Abstract Title Page

Title: Starting Strong: A Randomized Controlled Trial of the Building Assets Reducing Risks (BARR) Model in 9th Grade

Authors and Affiliations: Hinojosa, T., Bos, J., O'Brien, B., Park, S., Liu, F., American Institutes for Research; Corsello, M., University of New England; Jerabek, A., St. Louis Park School District

Background / Context: Students beginning high school commonly experience increased stress and behavior problems alongside declines in grades, attendance, interest in school, and perceptions of academic competence and self-esteem (Alvidrez & Weinstein, 1993; Reyes et al., 2000). Moreover, research indicates that, relative to students who graduate from high school, those who leave school prematurely are more likely to have experienced deeper 9th grade declines (Reyes et al., 2000; Roderick, 1995).

Research also demonstrates positive school climates, positive relationships between students and staff, and among staff, are essential ingredients for turning around low performing schools (Gordon, 2006; National Research Council, 2004; Cohen, 2006; Jerald, 2006; De La Ossa, 2005). More specifically, research is growing on the effectiveness of student-teacher relationships in producing increased attendance and academic performance, and decreased behavior problems. For example, Allen, Pianta, Gregory, Mikami, and Lun (2011) conducted a randomized controlled trial (RCT) in which secondary school teachers were given a year of coaching on effective teaching and student-teacher interactions. After a year of training, students with teachers in the experimental group scored significantly higher on year-end achievement tests than did students in the control group. Quality of student-teacher interaction was a significant mediator of student achievement.

With core components derived from research, the Building Assets Reducing Risks (BARR) model© targets students at this critical juncture in their academic career—9th grade. BARR addresses developmental, academic, and structural challenges in the 9th grade by combining student asset building, teachers' real-time analysis of student data, and intensive teacher collaboration to prevent course failure. BARR develops positive student-teacher relationships and integrates student supports into a school's existing model for addressing nonacademic barriers to learning.

The BARR model was developed more than 15 years ago in a St. Louis Park (MN) high school in response to high failure and dropout rates of 9th grade students. It is currently implemented in 25 schools, in seven states, across multiple geographic regions.

With funding from the Investing in Innovation program, BARR has been rigorously studied using an RCT and is now under scientific investigation in schools of varying locales and geographic regions to better understand the broader context under which BARR impacts students transitioning to the 9th grade. This current study is a multisite randomized controlled trial that explores replication of BARR across three cohorts of schools.

Results from the first RCT demonstrated positive impact on 9th grade students' standardized test scores, credits earned, and overall failure rate. Carried out in a large urban school in 2011–12, 9th grade students who were randomly assigned to participate in BARR earned 0.26 credits ($p<.001$) more than students not assigned to BARR, scored higher on both reading and mathematics standardized tests ($p<.001$), and had fewer course failures than those not assigned to BARR (21 percent compared to 31.9 percent) (Corsello, Sharma, & Jerabek, 2015). Furthermore, qualitative data collected from BARR teachers and counselors suggest staff felt more connected to students, other teachers, and their schools, and they reported higher levels of teacher effectiveness after only one year of BARR implementation.

P**urpose / Objective / Research Question / Focus of Study:** This study, a multisite 11-school RCT, explores the replicability of findings across a broader set of demographic and geographical contexts than previously studied. It is being carried out across three academic years (2014–15, 2015–16, 2016–17) in three cohorts of schools, with each school participating for one year. Because of the within-school random assignment, each cohort has sufficient statistical

power to constitute a stand-alone replication study of BARR. In this presentation, we report findings from the first cohort of schools participating in this trial.

The design of this study allows for causal interpretation of the impact of BARR on two standardized measures of student academic performance: Northwest Education Association's (NWEA) Measures of Academic Progress (MAP) scores in mathematics and reading. In addition, we present results on the effects of BARR on credit accumulation, course failure rates, and students' and teachers' self-report ratings on proximal outcomes of student achievement, for those involved in the BARR program relative to those not involved. The study is designed to answer the following research questions:

1. What is the impact of BARR on 9th grade students' academic achievement as measured by NWEA's mathematics and reading tests?
2. What is the impact of BARR on 9th grade students' educational attainment as measured by the percentage of credits earned in core subjects?
3. To what extent do ratings of proximal measures of student achievement (e.g., student behavior, teacher self-efficacy, use of data) differ between BARR and non-BARR teachers?
4. To what extent do students' self-ratings of proximal measures of student achievement (e.g., engagement, sense of belonging, grit) differ between BARR and non-BARR students?

Setting: This phase of the study was carried out in three high schools (with approximately 1,000 9th grade students): two schools in California, one urban and one suburban, and one rural school in Maine. The schools in California are predominantly high poverty (80 and 92 percent eligible for free or reduced price lunch) and comprised in large part of minority students (73 and 95 percent).

Population / Participants / Subjects: Participants in the study included all 9th grade students at each school, with the exception of students who receive core-subject instruction in self-contained classrooms (e.g., students with severe learning disabilities, some English language learners). The demographic composition of 9th grade students in the study mirrored that of the schools. Across the first cohort, 72 percent are minority, 9 percent are classified as English language learners, 7 percent are eligible for special education services and 74 percent are eligible for free or reduced price lunch.

Also participating in this study were teachers, counselors, principals and assistant principals. In total 19 teachers, 5 counselors, and 3 assistant principals, were trained in the BARR method. Principals also participated in the study through various mechanisms of involvement (e.g., attending team meetings). In addition, data were collected from a broad sample of teachers who did not participate in any BARR activities in order to contrast treatment and control conditions with regard to teacher experience.

Intervention / Program / Practice: The BARR model consists of eight strategies that are interconnected and function as a whole. Each of these strategies is a necessary component of the model. They include: 1) relationship-building professional development for teachers, counselors and administrators – onsite two day training and ongoing weekly and monthly meetings; 2) restructuring the high school course schedule to enable groups of students to be shared by a team of teachers; 3) contextual support (focused on leadership); 4) parent involvement to support high school reform; 5) whole student emphasis in instructional reform; 6) developmental assets curriculum (I-Time); 7) teacher and staff block meetings with collaborative problem solving; and 8) risk review for persistently failing students.

Research Design: For this study, 9th grade students were randomly assigned, within schools, to one of two groups (i.e., treatment or control) expected to be, on average, equivalent with respect to things that can be measured, such as baseline achievement levels, race/ethnicity, and socioeconomic status. Prior to the start of the school year, the research team randomly assigned eligible 9th grade students to receive BARR or not. Once assigned, school administrators were tasked with creating master school schedules such that students in the BARR condition would receive instruction from at least three core academic teachers trained in the BARR method. Students in the control condition would receive instruction from teachers not trained in BARR.

Data Collection and Analysis: Demographic and outcome data were extracted from school databases by a school-based BARR coordinator and submitted directly to the research team at three points in time: in the summer prior to the study, in the fall, and in the spring of the study year. Survey data were collected electronically or in some cases on paper.

Extant data collected included: student demographic data on race or ethnicity, ELL status, special education status, free or reduced price lunch status, and gender; student standardized achievement data including NWEA MAP reading and mathematics test scores administered in the fall and spring to all ninth-grade students; student achievement data including fall and spring semester credit data for core courses (e.g., English, mathematics, science, social studies) and fall and spring student grades. Student data were individually identified with a unique ID so that data could be linked across all data sets including a student survey (discussed below).

Two surveys were used to collect teacher and student measures of perceptions, attitudes, and behavior often linked to students' academic success. The teacher survey was constructed from several published scales and included 55 items that captured teachers' perception on eight constructs: thoughts about actual student behavior; perception of student behavior; views of the school; parent-teacher interactions; teacher self-efficacy; student accountability; interactions with colleagues; and use of data. The student survey was also constructed using published scales and included 46 items across six constructs measuring: expectations and rigor; student engagement; supportive relationships; socio-emotional learning; sense of belonging; and grit.

Interviews were carried out with 33 teachers (14 control and 19 treatment) using a semi-structured protocol and designed so that both treatment and control teachers could respond about their experiences teaching 9th grade. For example, teachers were asked about their use of data, professional development they had received, the structuring of 9th grade classes, and collaborative efforts among teachers, to name just a few. In addition, all teachers were asked to provide a rating related to the impact on their teaching of the ideas contained within the 8 core strategies that comprise BARR (e.g., professional development, contextual support).

We estimated impacts for students assigned to receive BARR services relative to those not assigned on a number of outcomes including NWEA mathematics and reading scale scores; credit accumulation; and five survey constructs. Survey data were psychometrically evaluated and fit with the Rasch model yielding interval scale scores for each student, for each construct. Specifically, we fit each outcome model with an ordinary least squares, covariate adjusted, regression. Each model controlled for a student's starting academic skill level (math or reading assessed via the fall administration of NWEAs), demographic identifiers including race/ethnicity, gender, special education, ELL, and free or reduced price lunch status, a school level indicator, and a treatment indicator. A dummy variable approach was used to account for missing data (set to 0 for missing), which was only missing for a very small percentage of pretest data.

Findings / Results: Prior to analysis of outcome measures, both baseline equivalency and attrition were explored. Baseline equivalence tests suggest that random assignment was successful in creating groups that were functionally equivalent on observed measures. No

statistical differences were found on any of the pretest or demographic measures between the treatment and control group. Calculation of overall and differential attrition rates for standardized tests yielded results that suggest this study would meet What Works Clearinghouse Standards without reservation (http://ies.ed.gov/ncee/wwc/default.aspx). Overall attrition for the NWEAs was approximately 25 percent for each test and differential attrition was 1.6 percent for the reading test and 3 percent for the mathematics test. Moreover, the majority of the overall attrition is accounted for by a single factor exogenous to study implementation (students never enrolled in the study schools). Data on credit accumulation has not been finalized at the time of writing this, but it is expected that similar baseline equivalency and attrition findings will emerge for these data.

The analytic results of BARR's impact on NWEA's reading and mathematics achievement include pre- and post-test scores, which are standardized with a mean of 0 and a standard deviation of 1. Thus, impact estimates can be interpreted as effect sizes (ES) which are expressed in standard deviation units. Results indicate that BARR had a statistically significant effect on reading achievement (ES=.11; $p<.05$) but not on mathematics achievement (insert tables 1 and 2 here). A full analysis of credit accumulation could not be completed at time of this writing but a preliminary examination of course failure rates suggest BARR positively impacts the percentage of courses successfully completed. In the study year a greater percentage of BARR students (61.5 percent) passed all of their fall and spring semester core courses (i.e., ELA, mathematics, and science) compared to non-BARR students (46.9 percent). A Pearson Chi Square test indicated that this group difference was significant, $X^2_{(1,\, n = 1053)} = 22.70$, $p<.001$. The remaining percentage of students (38.5 percent BARR, and 53.1 percent non-BARR students) failed at least one core course during the fall or spring semester of their 9th grade year.

Additional analyses will be performed to explore the impact of BARR at the three different locales (i.e., rural, urban, and suburban) and for various subgroups of students (e.g., ELLs, low income students).

Preliminary results of student survey data suggest students in the BARR condition provide on average statistically significant higher ratings on a measure of supportive relationships ($p<.001$). Analysis of the teacher survey data show that BARR teachers provide on average significantly higher ratings on three survey scale measures: the ways in which teachers think about students' actual behavior, commitment, and attitudes ($p<.05$); how teachers collaborate with and view their colleagues ($p<.05$), and their use of data ($p<.05$) (insert table 3).

Interview data with teachers and the site coordinator suggest the BARR model was implemented with high fidelity in each school. Most teachers participating in BARR report that they perceived a positive impact on teacher-student, teacher-teacher, and student-student relationships. Comparing the self-reported ratings on the impact of BARR on their effectiveness as a teacher this year, BARR teachers had significantly higher ratings than non-BARR teachers on two measures: ninth-grade structure ($p<.01$) and contextual support ($p<.001$) (insert table 4).

Conclusions: Given the importance of a successful transition to high school and of forming positive school relationships, these statistically significant findings that BARR impacts students' academic achievement and socio-emotional development, suggest deeper exploration of this model is warranted. Specifically, additional replication studies, student sub-group analyses, and potential impact differences for various locales will need to be carried out in order to fully understand the conditions under which, and for whom, BARR can be impactful.

Appendix A. References

Allen, J.P., Pianta, R.C., Gregory, A., Mikami, A.Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*, 1034-1037.

Alvidrez, J., & Weinstein, R.S. (1993). The nature of "schooling" in school transitions: A critical re-examination. *Prevention in Human Services, 10*, 7-26.

Cohen, J. (2006). Social, emotional, ethical and academic education: Creating a climate for learning, participation in democracy, and well-being. *Harvard Educational Review 76(*2), 201-239.

Corsello, M., Sharma, A., & Jerabek, A. (2015). Successful transition to high school: A randomized controlled trial of the BARR Model with 9th grade students. Paper presented at the Society for Research on Educational Effectiveness, Spring Conference, Washington, D.C. Mar 5-7.

De La Ossa, P. (2005). "Hear my voice:" Alternative high school student's perceptions and implications for school change. *American Secondary Education 34(*1), 24-40.

Gordon, G., & Crabtree, S. (2006). *Building engaged schools: Getting the most out of America's classrooms.* New York: Gallup Press.

Jerald, C. (2006). *School Culture: The Hidden Curriculum.* Center for Comprehensive School Reform and Improvement Issue Brief. Washington, D.C.: Learning Point Associates.

National Research Council. (2004). *Engaging schools.* Washington D.C.: National Academy Press.

Reyes, O., Gillock, K., Kobus, K., & Sanchez, B. (2000). A longitudinal examination of the transition into senior high school for adolescents from urban, low-income status, and predominantly minority backgrounds. *American Journal of Community Psychology, 28*, 519-544.

Roderick, M.R. (1995, December). Grade retention and school dropout: Policy debate and research questions. *Phi Delta Kappa Research Bulletin, 15*, 1-6.

Appendix B. Tables and Figures

Table 1. NWEA Reading Impact Analysis

Coefficients	Estimate	S.E.	*t*
Intercept	0.35	0.13	2.78**
School2	-0.10	0.14	-0.75
School3	-0.39	0.08	-5.05***
Treatment	0.11	0.05	2.14*
Pretest Reading Z-score	0.58	0.04	15.83***
Pretest Reading Dummy Variable	-0.04	0.08	-0.56
Female	0.02	0.05	0.31
Minority Status	-0.05	0.08	-0.65
English Language Learner	-0.75	0.11	-7.01***
Special Education	-0.51	0.11	-4.61***
Free and Reduced Price Lunch	-0.19	0.10	-1.94

Note. * = $p \leq .05$, ** = $p \leq .01$, *** $p = \leq .001$. S.E. = Standard Error.

Table 2. NWEA Mathematics Impact Analysis

Coefficients	Estimate	S.E.	*t*
Intercept	0.41	0.12	3.28***
School2	-0.07	0.14	-0.48
School3	-0.13	0.09	-1.49
Treatment	0.01	0.05	0.19
Pretest Math Z-score	0.75	0.04	20.89***
Pretest Math Dummy Variable	-0.20	0.08	-2.36*
Female	-0.10	0.05	-2.04*
Minority Status	-0.06	0.08	-0.76
English Language Learner	-0.64	0.10	-6.55***
Special Education	-0.59	0.10	-5.75***
Free and Reduced Price Lunch	-0.15	0.09	-1.64

Note. * = $p \leq .05$, ** = $p \leq .01$, *** $p = \leq .001$. S.E. = Standard Error.

Table 3. Teacher Survey Results

	Scale Score Means (SD)			
	BARR	Non-BARR	*t*	*df*
Teachers View Students' Actual Behavior	0.28 (2.38)	-1.92 (2.50)	2.55*	32
Teachers Perception of Students' Behavior	1.69 (2.83)	0.11 (3.81)	1.30	31
Teachers View of the School	3.92 (2.96)	2.35 (3.86)	1.29	32
Teachers and Parents Interaction	-0.74 (1.70)	-0.57 (3.80)	-0.17	32
Teachers Self-Efficacy	2.49 (2.01)	1.32 (2.37)	1.51	32
Teachers View of Student Accountability	0.48 (1.05)	0.39 (0.97)	0.26	32
Teachers Cooperate and View Colleagues	3.83 (2.26)	1.25 (3.15)	2.66*	32
Teachers Use Data	2.78 (2.58)	0.80 (2.28)	2.29*	32

Note. * = $p \leq .05$. SD = Standard Deviation.

Table 4. Teacher Impact Ratings on Effectiveness

	Means (SD)			
	BARR	Non-BARR	*t*	*df*
Professional Development	3.68 (0.95)	3.14 (1.10)	1.52	31
9th Grade Structure	3.95 (0.78)	2.57 (1.34)	3.43**	19.4
Whole Student Approach	4.00 (0.88)	3.43 (0.94)	1.79	31
Block/Team Meetings	4.53 (0.70)	3.79 (1.25)	2.00	18.9
I-Time/Youth Development	4.16 (0.96)	3.57 (1.28)	1.50	31
Risk Review/Student Referral	3.21 (1.36)	2.36 (1.39)	1.77	31
Contextual Support	4.53 (0.70)	2.64 (1.34)	4.13***	16.66
Parent Involvement	3.21 (1.13)	2.64 (1.34)	1.32	31

Note. ** = $p \leq .01$, *** $p = \leq .001$. SD = Standard Deviation.