy is aggressive or hostile (rated 1; e.g., spill paint on the child).

Behavioral symptomatology. *Child aggression* was measured with teacher reports of aggression on the Behavioral Assessment System for Children (Reynolds & Kamphaus, 1998). Teachers report on 13 aggressive behavior items (e.g., physically aggressive, argumentative, threatening, or critical of others) observed within the past 30 days, and items are rated on a 4-point scale (1 = *never*, 2 = *sometimes*, 3 = *often*, 4 = *almost always*).

Child social competence was measured with teacher reports on the 18-item Social Competence Scale (CPPRG, 1999). The measure includes two subscales: Prosocial Behaviors (e.g., is good at understanding other people's feelings) and Emotion Regulation (e.g., can calm down when excited or all wound up).

Depressive symptoms were assessed via self-reports of depressive symptoms on the Diagnostic Interview Schedule for Children Predictive Scales (Lucas et al., 2001). This questionnaire contains six items that ask children about whether they have experienced

particular depressive symptoms (e.g., lethargy, anhedonia, suicidal thoughts) in the past year (for Wave 1) or since the new year (for Wave 2). Items are rated on a 2-point scale (0 = *no*, 1 = *yes*).

Attention/hyperactivity symptoms were measured from teacher reports on nine items from the ADHD Symptomatology Scale (Milch, Loney, & Landau, 1982). Teachers report on children's hyperactivity (e.g., is excitable or impulsive) and inattentiveness (e.g., has difficulty organizing tasks or activities) within the past 30 days, with items rated on a 4-point scale (1 = *never* to 4 = *almost always*).

Literacy skills and academic achievement. *Academic skills* were measured from teacher reports on items adapted from the Early Childhood Longitudinal Study, Kindergarten Cohort of 1998–1999, third-grade assessment (National Center for Education Statistics, 2002). Nine items tap the degree to which a child has acquired and demonstrates the targeted language and literacy skills, knowledge, and behaviors appropriate for third graders (e.g., reads fluently, conveys ideas clearly, composes multiparagraph stories). Items are rated on a 5-point scale (1 = *not yet* to 5 = *proficient*).

Standardized math and reading achievement were measured with children's scaled scores on the New York State standardized assessments of math and reading achievement at the end of third grade for the 2004–2005 school year. These tests were not administered the prior year when children were in second grade.

Attendance rate was calculated as the proportion of full days present in school during the year prior to the study (2003–2004) and children's third-grade school year (2004–2005).

Baseline covariates. *Household socioeconomic (SES) risk index* was calculated as the sum of four parent-reported, dichotomous demographic characteristics at baseline: single-parent household, less than high school education, poverty at or below 100% of the federal poverty level, and unemployment. For children with missing parent-reported risk indicators, these values were imputed based on child school, gender, and race/ethnicity. Among the parents who consented to participate at each assessment point (85.7%, 80.96% of the total child sample at Waves 1 and 2, respectively), 13.4% ($n = 94$) and 20.3% ($n = 152$) of parent-report data were missing at Waves 1 and 2.

Community risk was measured from parent reports on the Community Risks and Resources Questionnaire (Forehand et al., 2000). Parents are asked to rate how well seven statements describe their neighborhood (e.g., drugs are sold or used, houses/apartments are in poor condition). Items are rated on a 4-point scale (1 = *not at all* to 4 = *lot*), and a high score indicates greater risk. As with household SES risk, for children with missing parent-reported community risk indicators these values were imputed based on child school, gender, and race/ethnicity.

Child aggression index was measured from teacher reports of aggression and conduct problems on the Behavioral Assessment System for Children (Reynolds & Kamphaus, 1998). Teachers report on 13 aggressive behaviors (e.g., physically aggressive, argumentative) and 11 conduct-disordered behaviors (e.g., steals, truancy) observed within the past 30 days, with items rated on a 4-point scale (1 = *never* to 4 = *almost always*). The index was calculated based on a nationally normed t score ($t = 63.5$ and 62.9 for aggressive behavior and conduct disorder, respectively). Children were grouped based on whether they were below the t -score cutoff on both scales (rated 0), at or above the cutoff on one scale (rated 1), or above the cutoff on both scales (rated 2).

Results

Means and standard deviations for the 13 focal child outcomes at Waves 1 and 2 are presented for the sample overall and by intervention group (see Table 2). There were no differences in mean levels of the outcomes at baseline (Wave 1) by intervention group. Intercorrelations among the dependent variables at baseline as well as stability in the dependent variables from Wave 1 (baseline) to Wave 2 are available from the first author upon request.

Main analytic strategy. Because schools were matched into pairs prior to randomization, the data presented here are nested: Children are nested in schools, and schools are nested in their matched pairs. To accommodate these design features, we calculated estimates of intervention impact on change in the primary child outcomes from preintervention baseline (fall 2004, Wave 1) to the first follow-up (spring 2005, Wave 2) using a series of two-level hierarchical linear models with random effects in HLM 6.02. In these models, Level 1 (child) included the preintervention

baseline of the dependent variable (with the exception of math and reading achievement, for which no prior year score was available), and the child-level baseline covariates (child gender, race/ethnic background, family SES risk, community risk, and child aggression index), and Level 2 (school) included a dummy variable indicating intervention condition as well as eight dummy variables representing the school pair matches. In addition, interactions between intervention status and the baseline child aggression index were examined in a second set of models. In interpreting the results we consider an alpha level of $p < .05$ as significant but note as trends effects up to the .10 level in the case of interactions (McClelland & Judd, 1993). In addition, for main effects and any significant interactions, we report effect sizes (e.g., McCartney & Rosenthal, 2000). Effect sizes were calculated by dividing the estimate of the intervention effect by the raw standard deviation of the dependent variable at Wave 2 for the control group (a variant of Cohen's d , attributed as Glass's Δ , designated in the results as ES; Cohen, 1992).

Main effects of intervention. As shown in Tables 3 and 4, there were significant main effects of the 4Rs intervention for two of the nine child outcomes in the social-cognitive processes and behavioral symptomatology domains. We found intervention effects for children's self-reports of hostile attributional bias (ES = .20), such that children in the intervention group reported significantly lower levels of hostile attributional bias in spring 2005 (Wave 2) than did those in the control group after we accounted for their baseline scores in fall 2004. In addition, we found intervention effects for children's self-reports of depressive symptoms (ES = .24), such that children in the intervention group reported significantly lower levels of depressive symptoms in spring 2005 (Wave 2) than did those in the control group after we accounted for their baseline scores in fall 2004. There were no main effects for the remaining outcomes in these domains after the first year of exposure to the 4Rs intervention. In addition, there were no main effects for the four outcomes in the literacy skills and academic achievement domain after the first year of exposure to the 4Rs intervention (see Table 4).

Interaction of intervention status with baseline child aggression. As indicated above, cross-level interactions of intervention status by the baseline child aggression index were included in a second set of models to address our second question about the degree to which the effects of the 4Rs intervention were moderated by baseline levels of child aggression (Rothwell, 2005). As shown in Tables 3 and 4, there were statistically significant interactions for two of the 13 outcomes examined (teacher report of academic skills and attendance), and there were two additional interactions with p values less than .10 (aggressive fantasies and reading achievement). To interpret the statistically significant interactions, we generated plots of the adjusted means for the relevant outcomes for each level of the baseline child aggression index by intervention group. We also calculated effect sizes for the mean differences between the intervention and control groups at each level of the baseline child aggression index. As shown in Figure 2, children with the highest level of baseline behavioral risk (based on teacher reports) showed the greatest positive difference in teacher report of academic skills between the intervention and control schools, with effect sizes of .40 for a score of 2 on the baseline child aggression index, .17 for a score of 1, and $-.05$ for a score of 0. As shown in Figure 3, children's attendance rate followed a similar pattern,

Table 3
Unstandardized Estimates (and Standard Errors) of 4R's Impact on Children's Social-Cognitive Processes in Spring of Third Grade

Model variable	Social-cognitive processes				
	HAB (CR)	AINS (CR)	NOBAGS (CR)	AGGFANT (CR)	PROFANT (CR)
Intercept (8 <i>df</i>)	0.19* (0.07)	0.18* (0.08)	1.02** (0.22)	0.60*** (0.12)	0.54** (0.12)
Intervention status (8 <i>df</i>)	-0.07* (0.02)	0.00 (0.04)	0.04 (0.07)	0.02 (0.05)	-0.01 (0.05)
[CI for intervention effect] and effect size	[-0.03, -0.11] .20	[0.08, -0.08] .00	[0.18, -0.10] .06	[0.12, -0.08] .04	[0.07, -0.09] .02
DV at baseline (17 <i>df</i>)	0.55*** (0.04)	0.57*** (0.04)	0.44*** (0.04)	0.45*** (0.04)	0.45*** (0.03)
Child gender, 1 = girl (17 <i>df</i>)	-0.02 (0.03)	-0.09** (0.03)	-0.22*** (0.05)	-0.19*** (0.04)	0.08 (0.04)
Child Hispanic vs. White (17 <i>df</i>)	-0.01 (0.06)	0.01 (0.05)	-0.11 (0.14)	-0.03 (0.09)	0.03 (0.08)
Child Black vs. White (17 <i>df</i>)	0.02 (0.07)	0.06 (0.06)	-0.05 (0.15)	0.13 (0.10)	-0.00 (0.09)
Child Other vs. White (17 <i>df</i>)	-0.01 (0.06)	-0.08 (0.07)	-0.21 (0.16)	-0.12 (0.11)	-0.02 (0.10)
SES risk (17 <i>df</i>)	0.01 (0.01)	0.00 (0.01)	0.02 (0.02)	-0.00 (0.02)	0.01 (0.02)
Community risk (17 <i>df</i>)	0.00 (0.02)	0.01 (0.02)	-0.00 (0.05)	-0.01 (0.03)	-0.00 (0.03)
Baseline child aggression index (17 <i>df</i>)	-0.02 (0.04)	-0.01 (0.04)	-0.05 (0.10)	0.12 (0.08)	0.07 (0.06)
Intervention of intervention status with levels of aggression at baseline (16 <i>df</i>)	0.04 (0.05)	0.02 (0.05)	-0.03 (0.14)	-0.17 [†] (0.10)	-0.06 (0.07)

Note. Eight dummy variables representing eight of the nine school-level matched pairs are included in all models at the school level (Level 2), with pair 9 serving as the referent group. The interaction term is added in a second model. HAB = hostile attribution biases; AINS = aggressive interpersonal negotiation strategies; NOBAGS = normative beliefs about aggression; AGGFANT = aggressive fantasies; PROFANT = prosocial fantasies; *df* = degrees of freedom; CI = confidence interval.

* $p < .05$. ** $p < .01$. *** $p < .001$. [†] $p < .10$ (interactions only).

with children in the highest risk group showing the largest positive impact of the intervention (ES = .32 for a score of 2 on behavioral risk, .14 for a score of 1, and -.10 for a score of 0).

Discussion

This paper presents experimental impacts after 1 school year of a universal, integrated intervention in ethnically diverse, urban elementary schools that embeds instruction in social-emotional learning into a balanced literacy curriculum. It is unique in that it provides an experimental test of the short-term causal impact of a universal integrated intervention on key mental processes as well as behavioral and academic outcomes. As described in the introduction, building on program and developmental cascades theories and on our prior work, we posited both short-term and longer term effects of the 4Rs: (a) short-term effects on children's social-cognitive processes and literacy skills, (b) potential early cascading or spillover effects on depression and attention, (c) variation in the strength of effects by children's baseline levels of aggression, and (d) longer term effects on aggression, social competence, and achievement.

With regard to the results after 1 school year, we did not see strong evidence of short-term change in children's social-emotional, behavioral, and academic functioning. We found (a) a single statistically significant main effect on children's hostile attributional biases out of five social-cognitive processes examined and no main effect for children's literacy skills; (b) a statistically significant main effect on children's self reports of depression but not on attention problems; (c) statistically significant cross-level interaction effects by baseline child aggression for two outcomes (academic skills and attendance) and trend-level effects for two others (standardized tests of reading and aggressive fantasies); and (d) no main effects for aggression, social competence, and the measures of academic achievement. Thus, we found statistically significant main effects for only two outcomes out of the

13 examined, a rate better than chance but only slightly. The same rate holds for the interaction effects.

Of the statistically significant main effects that were identified, hostile attributional bias is a core social-cognitive process considered central to the 4Rs Program theory and an outcome for which we expected to see short-term change. Children's self-reports of depressive symptoms is an outcome for which we expect longer term spillover effects, perhaps after more years of exposure to the program. The size of the effects on these two outcomes (ES = .20 and .24, respectively) is generally consistent with findings after the first year of our evaluation of the RCCP (as noted above, the RCCP lessons represent the majority of the social-emotional learning core of the 4Rs Program). Despite their absolute size, these findings were equivalent to offsetting approximately one year of decline in the social-cognitive and social-emotional processes tied to later behavioral outcomes (Aber et al., 1998).

We expect, consistent with the theory underlying the 4Rs Program, the design of the evaluation, and our previous work and the work of others (e.g., Aber et al., 1998, 2003; Lansford et al., 2006), that exposure to the 4Rs Program will first influence the most proximal targets (social-cognitive processes), and these in turn will influence the developmental course of a broader set of mental, emotional, behavioral, and academic outcomes. In particular, we expect that the early changes we see for 4Rs children in hostile attributional biases will translate into changes in aggressive and delinquent behavior in later elementary school, when these more serious behaviors begin to increase (e.g., Smolkowski et al., 2005). In addition, consistent with developmental cascades theory (Masten & Cicchetti, 2010), some of the unanticipated early effects we are observing in depressive symptoms might be expected to translate into universal benefits in other, related developmental domains later in elementary school (reduced academic disengagement/failure and delayed onset of substance use; e.g., Eddy, Reid, Stoolmiller, & Fetrow, 2003; Lochman & Wells, 2004).

Table 4
Unstandardized Estimates (and Standard Errors) of 4R's Impact on Children's Behavioral Symptomatology, Literacy Skills, and Academic Achievement in Spring of Third Grade

Model variable	Behavioral symptomatology				Literacy skills and academic achievement			
	DEP (CR)	ATTENT (TR)	AGG (TR)	S-COMP (TR)	AC SKILL (TR)	MATH (SR)	READ (SR)	ATTEND (SR)
Intercept (8 df)	0.34*** (0.06)	0.48* (0.16)	0.22 (0.12)	0.71* (0.22)	0.86* (0.28)	631.58*** (13.17)	651.31*** (9.92)	96.51*** (1.60)
Intervention status (8 df)	-0.06* (0.02)	0.01 (0.06)	0.02 (0.05)	0.09 (0.08)	-0.02 (0.08)	-7.40 (5.69)	-5.35 (4.96)	-0.86 (0.66)
[CI for interaction effect] and effect size	[-0.02, -0.10], .24	[0.13, -0.11], .01	[0.12, -0.08], .03	[0.25, -0.07], .12	[0.14, -0.18], .02	[3.7, -18.5], .12	[4.3, -15.1], .18	[0.43, -2.1], .16
DV at baseline (17 df)	0.30*** (0.04)	0.76*** (0.05)	0.95*** (0.05)	0.73*** (0.04)	0.78*** (0.03)			
Child gender, 1 = girl (17 df)	0.05* (0.02)	-0.13* (0.05)	-0.07 (0.03)	0.12* (0.04)	0.09 (0.05)	2.62 (3.31)	12.40*** (2.99)	0.50 (0.67)
Child Hispanic vs. White (17 df)	0.08 (0.05)	0.04 (0.10)	0.00 (0.08)	-0.08 (0.08)	-0.02 (0.13)	1.22 (9.48)	-5.16 (6.41)	-0.91 (1.35)
Child Black vs. White (17 df)	0.10 (0.05)	0.14 (0.10)	0.11 (0.08)	-0.15 (0.09)	-0.09 (0.13)	-0.83 (9.25)	-7.81 (6.37)	-0.33 (1.37)
Child Other vs. White (17 df)	-0.01 (0.06)	-0.03 (0.10)	-0.06 (0.09)	-0.07 (0.10)	0.02 (0.14)	14.14 (9.22)	0.56 (7.36)	1.99 (1.42)
SES risk (17 df)	-0.01 (0.01)	0.02 (0.02)	0.02 (0.02)	0.01 (0.03)	-0.02 (0.03)	-5.32** (1.65)	-4.67** (1.31)	-1.08** (0.26)
Community risk (17 df)	0.00 (0.01)	0.02 (0.03)	0.01 (0.03)	-0.00 (0.03)	0.03 (0.04)	-1.08 (2.27)	-0.94 (2.09)	-0.65 (0.48)
Baseline child aggression index (17 df)	0.01 (0.03)	0.04 (0.08)	-0.15 (0.09)	-0.12 (0.06)	-0.15 (0.09)	-10.71* (4.86)	-12.07** (4.02)	-3.31** (0.55)
Interaction of intervention status with levels of aggression at baseline (16 df)	0.03 (0.04)	-0.06 (0.10)	-0.03 (0.11)	0.03 (0.07)	0.21* (0.09)	4.77 (5.97)	8.89 (4.99) [†]	2.14* (0.96)

Note. Eight dummy variables representing eight of the nine school-level matched pairs are included in all models at the school level (Level 2), with pair 9 serving as the referent group. The interaction term is added in a second model. DEP = depression; ATTENT = attention/hyperactivity symptoms; AGG = aggression; S-COMP = social competence; AC SKILL = academic skills; MATH = math achievement; READ = reading achievement; ATTEND = attendance rate; CR = child report; TR = teacher report; SR = school records; df = degrees of freedom; CI = confidence interval; SES = socioeconomic status. * $p < .05$. ** $p < .01$. *** $p < .001$. [†] $p < .10$ (interactions only).

With regard to the interaction effects, the effect sizes for the subgroup of children with the highest scores on baseline aggression range from .59 for aggressive fantasies to .32 for attendance. These findings are consistent with other experimental evaluations of school-based interventions that target primarily children's externalizing behavior problems. For example, the most recent report of results from the evaluation of Fast Track showed significant effects for diagnoses and behavior in the externalizing domain for children after Grades 3, 6, and 9 but only for children identified at kindergarten at highest behavioral risk (using parent and teacher reports; CPPRG, 2007). In addition, Van Lier, Muthen, van der Sar, and Crijnen (2004) reported positive impacts of the Good Behavior Game on trajectories of conduct problems from Grades 1 to 3 for children with the highest levels of disruptive behavior at baseline (Fall, first grade), with only somewhat larger effects than those reported here (ES = .55).

Each of these studies reported the impact of interventions that target one domain of functioning (i.e., externalizing behavior problems) on outcomes in the same domain for children at highest risk at baseline. The present study, in contrast, reports significant experimental impacts of a universal, integrated, school-based social-emotional learning and literacy intervention for a subgroup of children identified based on functioning in one developmental domain on outcomes in another developmental domain. In other words, here we report effects of this integrated intervention on children identified at baseline as having serious behavioral difficulties on outcomes in the academic domain. The effects of 4Rs for children with the highest scores on the baseline aggression index on outcomes in the academic domain ranged in size from .31 (reading achievement) to .40 for teacher reports of academic skills, which represents approximately one third of a year of regular gain in reading for children of similar age (Hill, Bloom, Black, & Lipsey, 2008).

We are still several years away from being able to report on the full developmental cascade effects of the 4Rs intervention, but this report of effects on individual children across domains of functioning after 1 year represents an important, and we believe valuable, first step in establishing a better understanding of what is achievable by a schoolwide intervention, such as the 4Rs, in its earliest stages of unfolding. This is especially true in light of the relation of these results after 1 year to the results after 2 years of intervention. We now know that, after 2 years of 4Rs implementation, the positive main effects reported here regarding children's hostile attributional bias and depressive symptoms are not only maintained but have expanded to three additional outcomes for which we expect longer term or cascading effects: teacher reports of children's aggressive and socially competent behavior and children's attention problems. These results provide initial support for the developmental cascades concept. In addition, interaction effects by baseline child aggression persist for standardized reading and are extended to math scores (Jones et al., in press).

Although the present study has several strengths, our conclusions are tempered by several limitations. First, our ability to draw definitive conclusions from this set of findings is limited, given that we examined 13 outcomes and found main effects on only two outcomes and interaction effects on four other outcomes. Furthermore, there was only one significant effect in the social-cognitive processes domain identified as highly proximal to the intervention's program theory. Together with our 2-year results (Jones et

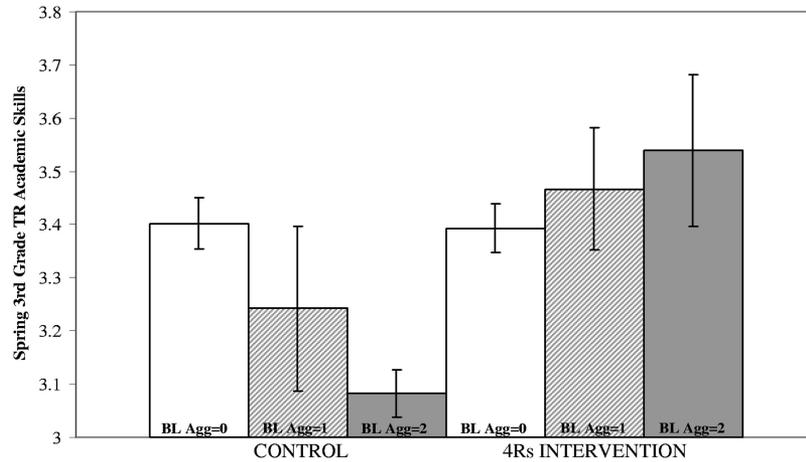


Figure 2. Intervention group and baseline child aggression plotted against Wave 2 (spring 2005) teacher reports of children’s academic skills. TR = teacher report; BL Agg = baseline child aggression index. Error bars represent standard errors.

al., in press), these findings suggest that aspects of children’s social–cognitive processes and mental health outcomes may themselves operate in a cascading manner requiring longer term, sustained intervention.

A second factor to consider is program implementation. As described in the Method section, there was variability in implementation of the 4Rs Program and the quantity of program implementation was below benchmarks, with teachers on average implementing 4Rs lessons for 40 min (rather than 60 min) per week and completing on average three quarters of a lesson (instead of one). In addition, substantial program-like activity in the broad domain of social and character development was reported in control schools. This is not surprising, given the dramatic growth in interest in social–emotional learning in the last decade (Payton et al., 2008). Based on these data, intervention schools still reported implementing significantly more social and character activities,

particularly and not surprisingly activities focused on violence prevention and social and emotional development (we assume with better training and more ongoing support). These implementation data underscore the need to examine intervention effects in a manner that accounts both for dosage and for implementation of similar forms of intervention activities in control schools using sophisticated methodological tools (e.g., propensity scores; Lochman, Boxmeyer, Powell, Roth, & Windle, 2006).

There are a number of additional limitations to this study, including the relatively low consent rate (~62%) and the rate of attrition between baseline and follow-up (7.2%). It is important to note, however, that these rates did not differ significantly by intervention group and that other sites participating in the Social and Character Development Research Network reported similar experiences. Finally, some of our measures had relatively low reliability (e.g., aggressive fantasies and depressive symptoms),

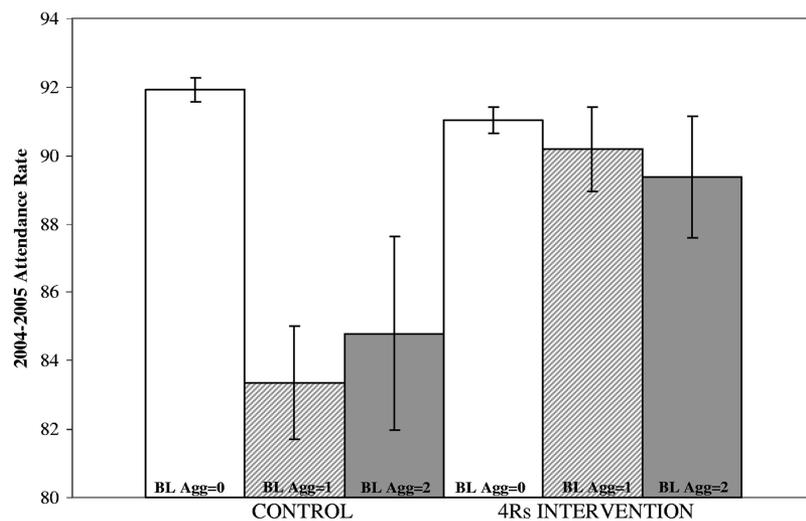


Figure 3. Intervention group and baseline child aggression plotted against Wave 2 (spring 2005) children’s attendance rate. BL Agg = baseline child aggression index. Error bars represent standard errors.

and this potentially undermines precision in our intervention estimates due to measurement error.

Despite these limitations, this report of first-year impacts of an integrated, social-emotional and literacy intervention, combined with our knowledge of sustained effects on hostile attributional bias and depression as well as expanded effects after a second year on aggression, social competence, and attention problems (Jones et al., in press), provides evidence that this intervention may be initiating a positive developmental cascade both in the general population of students and among those at highest behavioral risk. After 1 school year, 4Rs had positive effects for our population of children on two outcomes and subgroup effects in the academic domain for children identified as aggressive at the outset of the study. In addition, these effects appear to expand to other outcomes after 2 years. Together, the findings underscore the potential clinical significance of the 4Rs Program in two important ways. First, clinical treatments for mental health problems are typically costly and can be prohibitively so for low-income and underserved populations. The body of findings emerging from our research suggests the 4Rs Program may prove to be a low-cost preventive intervention strategy that could reduce known social-cognitive, social-emotional, and behavioral risks for long-term mental health outcomes among the general population of students. Relative to other school-based social-emotional learning programs, the 4Rs carries a comparable or lower price per student (Foster, Jones, & Conduct Problems Prevention Research Group, 2006). An important cautionary note, however, is that despite its relatively low price, the 4Rs Program does require a commitment and investment of time and resources that could limit its generalizability and replicability outside an evaluation study such as this.

Second, given the mounting evidence linking improved academic performance with social competencies, attentional capacities, and lower externalizing behavior, our findings suggest that integrating social-emotional learning with literacy has the potential to address academic and mental health constraints on learning as well as the broader adaptation of children at greatest risk based on early behavior problems. In sum, whole-school integrated social-emotional learning and literacy interventions can be an effective way for schools to address the social-emotional well-being of all students and, by improving academic functioning, reduce the likelihood of future mental health and academic problems.

References

- Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology, 39*, 324–348.
- 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.
- Bierman, K. L., Domitrovich, R. L., Nix, S. D., Gest, J. A. W., Greenberg, M. T., Blair, C., . . . Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start program. *Child Development, 79*, 1802–1817.
- 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.
- Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology, 112*, 155–159.
- Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 779–862). New York, NY: Wiley.
- Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology, 67*, 648–657.
- Conduct Problems Prevention Research Group. (2004). The effects of the Fast Track program on serious problem outcomes at the end of elementary school. *Journal of Child and Adolescent Psychology, 33*, 650–661.
- Conduct Problems Prevention Research Group. (2007). Fast Track randomized controlled trial to prevent externalizing psychiatric disorders: Findings from Grades 3 to 9. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*, 1250–1262.
- Cox, M. J., Mills-Koonce, R., Propper, C., & Garipey, J.-L. (2010). Systems theory and cascades in developmental psychopathology. *Development and Psychopathology, 22*, 497–506.
- Dalberg, L. L., Toal, S. B., & Behrens, C. B. (1998). *Measuring violence-related attitudes, beliefs, and behaviors among youths: A compendium of assessment tools*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.
- Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990, December 21). Mechanisms in the cycle of violence. *Science, 250*, 1678–1683.
- Dodge, K. A., Laird, R., Lochman, J. E., & Zelli, A. (2002). Multidimensional latent-construct analysis of children's social-information processing patterns: Correlations with aggressive behavior problems. *Psychological Assessment, 14*, 60–73.
- Domitrovich, C. E., Bradshaw, C. P., Greenberg, M. T., Embry, D., Poduska, J. M., & Jalongo, N. S. (2010). Integrated models of school-based prevention: Logic and theory. *Psychology in the Schools, 47*, 71–88.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. (in press). The effects of social and emotional learning on the behavior and academic performance of schoolchildren. *Child Development*.
- Eddy, J. M., Reid, J. B., Stoolmiller, M., & Fetrow, R. A. (2003). Outcomes during middle school for an elementary school-based preventive intervention for conduct problems: Follow-up results from a randomized trial. *Behavior Therapy, 34*, 525–552.
- Farrell, A. D., Meyer, A. L., Sullivan, T. N., & Kung, E. M. (2003). Evaluation of the Responding in Peaceful and Positive Ways (RIPP) seventh-grade violence prevention curriculum. *Journal of Child and Family Studies, 12*, 101–120.
- Farrell, A. D., & Vulin-Reynolds, M. (2007). Violent behavior and the science of prevention. In D. J. Flannery, A. T. Vazsonyi, & I. D. Waldman (Eds.), *The Cambridge handbook of violent behavior and aggression* (pp. 767–786). New York, NY: Cambridge University Press.
- Forehand, R., Brody, G. H., Armistead, L., Dorsey, S., Morse, E., Morse, P. S., & Stock, M. (2000). The role of community risks and resources in the psychosocial adjustment of at-risk children: An examination across two community contexts and two informants. *Behavior Therapy, 13*, 395–414.
- Foster, E. M., Jones, D. E., & Conduct Problems Prevention Research Group. (2006). Can a costly intervention be cost-effective? An analysis of violence prevention. *Archives of General Psychiatry, 63*, 1284–1291.
- Guerra, N. G., & Bradshaw, C. P. (2008). Linking the prevention of problem behaviors and positive youth development: Core competencies

- for positive youth development. *New Directions in Child and Adolescent Development*, 122, 1–17.
- Hahn, R., Fuqua-Whitley, D., Wethington, H., Lowy, J., Crosby, A., Fullilove, M., . . . Dahlberg, L. (2007). Effectiveness of universal school-based programs to prevent violent and aggressive behavior: A systematic review. *American Journal of Preventive Medicine*, 33, s114–s129.
- Hawkins, J. D., Guo, J., Hill, K. G., Battin-Pearson, S., & Abbott, R. D. (2001). Long-term effects of the Seattle Social Development Intervention on school bonding trajectories. *Applied Developmental Science*, 5, 225–236.
- Hill, C. J., Bloom, H. S., Black, A. B., & Lipsey, M. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2, 172–177.
- Hinshaw, S. (1992). Academic underachievement, attention deficits, and aggression: Comorbidity and implications for intervention. *Journal of Consulting and Clinical Psychology*, 60, 893–903.
- Hoagwood, K. E., Olin, S. S., Kerker, B. D., Kratochwill, T. R., Crowe, M., & Saka, N. (2007). Empirically based school interventions targeted at academic and mental health functioning. *Journal of Emotional and Behavioral Disorders*, 15(2), 66–92.
- Huesmann, L. R., & Guerra, N. C. (1997). Children's normative beliefs about aggression and aggressive behavior. *Journal of Personality and Social Psychology*, 72, 408–418.
- Huesmann, L. R., Guerra, N. G., Miller, L., & Zelli, A. (1992). The role of social norms in the development of aggression. In H. Zumkley & A. Fraczek (Eds.), *Socialization and aggression* (pp. 139–152). New York, NY: Springer.
- Hundert, J., Boyle, M. H., Cunningham, C. E., Duku, E., Heale, J., McDonald, J., . . . Racine, Y. (1999). Helping children adjust—a Tri-Ministry Study: II. Program effects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40, 1061–1073.
- Ialongo, N., Poduska, J., Werthamer, L., & Kellam, S. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders*, 9, 146–160.
- Jones, S. M., Brown, J. L., & Aber, J. L. (2008). Classroom settings as targets of intervention and research. In M. Shinn & H. Yoshikawa (Eds.), *Changing schools and community organizations to foster positive youth development* (pp. 58–77). New York, NY: Oxford University Press.
- Jones, S. M., Brown, J. L., & Aber, J. L. (in press). The longitudinal impact of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*.
- Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, 10, 165–186.
- Lansford, J. E., Malone, P. S., Dodge, K. A., Crozier, J. C., Petit, G. S., & Bates, J. E. (2006). A 12-year prospective study of patterns of social information processing problems and externalizing behaviors. *Journal of Abnormal Child Psychology*, 34, 715–724.
- Leadbeater, B. J., Hellner, I., Allen, J. P., & Aber, J. L. (1989). Assessment of interpersonal negotiation strategies in youth engaged in problem behaviors. *Developmental Psychology*, 25, 465–472.
- Lochman, J. E., Boxmeyer, C., Powell, N., Roth, D. L., & Windle, M. (2006). Masked intervention effects: Analytic methods for addressing low dosage of intervention. *New Directions for Evaluation*, 110, 19–32.
- Lochman, J. E., & Wells, K. C. (2004). The Coping Power Program for preadolescent aggressive boys and their parents: Outcome effects at the 1-year follow-up. *Journal of Consulting and Clinical Psychology*, 72, 571–578.
- Lucas, C. P., Zhang, H., Fisher, P. W., Shaffer, D., Regier, D. A., Narrow, W. E., . . . Friman, P. (2001). The DISC Predictive Scales (DPS): Efficiently screening for diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 443–449.
- Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22, 491–495.
- Masten, A. S., Long, J. D., Kuo, S. I.-C., McCormick, C. M., & Desjardins, C. D. (2009). Developmental models of strategic interventions. *European Journal of Developmental Science*, 3, 282–291.
- McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development*, 71, 173–180.
- McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114, 376–390.
- Milch, R., Loney, J., & Landau, S. (1982). Independent dimensions of hyperactivity and aggression: A validation with playground observation data. *Journal of Abnormal Child Psychology*, 91, 183–198.
- Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development*, 77, 103–117.
- National Center for Education Statistics. (2002). *Early Childhood Longitudinal Study, Kindergarten Cohort of 1998–1999, third-grade assessment*. Retrieved from <http://nces.ed.gov/ecls/Kindergarten.asp>
- National Research Council. (1993). *Understanding and preventing violence*. Washington, DC: National Academy Press.
- O'Connell, M. E., Boat, T., & Warner, K. E. (Eds.). (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.
- Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 77, 302–316.
- Reynolds, C. R., & Kamphaus, R. W. (1998). *Behavioral Assessment System for Children*. Circle Pines, MN: American Guidance Service.
- Rogoff, B. (1990). *Apprenticeship in thinking*. New York, NY: Oxford University Press.
- Rosenfeld, E., Huesmann, L. R., Eron, L. D., & Torney-Purta, J. V. (1982). Measuring patterns of fantasy behavior in children. *Journal of Personality and Social Psychology*, 42, 347–366.
- Rothwell, P. M. (2005). Subgroup analysis in randomised controlled trials: Importance, indications, and interpretation. *Lancet*, 365, 176–186.
- Smolkowski, K., Biglan, A., Barrera, M., Taylor, T., Black, C., & Blair, J. (2005). Schools and Homes in Partnership (SHIP): Long-term effects of a preventive intervention focused on social behavior and reading skill in early elementary school. *Prevention Science*, 6, 113–125.
- Social and Character Development Research Consortium. (2008). *Efficacy of school-wide programs to promote social and character development and reduce problem behavior in elementary school children*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- Tolan, P. H., & Gorman-Smith, D. (2002). What violence prevention research can tell us about developmental psychopathology. *Development and Psychopathology*, 14, 713–729.
- Tremblay, R. E., Masse, L. C., Vitaro, F., & Dobkin, P. L. (1995). The impact of friends' deviant behavior on early onset of delinquency: Longitudinal data from 6 to 13 years of age. *Development and Psychopathology*, 7, 649–667.
- Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. *American Journal of Community Psychology*, 39, 217–228.

Van Lier, P., Muthen, B., van der Sar, R. M., & Crijnen, A. (2004). Preventing disruptive behavior in elementary school children: Impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology, 72*, 467-478.

Vuchinich, S., Flay, B. R., Aber, J. L., Bickman, L., Farmer, T., Gottfredson, G., . . . Pelham, W. (2010). *The effects of person mobility on the design and analysis of place-randomized trials*. Manuscript submitted for publication.

Wentzel, K. R., & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development, 66*, 754-763.

Received December 1, 2008
 Revision received August 18, 2010
 Accepted August 31, 2010 ■

UNITED STATES POSTAL SERVICE® (All Periodicals Publications Except Requester Publications)

Statement of Ownership, Management, and Circulation

1. Publication Title: Journal of Consulting and Clinical Psychology

2. Issue Frequency: Bi-monthly

3. Filing Date: October 2010

4. Issue Frequency: 6

5. Number of Issues Published Annually: 6

6. Annual Subscription Price: Indiv \$261; Inst \$829

7. Complete Mailing Address of Known Office of Publication (Not printer) (Street, city, county, state, and ZIP+4®):
750 First Street, N.E., Washington, D.C. 20002-4242

8. Complete Mailing Address of Headquarters or General Business Office of Publisher (Not printer):
750 First Street, N.E., Washington, D.C. 20002-4242

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor (Do not leave blank):
 Publisher (Name and complete mailing address):
American Psychological Association
750 First Street, N.E.
Washington, D.C. 20002-4242
 Editor (Name and complete mailing address):
Annette M. LaGreca, PhD, Professor of Psychology and Pediatrics
Department of Psychology, University of Miami
P.O. Box 249229, Coral Gables, FL 33124-0751
 Managing Editor (Name and complete mailing address):
Shirah J.A. Harris
American Psychological Association
750 First Street, N.E., Washington, D.C. 20002-4242

10. Owner (Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all individuals owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of all individuals owning or holding 1 percent or more of the total amount of stock. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)
 Full Name: American Psychological Association
 Complete Mailing Address: 750 First Street, N.E., Washington, D.C. 20002-4242

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box None
 Full Name: _____
 Complete Mailing Address: _____

12. Tax Status (For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one):
 Has Not Changed During Preceding 12 Months
 Has Changed During Preceding 12 Months (Publisher must submit explanation of change with this statement)

PS Form 3526, September 2007 (Page 1 of 3) (Instructions Page 2) PSN 7530-01-000-9001 PRIVACY NOTICE: See our privacy policy on www.usps.com

Journal of Consulting and Clinical Psychology

14. Issue Date for Circulation Data Below: August 2010

15. Extent and Nature of Circulation

	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies (Net press run)		
	2930	2600
b. Paid Circulation (By Mail and Outside the Mail)		
(1) Mailed Outside-County Paid Subscriptions Stated on PS Form 3541 (Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)	1897	1862
(2) Mailed In-County Paid Subscriptions Stated on PS Form 3541 (Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)		
(3) Paid Distribution Outside the Mails including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®	623	518
(4) Paid Distribution by Other Classes of Mail Through the USPS (e.g. First-Class Mail®)		
c. Total Paid Distribution (Sum of 15b (1), (2), (3), and (4))		
	2520	2380
d. Free or Nominal Rate (By Mail and Outside the Mail)		
(1) Free or Nominal Rate Outside-County Copies included on PS Form 3541		
(2) Free or Nominal Rate In-County Copies included on PS Form 3541		
(3) Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g. First-Class Mail)		
(4) Free or Nominal Rate Distribution Outside the Mail (Carriers or other means)	17	17
e. Total Free or Nominal Rate Distribution (Sum of 15d (1), (2), (3), and (4))		
	17	17
f. Total Distribution (Sum of 15c and 15e)		
	2537	2397
g. Copies not Distributed (See Instructions to Publishers at page #3)		
	393	203
h. Total (Sum of 15f and g)		
	2930	2600
i. Percent Paid (15c divided by 15h times 100)		
	99%	99%

16. Publication of Statement of Ownership
 If the publication is a general publication, publication of this statement is required. Will be printed in the December 2010 issue of this publication. Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner
Barbara Spruill
 Dir. Service Center Operations
 Date: 10/18/2010

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

PS Form 3526, September 2007 (Page 2 of 3)