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

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 5th 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 1st to 5th 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 1st through 5th 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 (1st through 5th), 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 3rd 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 3rd grade (2004) and follow-up data were collected in the spring of 3rd grade and in the fall and spring of 4th and 5th 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 1st through 3rd grades at baseline were exposed to 3 years of 4Rs by June 2007; while children who were in 4th through 5th 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 3rd grade, and because children who were in 4th and 5th 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 1st graders at baseline, 3rd grade achievement scores will be used. For 2nd grade students, 4th grade achievement scores will be used; and for 3rd grade students, 5th grade achievement scores will be used. Each of these scores is based on achievement tests administered in June, 2007. For 4th and 5th graders, 5th 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 1st through 3rd grades at baseline are not comparable with those of children who were in 4th and 5th 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

	Math			Reading		
Grade	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

**$p < .0001$