

**Symposium Title:** Causal impacts of the 4Rs Program on children, classrooms, and schools: Using dynamic, multilevel analyses to inform theory, practice, and policy.

*Not included in page count.*

**Paper 1**

**Title:** Three Year Cumulative Impacts of the 4Rs Program on Children's Social-Emotional, Behavioral, and Academic Outcomes.

**Author(s):**

Stephanie M. Jones, Ph.D.  
Assistant Professor  
Harvard Graduate School of Education  
14 Appian Way, Larsen 603  
Cambridge, MA 02138  
617-496-2223 phone  
email: jonesst@gse.harvard.edu

Joshua L. Brown, Ph.D.  
Assistant Professor  
Department of Psychology  
Fordham University  
441 East Fordham Road  
Bronx, NY 10458-9993  
718-817-4069 phone  
718-817-3785 fax  
email: cjobrown@fordham.edu

J. Lawrence Aber, Ph.D.  
Department of Applied Psychology  
Steinhardt School of Culture, Education and Human Development  
New York University  
Kimball Hall, Room 417E  
246 Greene Street  
New York, N.Y. 10003  
tel: 212-998-5410  
fax: 212-995-4215  
email: lawrence.aber@nyu.edu

## **Abstract Body**

*Limit 5 pages single spaced.*

### **Background/context:**

*Description of prior research, its intellectual context and its policy context.*

Over the last two decades, developmental science has made significant progress in understanding children's trajectories toward social-emotional and academic outcomes (e.g., Duncan et al., 2007; Miles & Stipek, 2006). At the same time, there has been dramatic growth in the design, implementation, and rigorous evaluation of school-based interventions to promote positive social-emotional development and/or academic achievement (e.g., IES, 2003; Embry, 2002; Kellam et al., 2008). The present study contributes to ongoing scholarship in the school-based prevention of social-emotional, behavioral, and academic problems by reporting three-year longitudinal, experimental impacts of a novel social-emotional learning and literacy development intervention (the 4Rs Program, "Reading, Writing, Respect and Resolution") on a cohort of 3rd grade children's social-emotional, behavioral, and academic functioning after three consecutive years of exposure to the intervention. Findings from this work have implications both for basic theory in developmental science and for the continued evolution of school-based strategies to promote children's social-emotional learning and thereby prevent some of the most ubiquitous mental health problems of middle childhood.

Over the past ten years there has been a burgeoning of programs focused on enhancing children's social and emotional skills to reduce aggression and violence and promote positive interaction among youth (Payton et al., 2008; Wilson, Lipsey & Derzon, 2003). This period has also witnessed a growing convergence of developmental science and prevention science in guiding the design and evaluation of interventions aimed at preventing future aggressive and violent behavior in children and youth (Maggs & Schulenberg, 2001). From developmental science, knowledge has grown about the mechanisms by which exposure to violence affects children's risk for such outcomes (Dodge, 2006). Our greater understanding of these causal mechanisms has led to improvements in both the design and evaluation of preventive interventions, which increasingly target these mechanisms as their focus of change (e.g., Dodge, 2001; Hudley & Graham, 1993). Indeed several of these causal mechanisms are central to the 4Rs Program and to the evaluation design of this study (see below).

From prevention science, knowledge has grown about the effectiveness of these intervention strategies at reducing children's risk for future aggressive and violent behavior (Aber et al., 2003; Conduct Problems Prevention Research Group, 1999; 2007; Kellam et al., 1998). Although the literature on school-based preventive interventions is rich with studies of interventions targeted at subgroups of high-risk children (Vitaro, Brendgen, & Tremblay, 2001), with a few exceptions (e.g., the Good Behavior Game; Kellam, et al., 2008), it has only recently expanded in the area of large-scale evaluations of universal interventions implemented with general populations of students (Payton et al., 2008). Furthermore, among the best studies of universal school-based interventions designed to reduce risk for future aggression and violence, a variety of methodological challenges limit the quality and generalizability of the knowledge base. With but a few exceptions to date (e.g., CPPRG, 1999; 2002; 2004; 2007; Kellam et al., 1998), these earlier studies rarely employ both experimental random assignment designs and analytic methods appropriate to the design (e.g., multi-level modeling with intervention status modeled at the relevant level of random assignment, frequently the school-level) that enable one to make definitive causal statements about the impact of the intervention. This study is distinct in

that it provides an experimental test of the causal impact of a universal integrated intervention model in which social-emotional learning and skill building is embedded in a balanced literacy curriculum.

**Purpose / objective / research question / focus of study:**

*Description of what the research focused on and why.*

The present study is the first report of the experimental effects of the 4Rs Program on change over three years using six repeated assessments in children's social-cognitive, social-emotional, behavioral, and academic outcomes. The primary questions addressed in the paper are: (1) What is the experimental impact of the 4Rs Program on three-year change (from 3rd to 5th grade) in children's social-cognitive processes, social-emotional symptomatology, their aggressive and socially competent behavior, and academic functioning, controlling for key demographic covariates? (2) Building on recent findings from intervention studies that demonstrate significantly stronger impacts for families facing a greater versus smaller number of poverty-related risks, is the three-year impact of the 4Rs Program moderated by child-level demographic baseline covariates including child gender, race/ethnic background, family socioeconomic risk, and community risk? (3) Finally, building on evidence from the first year of our 4Rs evaluation (Jones et al., under review) and other experimental evaluations of universal school-based prevention programs (CPPRG, 2007), is the three-year impact of the 4Rs Program moderated by children's baseline behavioral risk?

**Setting:**

*Description of where the research took place.*

This project took place over three consecutive school years in 18 public elementary schools in New York City (in four of the five boroughs).

**Population / Participants / Subjects:**

*Description of participants in the study: who (or what) how many, key features (or characteristics).*

Participants were 1184 children (49% boys; average age at time 1 = 8.17 yrs, SD = 0.7), and 146 teachers (88% female; average age = 35) in 18 public inner-city elementary schools in a large metropolitan city in the Eastern United States. The children and teachers are part of an ongoing, longitudinal evaluation of a universal, school-wide literacy and social-emotional learning prevention program (4Rs: Reading, Writing, Respect and Resolution) implemented for three consecutive years in 9 intervention (n = 630; 53.2%) and 9 control (n = 554; 46.8%) schools. As noted above, in this paper we describe impacts of 4Rs on children's developmental trajectories across 6 repeated time points, and three consecutive years of exposure. Data were gathered from children and their teachers over six longitudinal waves across 3 school years. Because the 4Rs Program was randomized at the school-level, children who moved out of a participating school were not followed (e.g., 58 children in Wave 2) and consent was requested for new children who moved into a participating school at each follow-up wave (e.g., 124 children at Wave 2). Attrition between waves was minimal (on average 8.6%) and was primarily due to student mobility out of participating schools. Refusals from parents to continue were rare (e.g., n = 1 in Wave 2). According to parent-reports at baseline, 52% (n = 425) of children lived in a single-parent household, 14.5% (n = 119) of parents were unemployed, 29.7% (n = 243) of parents had less than a high school diploma or GED, and 60.9% (n = 498) of households were at or below 100%

of the federal poverty level. Based on parent-reports at baseline and NYC Department of Education records when parent-reports were missing, children represented diverse racial/ethnic groups; 45.3% (n = 368) were Hispanic/Latino, 41.1% (n = 334) Black/African American, 5% (n = 41) non-Hispanic White, and 8.6% (n = 70) represented other racial/ethnic groups (e.g., Asian, Pacific Islander, Native American).

**Intervention / Program / Practice:**

*Description of the intervention, program or practice, including details of administration and duration.*

The 4Rs Program (Reading, Writing, Respect and Resolution) is a universal, school-based intervention in literacy development and social-emotional learning that integrates a focus on social and emotional development into the language arts curriculum for children in grades K-5. In this evaluation, Developed and run by a community-based non-profit organization called the Morningside Center for Teaching Social Responsibility, 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. By highlighting universal themes of conflict, feelings, relationships, and community, the 4Rs curriculum adds social and emotional meaning and skill building to rigorous literacy instruction. The 4Rs Program has two primary components: (1) a comprehensive 7-unit, 21-35 lesson, literacy-based curriculum in social-emotional learning and (2) 25 hours of training followed by ongoing coaching of teachers to support them in teaching the 4Rs curriculum with a minimum of 12 contacts in one school year.

**Research Design:**

*Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).*

Forty-one schools representative of the population of NYC elementary schools were originally identified as potential participants in the 4Rs evaluation. Of these 41, 24 agreed to the process of matching and randomization. Prior to randomization, a pairwise matching procedure was used to maximize demographic similarity of intervention and control groups. An algorithm was used to compute the distance from each school to every other school along 20 demographic and school characteristics. These variables were drawn primarily from the 2001-2002 administrative databases kept by the city's Department of Education and were selected to represent a number of important dimensions related to the criterion outcomes. To conduct random assignment of matched pairs to 4Rs intervention and control groups, a MatLab uniform random numbers generator was employed to generate, in sequence, 12 random numbers ranging from 0-1 that were assigned to the first school in each of the 12 pairs (note, a total of 24 schools were recruited to participate in this study and were matched into 12 pairs, the 9 best matching pairs were kept as study schools and 3 pairs were kept as back-ups). The first school in each pair was assigned to the intervention or control group based on the randomly generated number, and the second school in the pair was, therefore, assigned to the other group. After random assignment, the two groups were compared across the 20 demographic characteristics employed in the matching procedures. As expected the two groups did not differ significantly on any of these characteristics and eta<sup>2</sup> values (the proportion of variance in the demographic characteristic explained by differences between the two groups) were minimal. Based on these statistics, the schools can be described as racially and ethnically diverse, composed primarily of students who

receive a free school lunch, and characterized by attendance rates over 89% and one-year stability rates that range from 86% to 95%.

### **Data Collection and Analysis:**

*Description of the methods for collecting and analyzing data.*

Consent packages (in English and Spanish) were sent home to all parents of third grade children in the 18 participating schools informing them of the study and seeking consent for their child to participate. The overall consent rate was 64.54% across schools (range = 44% to 79%); consent rates did not differ between treatment (65.2%) and control (63.7%) schools. Non-participants included children whose parents did not speak English or Spanish well-enough to consent to participate and special needs children who could not be interviewed even on an individual basis (e.g., due to autism). At each wave teachers completed questionnaires rating the language and literacy skills, as well as social competence and externalizing problems of each child in their class with consent to participate. Teachers were paid at the union wage of \$36.50/per hour for completion of the surveys at each assessment. At each wave, children also completed questionnaires rating their aggressive social-cognitions, pro-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 placement of responses and to answer the children's questions. Children who did not have consent to participate or who refused assent worked on an alternative activity with their classroom teacher.

To accommodate the nested nature of the design, estimates of intervention impact on change in the primary child outcomes from pre-intervention baseline (Wave 1, Fall 2004) to the 6th time point (Wave 6, Spring 2007) were calculated using a series of 3-level hierarchical linear growth models with school fixed effects in HLM 6.02. In these models, Level 1 represents time (i.e., the 6 repeated assessments of the constructs of interest for each child), Level 2 represents the child, and Level 3 represents schools. All child- and classroom-level covariates were included at Level 2. Level 3 included a treatment dummy as well as 8 school pair dummies to represent the school matches (the worst matching pair, #9, served as the referent). In addition, as indicated above, we examined a number of cross-level treatment by baseline covariate interactions.

### **Findings / Results:**

*Description of main findings with specific details*

Findings to date through the second year of the intervention are summarized below. Analyses examining intervention effects through the end of the third year will be complete by December 1, 2009. Findings are presented below by broad outcome domain.

Social Cognitive Processes and Social-Emotional Symptomatology. Over two consecutive years, children in 4Rs schools self-reported slower rates of increase in hostile attributional bias, a slowed rate of growth in aggressive interpersonal negotiation strategies that appears to begin toward the outset of the second year of exposure to intervention, and a steeper rate of decline in depressive and ADHD symptoms compared to children in the control schools.

Aggressive and Socially Competent Behavior. Teachers in 4Rs schools reported slower growth in children's aggressive behavior (compared to increases in control schools), and increases in social competence (compared to declines in control schools) over two school years.

Academic Functioning. While there were no main effects of treatment on teacher reports of

children's academic skills or on the three school records outcomes examined, there were treatment by baseline behavioral risk interactions for standardized math and reading achievement and for teacher reported academic skills. In short, children identified by teachers at greatest behavioral risk at baseline showed greater improvements as a result of exposure to 4Rs in their math and reading achievement and in teacher reports of their academic skills (see Figure 1 in Appendix B). Importantly, this set of treatment by baseline behavioral risk interactions were not evident for the social-emotional outcomes, regardless of the type of model examined (i.e., as a growth model estimating treatment effects on growth parameters, or a basic point-in-time model estimating treatment effects on Wave 4, controlling for baseline levels). This suggests that these treatment by risk interactions are not an artifact of the form of model applied, but instead are tied to the developmental domain examined: children's academic functioning and not, in this case, their social-emotional skills.

Over all, the treatment main effects reported after two years are small to moderate in size (ranging in size from .05 for the teacher-reported aggression slope to .22 for the depressive symptoms slope). In contrast, the interactions of treatment with baseline behavioral risk are represented by treatment effects for the highest risk group of moderate size (ranging from .56 for math achievement and academic skills to .60 for reading achievement).

### **Conclusions:**

*Description of conclusions and recommendations based on findings and overall study.*

Our findings to date on the impacts of an integrated, social-emotional and literacy program provides clear evidence that this universal intervention has both broad impacts on social-cognitive processes and behaviors in the social-emotional domain, and targeted impacts in the academic domain. This study provides good evidence that universal school-based interventions, delivered to whole populations of children, can result in substantial impacts on children's developmental health and well-being.

From the standpoint of practice, our current evidence suggests that integrating pedagogical attention to building social-emotional skills through simultaneously enriching literacy practices, as instantiated by the 4Rs Program, can promote positive development in both social-emotional and academic domains. These findings challenge schools and school-based program practitioners to continue to conceptualize and operationalize a variety of practice models in which social-emotional and academic development can be fostered both in the classroom and in the school as a whole. While the 4Rs has demonstrated some success towards this goal, developing and testing integrated models that may be more appropriate for other geographic locations and student and teacher populations will require considerable attention and resources. From the standpoint of policy, findings from the present study highlight the short-sightedness of educational policies that privilege and reward school and teacher attention to narrowly defined domains of development such as academic performance at the exclusion of attention to children's social-emotional development. A growing theoretical and empirical literature supports the inextricably connected links between development in these domains and so we must now face the challenge of developing and adopting educational policies that both acknowledge and support this reality.

## Appendices

*Not included in page count.*

### Appendix A. References

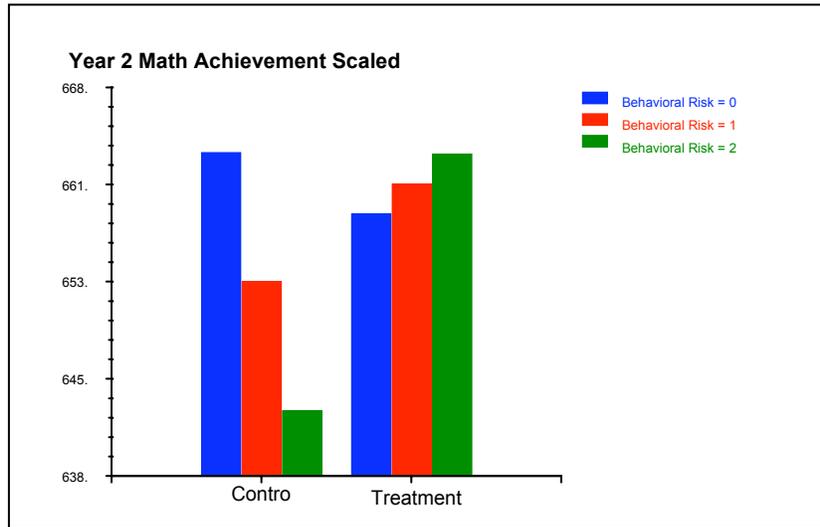
*References are to be in APA version 6 format.*

- Aber, J.L., Brown, J.L. & Jones, S.M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology, 39*(2), 324-348.
- Conduct Problems Prevention Research Group (CPPRG, 1999). Initial impact of the Fast Track Prevention Trial for Conduct Problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology, 67*, 648-657.
- CPPRG (2002). The implementation of the Fast Track Program, an example of a large-scale prevention science efficacy trial. *Journal of Abnormal Child Psychology, 30*(1), 1-17.
- CPPRG (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*(4), 650-661.
- CPPRG (2007). Fast Track randomized controlled trial to prevent externalizing psychiatric disorders: findings from grades 3 to 9. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(10), 1250-1262.
- Dodge, K. A. (2001). The science of youth violence prevention: Progressing from developmental epidemiology to efficacy to effectiveness to public policy. *American Journal of Preventive Medicine, 20*(1) (Suppl.), 63-70.
- Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology, 18*, 791-814.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Klebanov, P., Pagani, et al. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446.
- Embry, D.D. (2002). The Good Behavior Game: A best practice candidate as universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*(4), 273-297.
- Hudley, C.A. & Graham, S. (1993). An attributional intervention to reduce peer-directed aggression among African-American boys. *Child Development, 64*, 124-138.
- Institute of Educational Sciences. (2003). Social and character development research. <http://www.sacdprojects.net>.

- Kellam, S. G., Brown, C. H., Poduska, J.M., Ialongo, N.S., Wang, W., Toyinbo, P., Petras, H., Ford, C., Windham, A., & Wilcox, H. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95S*, S5-S28.
- 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*(2), 165-186.
- Maggs, J. L., & Schulenberg, J. (2001). Editors' introduction: Prevention as altering the course of development and the complimentary purposes of developmental and prevention sciences. *Applied Developmental Science, 5*(4), 196-200.
- 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*(1), 103-117.
- 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.
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*(1), 136-149.
- Vitaro, F., Brendgen, M., & Tremblay, R. E. (2001). Preventive intervention: Assessing its effects on the trajectories of delinquency and testing for mediational processes. *Applied Developmental Science, 5*(4), 201-213.

**Appendix B. Tables and Figures**  
*Not included in page count.*

Figure 1. Interaction of treatment and baseline behavioral risk on Year 2 math achievement



## **Paper 2**

*Not included in page count.*

**Title:** The Longitudinal Impact of a Universal School-Based Social-Emotional and Literacy Intervention on Classroom Climate and Teacher Processes and Practices

**Author(s):**

Joshua L. Brown, Ph.D.  
Assistant Professor  
Department of Psychology  
Fordham University  
441 East Fordham Road  
Bronx, NY 10458-9993  
718-817-4069 phone  
718-817-3785 fax  
email: [cjobrown@fordham.edu](mailto:cjobrown@fordham.edu)

Stephanie M. Jones, Ph.D.  
Assistant Professor  
Harvard Graduate School of Education  
14 Appian Way, Larsen 603  
Cambridge, MA 02138  
617-496-2223 phone  
email: [jonesst@gse.harvard.edu](mailto:jonesst@gse.harvard.edu)

J. Lawrence Aber, Ph.D.  
Department of Applied Psychology  
Steinhardt School of Culture, Education and Human Development.  
New York University  
Kimball Hall, Room 417E  
246 Greene Street  
New York, N.Y. 10003  
tel: 212-998-5410  
fax: 212-995-4215  
email: [lawrence.aber@nyu.edu](mailto:lawrence.aber@nyu.edu)

## Abstract Body

### **Background/context:**

Increasingly, school-based intervention and whole school reform efforts aim at identifying, assessing, and effecting changes in classroom-level processes associated with or predictive of children's social-emotional and academic development (Hamre & Pianta, 2005; Pianta, 2006; Raver et al., 2008; Rimm-Kaufman, LaParo, Downer, & Pianta, 2005). These classroom-processes are fundamentally social in nature, reflect the underlying quality of the interactions among and between teachers and students and encompass emotional, instructional, and organizational dimensions of classroom experience. Indeed, in the absence of improving such social processes, other resources such as qualified teachers or costly curricular materials may be ineffective in promoting learning and achievement (Cohen, Raudenbush & Ball, 2003; Fullan, 2001). Reliable and valid methods and measures for assessing important classroom-level social processes have recently been developed and are gaining use in the research community (Pianta, LaParo & Hamre, 2008). But little is currently known about the ability of school and classroom-based interventions to successfully alter these dimensions of classroom settings, particularly in poorly functioning classrooms.

Another important set of constructs that has been neglected in research examining the effects of social-emotional learning and literacy development programs may be considered together in the domain of teacher affective and pedagogical processes and practices. A teacher's orientation toward their own professional development (Selman, 2003), their perceptions of their role in attending to students' social-emotional needs (Ryan, Gheen, & Midgely, 1998), their experience of stress and feelings of job-burnout (Maslach, Jackson & Schwab, 1996; Yoon, 2002), their classroom management styles and strategies (Wentzel, 2002), and their perceptions of their own emotional abilities have each been identified as critical dimensions of teachers associated with the development of children's social and/or academic competence. Teachers' values, beliefs, and perceptions of ability for promoting students' social-emotional learning may also change through their involvement in intervention programs (Adalbjarnardottir, 1994). A recent literature review concluded (Daniels & Shumow, 2003) there is a critical need for more research on teachers' views of their role in addressing children's social-emotional needs. Whether teachers feel stressed and overwhelmed, believe in the importance of the intervention goals, and feel that they have relevant skills and knowledge all have direct implications both for the type of classroom setting they create for children and for the effectiveness of classroom- and school-based interventions.

### **Purpose / objective / research question / focus of study:**

This presentation capitalizes on a three-year, longitudinal, school-randomized trial of the 4Rs Program, a comprehensive, school-based social-emotional and literacy program for elementary schools, to test intervention induced changes in features of classroom climate and key dimensions of teacher affective and pedagogical processes and practices thought to influence children's social-emotional and academic development. Experimental evidence to date suggests the 4Rs Program is effective in improving the quality of classroom climate after one-year of intervention (Brown, Jones, LaRusso & Aber, accepted pending minor revisions) and positively altering the course of children's social-cognitive, emotional, behavioral and academic functioning after two years of intervention (Jones, Brown, & Aber, revision under review) [see Findings/Results section below for more detail]. While these results are encouraging, this

presentation will extend this work by testing the causal influence of the 4Rs Program on (1) the quality of classroom climate at the end of the second and third years of intervention, and (2) changes in teachers affective and pedagogical processes and practices over three consecutive years of intervention.

The role of classroom settings and teachers is highlighted in the intervention's theory of change, with the belief that more supportive classroom environments and positive changes in key teacher processes and outcomes underlies positive change in students. The 4Rs Program makes a significant early investment in teacher preparation beginning with 25 hours of high quality teacher training and ongoing professional support throughout the school year. The training process encourages teachers to utilize the ideas and skills of the 4Rs curriculum in their own lives. The philosophy of the program revolves around the idea that teachers who more strongly align their personal beliefs with the values promoted by the program will have classrooms where more positive child outcomes occur. As such, the current report on teacher outcomes will begin to illuminate the causal process by which the program has yielded positive results for children.

### **Setting:**

This presentation is based on research that took place between Fall 2004 and Spring 2007 in eighteen New York City public elementary schools that agreed to participate in a three-year, school-randomized study to test the impacts of the 4Rs Program on children, teachers and classrooms. Schools were located across four of the five boroughs of New York City and were demographically representative of the larger NYC elementary school population. Participating schools were pairwise matched and then randomly assigned (see Research Design section below) to intervention (9 schools) and control (9 schools) conditions. The 4Rs Program was implemented in the intervention schools for three consecutive years.

### **Population / Participants / Subjects:**

This study will examine the impact of the 4Rs Program on independent observations of classroom climate collected in 147 3<sup>rd</sup> and 4<sup>th</sup> grade classrooms during the 2005-2006 school year and 191 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade classrooms during the 2006-2007 school year. This study will also examine the impact of the 4Rs Program on teacher processes and practices, and will include 299 teachers across the 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> grades. The vast majority of teachers are female (89%), and slightly more than half (53%) self-reported as non-Hispanic White, 24% as Black or African-American, 15% as Hispanic, and the remaining teachers as Asian, American-Indian, or Alaskan Native. Teachers averaged 34 years of age (SD=10 years), 8 years of teaching experience overall (SD=6.9) and 5 years at their current school (SD=5). Sixty-eight per cent of teachers reported having a Master's degree, and 30% reported a Bachelor's degree. Most teachers (67%) reported having a New York State Teaching Certificate.

### **Intervention / Program / Practice:**

The 4Rs Program is a school-based intervention in literacy development, conflict resolution, and intergroup understanding that trains and supports all teachers in grades K-5 in how to integrate the teaching of social and emotional skills into the language arts curriculum. It is considered a universal intervention in that it targets and is implemented with the entire teacher and student population of a given school (Institute of Medicine, 1994). Through the program, teachers learn how to use 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. The 4Rs Program provides a pedagogical link between the teaching of conflict resolution and the teaching of fundamental academic skills, thereby capitalizing on their mutual influence on successful youth development (Hinshaw, 1992; Jones et al, revision under review).

The 4Rs Program has two primary components: (1) a comprehensive 7-unit, 21-35 lesson, literacy-based curriculum in conflict resolution and social-emotional learning and (2) 25 hours of training followed by ongoing coaching of teachers to support them in teaching the 4Rs curriculum with a minimum of 12 contacts in one school year.

At its core, the program's theory of change involves helping teachers more deeply assimilate, find utility in, and become skilled at practicing the concepts of the 4Rs Program in their own lives and teaching them in their classroom through the consistent delivery of lessons from the 4Rs curriculum and the provision of greater social-emotional learning opportunities in which students can practice the component skills and be supported in applying them in real life situations. The alignment of teachers' own values, beliefs, and perceptions of ability with the underlying pedagogy of a particular intervention is critical to their understanding, acceptance, and implementation of the intervention, and to the effectiveness of the intervention itself (CPPRG, 1999; Fullan & Stiegelbauer, 1991; Hauer, 2003). Teachers who practice good listening skills (e.g., direct eye contact, paraphrasing, acknowledging comprehension) during interactions with their students and other adults, and who can teach these skills and provide real-life, real-time examples of how they are effective, increase the likelihood their students will employ them in their own interactions. But it is not merely the practice of good listening skills by the teacher or any given student that is important; it is how the use of these skills reflects a set of transactional social processes in the classroom that enable teachers and students to develop closer, more intimate relationships (Pianta, 2006; Tseng & Seidman, 2007). Increases in the quality of the relationships among teachers and students in turn facilitates future positive communications by fostering a more responsive classroom overall. Therefore, central to the program's theory of change is that teachers are successfully engaged in serving as the gateway to changing broad characteristics of classrooms including relationships and climate, as well as in the development of individual children.

School-wide implementation of the two primary components of the 4Rs Program (curriculum delivery and teacher training and coaching) was systematically tracked and monitored during the course of the study. According to implementation data from Year 1, teachers in the 9 treatment schools received (a) on average 2.4 (SD=.33) days of training in the delivery of the 4Rs curriculum, and (b) an average of 38 (SD=9.6) days per school. On average, teachers delivered three-quarters of a lesson per week, with the majority closer to the benchmark of 1 lesson per week. The majority of teachers appear to have spent on average between 20-25 (~40 minutes/week) total hours during Year 1 on 4Rs. Year 2 implementation data revealed a slight decrease in training days, and a slight increase in coaching days and the average classroom lessons per week, and the amount of time spent on 4Rs per week. Our data also indicate that teachers who were trained in the first year of the study, and who remained in the school the following year, were even closer to program benchmarks (i.e., on average they implemented 1 lesson/week and spent ~50 minutes on 4Rs per week). While there is variability in 4Rs implementation between teachers and schools, this variation is not inconsistent with similar programs and evaluation studies that focus on public schools (e.g., Kam, Greenberg & Walls, 2003).

**Research Design:**

This is an experimental (school-randomized), longitudinal study of the 4Rs Program. An initial pool of 24 recommended public elementary schools were matched into 12 pairs on 20 different school-level characteristics (e.g., school size, percent of students receiving free lunch, racial/ethnic composition, student attendance and achievement, average spending per student, teacher experience) drawn from the 2001-2002 administrative database maintained by the New York City Department of Education. A uniform random numbers generator was employed to generate, in sequence, 12 random numbers ranging from 0-1 that were assigned to the first school in each of the 12 pairs (note, a total of 24 schools were recruited to participate in this study and were matched into 12 pairs, the 9 best matching pairs were kept as study schools and 3 pairs were kept as back-ups). The first school in each pair was assigned to the intervention or control group based on the randomly generated number, and the second school in the pair was, therefore, assigned to the other group. After random assignment, the two groups were compared across the 20 demographic characteristics employed in the matching procedures. As expected the two groups did not differ significantly on any of these characteristics and eta<sup>2</sup> values (the proportion of variance in the demographic characteristic explained by differences between the two groups) were minimal. Based on these statistics, the schools can be described as racially and ethnically diverse, composed primarily of students who receive a free school lunch, and characterized by attendance rates over 89% and one-year stability rates that range from 86% to 95%.

**Data Collection and Analysis:**

While the overall study involved the collection of data from children, parents, teachers, and via independent classroom observations, the focus of this presentation is limited to independent observations of classroom climate and teacher self-report data collected from 3<sup>rd</sup> through 5<sup>th</sup> grade classrooms/teachers over three consecutive school years.

*Classroom Climate* was assessed via independent observation using the Classroom Assessment Scoring System (Pianta, LaParo, & Hamre, 2008) in 3<sup>rd</sup> grade classrooms during the spring of the first year of the study (n=82), in 3<sup>rd</sup> and 4<sup>th</sup> grade classrooms during the fall and spring of the second year (n=147), and in 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> grade classrooms during the fall and spring of the third year (n=191). Observations were conducted by a multiracial/multiethnic field research team who received extensive training to reliability standards in the use of this instrument. Classrooms were observed for two hours, divided into four 20-minute periods of observation, each observation period followed by a 10-minute rating period. CLASS assesses three primary domains of classroom climate in preschool through fifth grade classrooms: Emotional Support, Classroom Organization, and Instructional Support. Each broad domain is comprised of several specific dimensions of interactions. Each dimension, in turn, is represented by a continuum of indicators of that dimension, each indicator including a behaviorally anchored, observable description of interactions in the classroom (teacher-student, student-student). Internal reliability was .90 for Emotional Support, .83 for Classroom Organization, and .90 for Instructional Support and .93 for a composite index of all three subscales. There is strong support for the psychometric properties of the CLASS, including demonstrated relationships to children's social and academic development (predictive validity) both during the preschool (e.g., Howes et al., 2008) and elementary school years (e.g., Pianta et al., 2008).

Assessments of *teacher affective and pedagogical processes and practices* were collected directly from teachers through paper and pencil survey questionnaires. Similar to the structure of classroom observations, teacher surveys were collected during the fall and spring from 3<sup>rd</sup> grade teachers during the first year of the study, from 3<sup>rd</sup> and 4<sup>th</sup> grade teachers during the second year of the study, and from 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade teachers during the third year of the study. Specific measures include: *Teaching strategies* (Teacher Strategies Questionnaire; Webster-Stratton, Reid, & Hammond, 2001), assessing both the Frequency and Usefulness of Positive Strategies (12 items, alphas = .76-.82) and Inappropriate Strategies (9 items each, alphas=.69-.77); *Teacher Perceptions of Role in Students' Social-Emotional Well-Being* (Ryan, Gheen, & Midgely, 1998), a 7-item measure assessing the degree to which teachers believe they play a central role in their student's social-emotional development (alpha=.76-.83); *Teacher Burnout* (Maslach Burnout Inventory-Educators Survey, Maslach, Jackson, & Schwab, 1996), including and overall burnout index (18 items, alphas=.88-.91) and subscales of Emotional Exhaustion (9 items, alphas=.88-.91), Depersonalization (5 items, alphas=.73-.81), and Personal Accomplishment (7 items, alphas=.82-.85); and *Perceived Emotional Intelligence Scale* (Brackett & Mayer, 2003), a 17-item measure assessing teacher's self-reported ability to perceive and understand the emotions of others and regulate their own emotions (alphas=.72-.87).

All analyses will be conducted using multi-level hierarchical linear modeling (HLM 6.02). Variation in classroom climate at the end of Year 2 (Spring 2006, after two years of intervention) and the end of Year 3 (Spring 2007, after three years of intervention) will be estimated as a function of teacher and classroom demographic control variables at level-1 (e.g., teacher experience, class size) and school factors at level-2 (including treatment status and eight dummy variables representing school matched-pair status). Estimates of intervention impact on change in the primary teacher processes and practices outcomes from pre-intervention baseline (Fall 2004) to the final assessment (Spring 2007) will be calculated using a series of 3-level hierarchical linear growth models with school fixed effects. In these models, Level 1 represents time (i.e., the 6 repeated assessments of the constructs of interest), Level 2 represents the teacher outcomes, and Level 3 represents schools (including treatment status and eight school pair dummies). Key teacher- and classroom-level covariates will be included at Level 2.

### **Findings / Results:**

Recent findings have demonstrated that the 4Rs Program positively influences both classroom-level outcomes. Brown et al (accepted pending minor revisions) found that after one year of intervention, classrooms in 4Rs schools were rated by independent observers as higher in overall quality compared to classrooms in control schools. Specifically, this work indicated that 4Rs classrooms had higher mean levels of classroom emotional support and instructional support (but not classroom organization) than control classrooms at the end of the first year of the intervention. Results of the analyses outlined in the previous section, and currently underway, will focus on the impacts of the 4Rs Program on classroom climate in Years 2 and 3 and on teacher processes and practices across all three years of the intervention.

### **Conclusions:**

Based on these compelling experimental results to date, it is critical to test the impact of the 4Rs Program on classroom- and teacher-level outcomes given their centrality as levers for student-level change in the 4Rs Program and in a majority of school-based social-emotional learning interventions.

## Appendices

*Not included in page count.*

### Appendix A. References

- Adalbjarnardottir, S. (1994). Understanding children and ourselves: Teachers' reflections on social development in the classroom. *Teaching and Teacher Education*, 10(4), 409-421.
- Brackett, M. A. & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin*, 29(9), 1147-1158.
- Brown, J.L., Jones, S.M. LaRusso, M. & Aber, J.L. (accepted, pending minor revision). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs Program. *Journal of Educational Psychology*.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 119-142.
- Conduct Problems Prevention Research Group (CPPRG, 1999). Initial impact of the Fast Track Prevention Trial for Conduct Problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, 67, 648-657.
- Daniels, D. H., & Shumow, L. (2003). Child development and classroom teaching: A review of the literature and implications for educating teachers. *Journal of Applied Developmental Psychology*, 23(5), 495-526.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York, NY: Teachers College Press.
- Fullan, M. & Stiegelbauer, S. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, 76(5), 949-967.
- Hauer, J. (2003). Educating for character and teachers' moral vitality. *Journal of Research in Character Education*, 1(1), 31-42.
- Hinshaw, S. P. (1992). *Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms*. *Psychological Bulletin*, 111(1) 127-155.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly*, 23, 27-50
- Institute of Medicine (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Jones, S. M., Brown, J. L., & Aber, J. L. (revision under review). The Longitudinal Impact of a Universal School-Based Social-Emotional and Literacy Intervention: An Experiment in Translational Developmental Research. *Child Development*.
- Maslach, C., Jackson, S. E., & Schwab, R. L. (1996). Maslach Burnout Inventory-Educators Survey (MBI-ES). In C. Maslach, S. E. Jackson, & M. P. Leiter, *MBI Manual (3rd ed.)*. Palo Alto, CA: Consulting Psychologists Press.

- Pianta, R. C. (2006). Schools, schooling, and developmental psychopathology. In D. Cicchetti & D. Cohen (Eds.). *Developmental psychopathology (2nd edition), Volume one: Theory and method* (pp. 494-529). John Wiley & Sons.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, M.D.: Paul H. Brookes Publishing Co., Inc.
- Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, 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.
- Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *Elementary School Journal*, 105(4), 377-394.
- Ryan, A. M., Gheen, M. H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social-emotional role, and the classroom goal structure. *Journal of Educational Psychology*, 88, 1-8.
- Selman, R. L. (2003). *The promotion of social awareness: powerful lessons from the partnership of developmental theory and classroom practice*. New York: Russell Sage Foundation.
- Tseng, V., & Seidman, E. [CE: (2007). A systems framework for understanding social settings. *American Journal of Community Psychology*, 39, 217 – 228.
- Webster-Stratton, C., Reid, J. M., & Hammond, M. (2001). Preventing Conduct Problems, Promoting Social Competence: A Parent and Teacher Training Partnership in Head Start. *Journal of Clinical Child Psychology*, 30(3), 283–302.
- Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, 73(41), 287-301.
- Yoon, J. S. (2002). Teacher characteristics as predictors of teacher-student relationships: Stress, negative affect, and self-efficacy. *Social Behavior and Personality*, 30(5), 485-493.

### **Paper 3**

*Not included in page count.*

**Title:**

Using administrative data to evaluate impacts in a school-randomized trial of the 4Rs Program

**Author(s):**

Juliette Berg, BSc.

Doctoral candidate.

New York University.

Steinhardt School of Culture, Education and Human Development.

Kimball Hall, Room 417E

246 Greene Street

New York, N.Y. 10003

tel: 212-998-5561

email: jberg@nyu.edu

Catalina Torrente, BSc.

Doctoral candidate.

New York University.

Steinhardt School of Culture, Education and Human Development.

Kimball Hall, Room 417E

246 Greene Street

New York, N.Y. 10003

tel: 212-998-5561

e-mail: cet254@nyu.edu

J. Lawrence Aber, Ph.D.

Department of Applied Psychology

New York University.

Steinhardt School of Culture, Education and Human Development.

Kimball Hall, Room 417E

246 Greene Street

New York, N.Y. 10003

tel: 212-998-5410

fax: 212-995-4215

email: lawrence.aber@nyu.edu

Stephanie M. Jones, Ph.D.

Assistant Professor

Harvard Graduate School of Education

14 Appian Way, Larsen 603

Cambridge, MA 02138

617-496-2223 phone

email: jonesst@gse.harvard.edu

Joshua L. Brown, Ph.D.  
Assistant Professor  
Department of Psychology  
Fordham University  
441 East Fordham Road  
Bronx, NY 10458-9993  
718-817-4069 phone  
718-817-3785 fax  
email: [cjobrown@fordham.edu](mailto:cjobrown@fordham.edu)

## **Abstract Body**

*Limit 5 pages single spaced.*

### **Background/context:**

*Description of prior research, its intellectual context and its policy context.*

Educators and policymakers continually search for effective strategies to improve the opportunities for learning and positive social-emotional development offered to children in schools. In spite of the recent nationwide proliferation of whole-school reform efforts that target these opportunities, our empirical base for understanding what makes for an effective program is still very limited. Most of the gaps relate to poor methodological designs that prevent researchers from making reliable and valid causal inferences about program effects. For example, in a meta-analysis on comprehensive-school reform (CSR) by Borman, Hewes, Overman & Brown (2003), only 7 out of 24 models reviewed were classified as having sufficient, reliable and generalizable information to estimate program effects. In addition, while the purpose of whole-school reform efforts is to promote school-wide improvements, these programs are often evaluated using a single cohort of students and the effects of the program on the whole student population's academic and behavioral outcomes are rarely examined. Finally, while the connection between academic and social development is being increasingly acknowledged (Blair & Diamond, 2008; Miles & Stipek, 2006), most whole-school reform efforts focus primarily on either academic outcomes (e.g., Success for All) or social-emotional outcomes (e.g., PATHS), with only a few initiatives targeting both sets of outcomes.

The 4Rs Program (Reading, Writing, Respect and Resolution) is a "dual focus" whole school universal intervention designed to promote literacy development and social-emotional learning, that is currently being rigorously evaluated using a school-randomized trial of 18 elementary schools (9 intervention, 9 control) in New York City. Analyses of program impacts across the first two years of the study suggest significant positive effects of the intervention on children's social-emotional outcomes, including aggression, depression, and social competence, and on math achievement for children at highest behavioral risk (see Jones, Brown, Hoglund, & Aber, under review). While the primary research method for these analyses has been to follow a small (n=900) cohort of 3rd grade students over three years (through the end of 5<sup>th</sup> grade) using intensive survey and observational research techniques, the present study utilizes administrative data to estimate the impact of 4Rs on multiple cohorts of students. This study will both allow us to estimate the impact of the 4Rs intervention on the whole student body, and to better understand how whole-school analyses based on data from the entire student body can complement findings from one specific cohort of students.

### **Purpose / objective / research question / focus of study:**

*Description of what the research focused on and why.*

The present study aims to estimate the impact of a social-learning and literacy development intervention (4Rs) on highly policy-relevant academic achievement outcomes using administrative data on children from multiple cohorts. In addition, it attempts to discuss the ways in which whole-school analyses using administrative data can complement cohort-specific analysis and inform efforts to improve the experiences of children in schools. Specifically, we ask:

1. What is the impact of the 4Rs program on the academic achievement (i.e., math and reading scaled scores and performance levels on the math and language arts exams) of children in 1<sup>st</sup> to 5<sup>th</sup> grades, adjusting for a set of individual and school level covariates?
2. Do the impacts of 4Rs vary as a function of child grade (cohort), race/ethnicity, gender, and free/reduced price lunch?

**Setting:**

*Description of where the research took place.*

The evaluation of the 4Rs program took place in 9 intervention and 9 control public elementary schools in low-income neighborhoods in 4 boroughs of New York City.

**Population / Participants / Subjects:**

*Description of participants in the study: who (or what) how many, key features (or characteristics).*

The sample includes all 1<sup>st</sup> through 5<sup>th</sup> grade children who were present in the 18 study schools in October of 2004 (N=7299). The children were 49.6% girls. The majority of children were Black/African American (46.2%) and Hispanic (42.8%), with 5.7% other race/ethnicity and 5.2% non-Hispanic White. In the first year of the study (2004-2005), 33.8% of these children were entitled to free or reduced price lunch (28.9% reduced price lunch, 4.9% free lunch). Children not assigned a specific grade code (1<sup>st</sup> through 5<sup>th</sup>), but who were assigned a special education code instead, were not included in the sample used for these analyses because they could not be assigned to a specific grade cohort.

The average school size from all 18 schools was 656 (SD = 110; R: 412-854) and the average classroom size was 22 (SD=2; R: 20-27).

**Intervention / Program / Practice:**

*Description of the intervention, program or practice, including details of administration and duration.*

The 4Rs program (Reading, Writing, Respect, and Resolution) is an evidenced-based universal violence preventive intervention delivered and evaluated over three-years in elementary schools. The program focuses on changing children's trajectories of problem behaviors and school disengagement and failure through a "balanced literacy" curriculum that promotes conflict resolution skills and social-emotional learning. The program involves a 7-unit, 21 lesson literacy-based curricula in conflict resolution and social emotional learning, each organized around a grade-appropriate book. Each unit involves a book reading and discussion, followed by three to five conflict resolution skill lessons in which children are able to practice specific skills in the context of a larger discussion of the book. The program also involves an intensive professional development component for teachers consisting of a 25-hour introductory training course, followed by ongoing classroom coaching. Finally, a parent component designed to increase family involvement consists of a parent guide (Family Connections) for teachers, a parent-child interactive homework activity for each unit sent home to parents and returned to teachers, and parent workshops.

**Research Design:**

*Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).*

A school-randomized field trial design was used for the evaluation. A set of 18 elementary schools in NYC were matched based on demographic characteristics and were randomized to the 4Rs intervention (9 schools) or to standard practice (9 schools). A cohort of 3<sup>rd</sup> grade children and their parents and teachers was followed over 3-years and 6 waves of data. Baseline data were collected in the fall of 3<sup>rd</sup> grade (2004) and follow-up data were collected in the spring of 3<sup>rd</sup> grade and in the fall and spring of 4<sup>th</sup> and 5<sup>th</sup> grades.

**Data Collection and Analysis:**

*Description of the methods for collecting and analyzing data.*

Administrative data on all children in the 18 schools participating in the 4Rs experimental evaluation were obtained from the New York City Department of Education. Data used in these analyses include individual students' scaled scores and performance levels on the math and language arts exams, eligibility for free and reduced-price lunch, race/ethnicity, gender, annual attendance and suspensions rates, and school and classroom size.

Analyses will be conducted separately for each grade cohort for the following reasons. First, children from different grades were exposed to different doses of treatment. Children from 1<sup>st</sup> through 3<sup>rd</sup> grades at baseline were exposed to 3 years of 4Rs by June 2007; while children who were in 4<sup>th</sup> through 5<sup>th</sup> grade at baseline were exposed to 2 years and 1 year of 4Rs, respectively. Second, because standardized tests are only administered to students beginning in 3<sup>rd</sup> grade, and because children who were in 4<sup>th</sup> and 5<sup>th</sup> grades at baseline were no longer in elementary school at the end of the third year of the study (June, 2007), achievement scores are not equivalently available for all cohorts. Therefore, for the purposes of this analysis, achievement scores at the end of children's exposure to 4Rs will be used. For students who were 1<sup>st</sup> graders at baseline, 3<sup>rd</sup> grade achievement scores will be used. For 2<sup>nd</sup> grade students, 4<sup>th</sup> grade achievement scores will be used; and for 3<sup>rd</sup> grade students, 5<sup>th</sup> grade achievement scores will be used. Each of these scores is based on achievement tests administered in June, 2007. For 4<sup>th</sup> and 5<sup>th</sup> graders, 5<sup>th</sup> grade achievement scores will be used; these scores are based on achievement tests administered in June of 2005 and 2006, respectively. Third, and most importantly, New York City math and language arts exams were modified in 2006, and the scoring system of the new exams is no longer comparable with that of previous years. This means that the standardized scores available for children who were in 1<sup>st</sup> through 3<sup>rd</sup> grades at baseline are not comparable with those of children who were in 4<sup>th</sup> and 5<sup>th</sup> grades. A second set of analyses using performance levels will be conducted in which all cohorts will be combined. This achievement data is equivalent across all years.

A series of two-level hierarchical linear models will be run to estimate the effects of 4Rs on academic achievement and school adaptation (i.e., attendance and suspensions), adjusting for a set of individual-level and school-level baseline covariates. A multi-level strategy is necessary because randomization to receive the 4Rs Program or to remain in a non-intervention control group was done at the school level. Therefore, in this study children are nested within schools and multilevel modeling allows for the simultaneous estimation of variance associated with between individuals and between schools variation based on the specification of fixed- and random-effect parameters (Bryk & Raudenbush, 1992). These analyses will allow us to assess the direct impact of both person-level (e.g., demographic characteristics of children), and school-

level (e.g., intervention vs. control and demographic characteristics of schools) variables on children's academic achievement and student adaptation outcomes.

First, unconditional baseline models in which no predictors are included will be analyzed for each of the outcome variables. Intraclass correlations will be computed for the variance estimates. Next, treatment assignment and school matches will be added to the second level equations, followed by child demographic characteristics (gender, race/ethnicity, free/reduced price lunch eligibility). Each of these covariates will be interacted with treatment to test whether 4Rs is most beneficial to children at highest risk. In the final model, school-level covariates from the 2003-2004 academic year will be added (i.e. average school size, average classroom size, average school attendance rate, average suspension rate).

### **Findings / Results:**

*Description of main findings with specific details.*

Between-school variation in math and language arts standardized achievement scores is significantly different from zero for all cohorts (see Table 1). As shown by the intraclass correlations, approximately 90% of the variance in achievement scores can be attributed to variation between children, while 10% is due to differences between schools. Our proposed analyses will allow us to estimate the proportion of variance that can be explained by treatment and to examine whether treatment differentially affects subgroups of children.

### **Conclusions:**

*Description of conclusions and recommendations based on findings and overall study.*

Preliminary findings suggest that the use of administrative data to evaluate whole-school reform efforts has the possibility of complementing high-cost developmental and survey data. Implications for what this type of analysis strategy can bring to the field of prevention science, including the cost-quality tradeoff and how to marry quality administrative data with a whole-school experimental approach to whole-school reform will be discussed.

## Appendices

*Not included in page count.*

### Appendix A. References

*References are to be in APA version 6 format.*

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*, 899-911.

Borman, G.D., Hewes, G., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*(2), 125-230.

Jones, S.M., Brown, J.L., Hoglund, W., & Aber, J.L. (Revision under review.) Impacts of an integrated school-based social-emotional learning and literacy intervention on third-grade social-emotional and academic outcomes. *Journal of Consulting and Clinical Psychology*.

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*(1), 103-117.

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

**Appendix B. Tables and Figures**  
*Not included in page count.*

Table 1  
*Intraclass Correlations*

Grade	Math			Reading		
	Total Variance	Between Schools Variance	Intraclass correlation	Total Variance	Between Schools Variance	Intraclass correlation
1st	1329.85	169.45**	0.87	1313.72	112.57**	0.91
2nd	1429.29	139.74**	0.90	1279.14	99.75**	0.92
3rd	106.032	109.64**	0.91	1009.72	78.79**	0.92
4th	1142.17	120.99**	0.89	1265.86	110.28**	0.91
5th	1433.09	139.61**	0.9	846.14	95.85**	0.89

\*\* $p < .0001$