Journal of Education and Human Development
June 2014, Vol. 3, No. 2, pp. 121-154
ISSN: 2334-296X (Print), 2334-2978 (Online)

Published by American Research Institute for Policy Development

A Study on Immediate Effects of the First-Year Intervention Under the Illinois Striving Readers Project

Dimiter M. Dimitrov[1], **Michael Frye**[2], **Sonia Jurich**[3], **Sarah Sayko**[4] **and Jill Lammert**[5]

Abstract

Under the Illinois Striving Readers (ISR) project, the *Passport Reading Journeys*III (*PRJ III*) is a supplemental reading intervention for students in ninth grade who were reading below grade level. The main goal of the study described in this paper was to evaluate immediate effects of the *PRJ III* intervention at the end of the first implementation year (2010/11). The research questions related to this goal are addressed through the use of hierarchical linear modeling. Although the results do not provide evidence of *PRJ III*intervention effects at the end of the first year, the study findings can provide a valuable feedback to the ISR and other projects or meta-analytic studies on *PRJIII* interventions.

Keywords: reading intervention, research design, hierarchical linear modeling

Currently, there is a collective effort underway in America to ready students for college and career after high school graduation (President's Council of Advisors on Science and Technology, 2010). To accomplish this goal, students must be adequately prepared to meet the increased literacy demands they will face in middle and high school and have the necessary supports and strategies in place for overcoming learning barriers at these levels.

[1] PhD, Professor, Graduate School of Education, George Mason University, MS 6D2, 4400 University Drive, Fairfax, VA 22030-4444. Email: ddimitro@gmu.edu, Phone: 703-993-3842, Fax: 703-993-2063
[2] Abt Associates, 4550 Montgomery Avenue, Suite 800 North, Bethesda, MD 20814. Email:michael_frye@abtassoc.com, ph: (301) 347-5579, fax: (617) 386-8436
[3] RMC Research Corporation, 1501 Wilson Blvd Suite 1250, Arlington VA 22209. ph: (703) 558-4809, fax: (703) 558-4823, Email:jurichs@rmcarl.com
[4] RMC Research Corporation, 1501 Wilson Blvd Suite 1250, Arlington VA 22209**.** ph: (703) 558-4830**,** fax: (703) 558-4823, Email: saykos@rmcarl.com
[5] Westat , 1600 Research Blvd RW2604, Rockville MD 20850, ph: (301) 610-8855. Email: JillLammert@westat.com

In reality, many adolescent readers struggle to master the most basic of literacy skills. According to the National Center for Education Statistics (NCES), more than 8 million students in grades 4-12 are struggling readers (USED, 2003). Reading achievement data from the National Assessment of Educational Progress (NAEP) indicate that results from the assessment have remained unchanged over time; in 2011, only 34 percent of eighth graders scored at or above proficiency (NCES, 2011). Carnevale (2001) found that high school students in the lowest 25 percent of their class are 20 times more likely to drop out than their highest performing peers. Of those who graduate, approximately 25 percent enroll in literacy remedial courses in postsecondary education (NCES, 2003).

Research indicates that the most serious challenges faced by struggling adolescent readers are difficulties with decoding, slow and labored reading, lack of background knowledge, and limited vocabulary (Beck, McKeown,& Kucan, 2002). Many of the poor and minority students who currently perform below the 30th percentile in reading skills entered school with academic vocabularies already only half the size of their middle-class counterparts (Beck, McKeown, & Kucan, 2002). Additionally, they have fewer opportunities to acquire the kinds of active reading comprehension strategies that become increasingly important as text increases in size and complexity after third grade (Torgesen, 2005). As a result, these deficits interfere with struggling adolescent readers' ability to develop higher level literacy skills and cause cognitive loads that are too much for most students to bear.

A crucial component to addressing struggling adolescent readers' needs is improving literacy instruction and intervention practices in schools. This includes providing explicit vocabulary instruction, direct and explicit comprehension strategy instruction, opportunities for extended discussion of text meaning and interpretation, and increasing student motivation and engagement in literacy learning (Kamil, Borman, Dole, Kral, Salinger, & Torgesen, 2008). In addition, intensive and individualized interventions for those who struggle most must be available and delivered by trained specialists (Kamil et al., 2008).

These interventions should target students' instructional needs, occur in small group settings during extended learning periods, happen with increased frequency, and incorporate opportunities to monitor student progress.

Alternatively, research suggests that it is important to avoid intervention programs that have insufficient intensity, weak instruction in word study skills, and little or no direct instruction in comprehension strategies, as they have limited effectiveness (Scammacca, Roberts, Vaughn, Edmonds, Wexler, Reutebuch, & Torgesen, 2007). Because interventions for students' still mastering basic reading skills tend to stabilize rather than remediate the relative reading deficiency (Scammacca et al., 2007), careful selection of research-based adolescent reading programs are essential.

The Illinois Striving Readers Project

Addressing research findings regarding the needs of adolescent who are struggling readers, the U.S. Department of Education (USED) started the Striving Readers program. The program had two main goals: (a) to address the challenges of improving reading skills for middle and high school students who were reading below grade level; and (b) to build a scientific base to identify effective strategies that improve adolescent literacy skills. Striving Readers was geared to Title I eligible schools that had significant percentages of students reading below grade level and/or schools that were not meeting or at-risk of not meeting adequate yearly progress (AYP) requirements under the *No Child Left Behind Act* (NCLB).

The program, which reflected a joint effort from the Office of Elementary and Secondary Education (OESE) and the Institute of Education Sciences (IES), included three key components: (a) supplemental literacy interventions targeted to students who were reading "significantly below grade level;" (b) cross-disciplinary strategies for improving adolescent literacy, including professional development and research-based reading and comprehension strategies; and (c) a required evaluation component using an experimental design (USED, 2008). In 2009, USED published a Request for Proposal for a second cohort of the Striving Readers grantees. The second cohort competition requested that states put together a coalition of schools that crossed school district boundaries and focus solely on interventions for struggling readers. The Illinois State Board of Education (ISBE) applied for and was awarded one of the eight Striving Readers grant funded as part of the second cohort competition.

The Illinois Striving Readers (ISR) Project had two purposes: (a) implement a supplemental reading intervention for students in ninth grade who were reading below grade level, and (b) study the impact of the intervention on students' performance on standardized assessments using a randomized control trial design. After reviewing existing research-based supplemental adolescent reading programs, ISBE staff selected *Passport Reading Journeys*III (*PRJ III*), published by Cambium Learning Group, as the intervention for the project. The ISBE invited Title I high schools to participate in the study. Of the schools that volunteered, the ISBE selected six high schools in four districts across Illinois to participate in the ISR Project. Although demographic characteristics at the schools varied, all had a high percentage of minority and low income students, and mobility rates were substantial. All 9th grade students in these high schools who were indicated as struggling readers based on reading assessments were eligible to participate in the study.

Table 1: Characteristics of Participating Schools

School	Enrollment (N)	Minority (%)	Low Income (%)	Mobility Rate (%)
School 1	1,296	53.5	57.4	35.7
School 2	1,112	51.1	45.2	25.3
School 3	1,220	40.9	66.1	35.2
School 4	1,606	47.6	53.5	29.8
School 5	1,185	80.1	72.5	22.0
School 6	1,149	53.4	52.7	37.7

Published by Cambium Learning Group, *Passport Reading Journeys* (*PRJ*), is a supplemental reading intervention that incorporates whole group, small group, computer-aided instruction, and individual instruction to support adolescent struggling readers. Across grade levels, the intervention maintains the same structure but the content and reading level change. *PRJ III,* the intervention implemented in the ISR project, is geared for students in grade 9. *PRJ III* is based on reading research and research in learning, including works from Baker, Simmons, & Kame'enui (2004), Beck, McKeown, & Kucan (2002), Biancarosa, & Snow (2006), Deshler, Palincsar, Biancarosa, & Nair (2007), Gersten, Fuchs, Williams, & Baker (2001), Graham & Perin (2007), Marzano (2004), Mastropieri, Scruggs, & Graetz (2003), Scammacca, Roberts, Vaughn, Edmonds, Wexler, Reutebuch, and Torgesen (2007), and Schatschneider, Buck, Torgesen, Wagner, Hassler, Hecht, & Powell-Smith (2004).

The intervention encompasses daily, 50-minute lessons that provide explicit, systematic instruction in critical reading skills on a topic related to science or social studies. The lessons are organized in Expeditions for a total of 14 Expeditions that are taught within one school year. Each Expedition is organized in two week, ten-lesson routines that mix teacher-led instruction and students' independent practice. Lessons one, three, six, and eight are organized around whole-group instruction in which students are introduced to new vocabulary and a new reading passage. After whole group instruction, students can individually practice vocabulary using the online technology component (*VocabJourney*) and/or selectbooks for independent reading. Lessons two, four, seven, and nine include whole-group review of the previous day's instruction and the opportunity for students to re-read the passage to build fluency, independently or with a partner. During this period of independent or small-group structured practice, the teachers work individually with students as needed. Students spend lessons five and ten of the Expedition on independent or paired practice in a variety of activities intended to review, extend, or assess previous learning. Teachers may do any or all of the activities selected from a menu provided by the publisher. The choice is expected to reflect students' learning needs. Since re-teaching may be necessary, lessons five and ten are intended to extend across multiple days to allow teachers to adequately address individual student needs. Cambium research personnel, in consultation with the ISR Director, set a two day limit for lessons five and ten to promote implementation consistency across the state. Therefore, in this project, the teachers were expected to complete the ten lessons that comprise each Expedition within twelve days.

PRJIII includes a number of assessments that are used for diagnostic and performance monitoring tools. Reading Benchmark (RB) I is administered before instruction starts and is used to place students with the appropriate level of reading materials, while RBII and III are used for monitoring and informing instruction. Comprehension and vocabulary assessments areadministered during lessons 5 and 10 of each Expedition, while online vocabulary technology self-assessments are used by the students to track their own progress on vocabulary, comprehension and content-specific text. *PRJIII* also includes semester exams. These4 are criterion-referenced assessments that focus on mastery of skills and content taught, and are administered online at the end of Expeditions 7 and 14.

Cambium Learning Group offers all new users two days of launch training intended to prepare teachers to implement the intervention with fidelity. Participants learn about the intervention, and are instructed in specific practices, such as administering the assessment measures, grouping students, setting up their classrooms, structuring small and large group instruction, and using intervention materials. The training includes time for practice of lesson delivery and instruction in *VocabJourney*. Training on the Voyager data management system (VPORT) and classroom management are also included. During the launch training, teachers receive a DVD showing footage of classroom instruction, illustrations of program features, and the measures to practice administering and scoring the assessments. Tutorial booklets introduce the key features and components of the program, present sample lessons at each grade level, review the assessment component, and provide suggestions for managing time and working with students with special needs.Available to all users are online training modules that cover topics such as curriculum, classroom management, assessment, and implementation, and provides links to a library of video segments. At the conclusion of each section in the module, the teachers take a quiz to check knowledge gained. They can redo the modules to improve knowledge, or come back to them later to refresh information.Further supports, including one-on-one mentoring and online modules on adolescent literacy require extended contracts. The ISR project contracts for five days of one-on-one support from Cambium experts, in addition to support from the project coordinator, an adolescent reading specialist.

Despite the program popularity, most research on Journey's has been done with the elementary level programs. Two studies were found that focused on middle or high school students.Denton (2008) used a quasi-experimental design to compare students receiving *PRJ III*(print and electronic) instruction to students matched for demographics and reading abilities whowere instructed with traditional (print only) curricula. The sample comprised approximately 200 students who attended grade 9 in the Dallas Independent School District. The study found that*PRJ*students made statistically significant gains when compared to their non-*PRJ*peerson both a norm-referenced test (Iowa Test of Basic Skills) and a criterion-referenced test (TexasAssessment of Knowledge and Skills). Shneyderman (2006) conducted a large-scale quasi-experimentalstudy involving approximately 1,400 students in Miami-Dade, Florida. He used apropensity score to select comparison schools; within these schools, students were randomlyselected for the comparison group.

Journeys students showed small positive gains on the FloridaComprehensive Assessment Test (FCAT) when compared to their peers. A meta-analysis of effective adolescent reading programs (or curricula) reviewedresearch on more than 100 adolescent reading programs and identified 14 programsthat had some evidence of effectiveness, with no program showing strong evidence (Slavin et al,2008). *Journeys* was classified as having limited evidence of effectiveness with effect sizes of0.22 and 0.12 in grades 9 and 10, respectively (Slavin et al., 2008).

Study Goal and Research Questions

The goal of this study was to investigate theimpact of the *ISR* project on student readingachievementas measured by the Grade 9 Gates-MacGinitie Reading Test (GMRT) 4th editionand Grade 9 EXPLORE®. Of primary interest was to assess the overall average of the immediate (pretest-posttest) effects directly targeted with the treatment (*PRJ III*) experimental design of the *ISR*study and the variability of these effects across schools.. Specifically, the following two research questions were addressed.

RQ1: What is the overall experimental (pretest-posttest) effect of the *PRJ III* supplemental literacy intervention on reading achievement of grade 9 students at the end of the first implementation year (SY 2010/11)?

RQ2. Do the experimental (pretest-posttest) effects of the *PRJ III* supplemental literacy intervention on reading achievement of grade 9 students at the end of the first implementation year (SY 2010/11) vary across schools?

Method

Participants

The ISRstudy involved six high schools in four school districts across Illinois. Allschools were Title I-eligible schools that had not made, or were at-risk of not making, adequate yearly progress requirements under the *No Child Left Behind Act* of 2001. Incoming 9th grade students identified as struggling adolescent readerswere eligible for the study. The identification criterion was scoring at the two lowest quartiles on grade 8 EXPLORE® reading assessment.

This criterion was deemed appropriate after inspection of EXPLORE® reading assessment score distributions indicated that these scores were normally distributed at individual schools and across schools.Students were excluded from the study if they had Individualized Education Plans (IEP) that precluded their participation in the study, or if their parents requested that they not participate.Of the 1,985 incoming ninth grade students across the six schools, a total of 855ultimately qualified as eligible and participated in the study. These students were randomly assigned to *PRJ III* treatment and control groups.

Random Assignment

The random assignment of students to intervention and control groups was conducted after obtaining the lists of eligible students with exclusions applied from the schools. A total of 855 students were included in this final list. The process for the random assignment included three steps. First the eligible students were assigned to pairs, with the students in each pair matched on relevant characteristics: EXPLORE® reading score, Limited English Proficiency (LEP), free and reduced meals (FARM) eligibility, special education status (SPED), gender, and ethnicity. Students were matched on as many criteria as possible. This matched pair strategy ensures that the intervention and control groups are balanced on observable characteristics. Hence, we minimize the threat of these observable characteristics moderating the effect of the intervention and avoid the inclusion of numerous interaction terms in our impact models. Then the students from each pair were randomly assigned to two different groups using a Bernoulli distribution function with $p = 0.5$. The process was completed by randomly assigning the two groups to the two treatment conditions (intervention or control). The resulting intent-to-treat (ITT) group, comprised of 855 students, represented 43 percent of the total enrollment across the six schools.

Lists of students assigned to treatment and control group were provided to the schools, with treatment students attending *PRJ III* while control students were to enroll in electives unrelated to reading supplemental intervention. ITT group sizes were similar across schools.

Table 2: Number of Students by Participating Schools

School	Treatment Condition		Non-Eligible	Total Enrollment
	PRJ III	Control		
School 1	87	88	240	415
School 2	54	54	187	295
School 3	77	75	163	315
School 4	63	64	173	300
School 5	62	62	161	285
School 6	84	85	206	375
TOTAL	427	428	1,130	1,985

Demographics of ITT group

The demographic makeup of the ITT group was primarily African American (58%). White students constituted 30 percent of the ITT group, while other races (American Indian, Asian, Hispanic, mixed, and other) represent 12 percent. The ITT group contained slightly more males than females, and the majority of students (85%) qualified for free and reduced lunch, which is used as the proxy for low socioeconomic status. The matching procedure used in the random assignment process ensured that the demographic percentages were equally distributed between the treatment and control groups.

Table 3: Demographic Breakdown of Study Participants

Subgroups		Percentage
Race/ Ethnicity	African American	58.0
	Hispanic	5.0
	White	30.0
	Other	7.0
Gender	Male	56.7
	Female	43.3
Free and Reduced Lunch		84.7
Special Education		17.6

Note. Total *N* = 855.

Description of Intervention

The implementation of *PRJIII* was to follow the model established by the developer. Each of the six participant schools hired one full time teacher to provide the lessons. The hiring of theseteachers followed the process used by the school districts to hire their regular teaching staff. The position was announced in local newspapers, and the applicants were interviewed by a panel that included the school principals, who made the final decision. Theteachers were required to have a valid Illinois teaching license with a reading endorsement, two to three years of teaching experience, an understanding of the Response to Intervention process, and proven classroom management skills. As part of the contract, the intervention teachers were required to dedicate at least 80 percent of their time to the intervention, attend all of the professional development activities related to the intervention, and implement *PRJ III* with fidelity.

Before the school year started, each teacherreceived individually eight hours of launch training from the publisher, and had available the online modules. Another eight hour training that dealt mostly with assessments and data analysis was provided about one month later. The teachers also received one-on-one mentoring from Cambium specialists, in addition to the support from the IRS project coordinator.

The publishers worked with IRS staff to adapt the 50-minute lessons for schools with 90-minute block scheduling. At these schools, teachers were expected to present two lessons in one day, thus modifying the expected 50 minute to 45 minutes for each lessons. To accommodate for the time, Cambium suggested eliminating some of the writing exercises within each task in the Expedition, placing the focus on the reading process rather than the writing material. Another suggestion was to use at least one of the extra days in lessons 5 and 10 to initiate a new lesson or start a new Expedition if students were showing strong results in the lesson-specific assessments.

Description of control group

Students in the control group were enrolled in elective classes that did not provide supplemental reading instruction, such as arts or foreign languages. Enrollment data for the *PRJ III* classrooms were monitored to ensure that control group students were not receiving the treatment.

Fidelity of Implementation

The evaluation of the ISR implementation aimed to assess how close the implemented intervention was to the *PRJ III*model. The implementation study was to be used as a descriptive tool to further the understanding of the findings from the impact study. The intervention's structure provided the framework upon which the research questions and the development of instruments for data collection were built. As previously discussed,*PRJ III* is a highly-structured intervention in which the teachers are expected to follow a scripted guide that details what, how,and when they teach. Diversions from the model are not welcomed, except for minor adaptations to adjust the required pacing within allocated classroom time.

The evaluators worked closely with Cambium's research department to define fidelity of implementation and ensure that the site visit rubric reflected the conceptual framework and format of *PRJ III*. The rubric included four components. The first component, Section A –classroom environment, provided a descriptive overview of classroom size, desk arrangements, technology elements, and materials required for the intervention. Sections B and Section C focused on the quality and amount of instruction and use of differentiation strategies. Section B provided an overview of the lesson's structure, while Section C was lesson-specific, thus the template changed according to the lesson number within the Expedition. Section D focused on the classroom management component. Since the observers would not be present during assessment time, all information from the assessments was obtained through VPORT.

Qualitative data were collected through interviews with the Voyager Implementation Specialists (who provide one-on-one mentoring to the teachers), the intervention teachers, and the ISBE Project Director. The teachers were contacted monthly by telephone on in-person, while the other staff was to be contacted twice – at the middle and end of the project. Between September 2010 and May 2011, a total of eight interviewswere completed for each teacher.

Additionally, four classroom visits were planned during the implementation year. The evaluation team who observed the *PRJ III* classrooms was comprised of an evaluator with a reading background, who had been trained on the intervention, and an evaluator with a methods background. The evaluation plan included four one-day visits to each school, at the beginning, middle, and end of the school year.

For each of the visits, the evaluators would conduct alternating half-day observations at each school (e.g. on day one, school A was observed during the morning and school B was observed in the afternoon; the following day, school A was observed during the afternoon and school B during the morning, and so on). Classrooms were visited during whole-group instruction portions of the lesson, as observation of lesson delivery during these portions is when adherence to the *PRJ III* model is directly observable. This process allowed for observing instruction at different times of the day and on different days of the week to cover the ten lessons routine that characterizes each Expedition. The evaluators took extensive notes during the interviews and used categorizing and connecting process for the analysis (Maxwell & Miller, 2008). Data were coded thematically to reflect the *PRJ III* components highlighted by the developer in its Index of Fidelity of Implementation (IFI). These components included: amount of instruction, quality of instruction, classroom management, use of assessment, and differentiation. Information from the interviews was used to understand the process of implementation in the different schools from the different actors' perspectives, including perspectives on the barriers and facilitators to implementation. Although the evaluators wanted to make unannounced visits, the complex reality of school scheduling made that impossible and eventually, the visits were coordinated with the teachers.

The first school visit occurred in October 2010, at the end of the first month of implementation, and the second visit occurred in February 2011, as the implementation entered its sixth month. A total of 27 observations were conducted.Two other visits were scheduled – one for mid-March and the other for the end of May. However, by April, The Congress eliminated funds for the Striving Readers grant. With the recommendation from USED to focus on the impact study, the two final visits were cancelled.

The process to calculate the fidelity score for the classroom model was as follows: (1) all observation rubrics completed by each of the evaluation team were entered into the observation database; (2) the different observations for each evaluation team member were combined to get an average score for each lesson; and (3) the scores for both rounds of observations were then combined to get an average score for each teacher. Because each school had only one teacher, the teacher score equals the school score. Weighting each Section of the rubric was done in agreement with Cambium's IFI rubric.

Table 3 displays the final model for calculating the index of fidelity of classroom implementation. Based on feedback from Cambium, the evaluators established the following fidelity levels:scores below 0.70 were defined as inadequate or low implementation, scores between 0.70 and 0.89 reflected medium fidelity; scores of 0.90 or above were considered high fidelity.

Table 4: Calculating the Classroom Implementation Fidelity Score

Section	Weight		Section Score	Total Possible Weighted Score
A	.20		$X_A/6$	.20
B	.30		$X_B/(12 - \text{number of N/A})$	.30
C	.30	x	$X_C/8$	.30
D	.20		$X_D/(\text{total time intervals} - \text{number of N/A})$	.20
Total possible score				1.00
Levels: 0.0 – 0.69 = low			0.70 – 0.89 medium	0.90 – 1.0 = high

The fidelity for the professional development model was calculated as hours of professional development attended divided by hours provided. All teachers completed the required professional development and this index will not be discussed in this paper for reasons of space.

The teachers were assigned permanent classrooms that were expected to have enough space to conduct whole group and small group instruction. Computers and DVD projectors were available for instruction. During the site visits, the evaluators deemed that only one school had a class that was too small to accommodate small group instruction. All other classrooms were deemed adequate in terms of space for instruction.*PRJ III* class size varied from 7 to 16 students, with an average of 12.7 students per class.

The actual intensity of the intervention the students received was influenced by three elements: (a) actual instruction time; (b) actual days of instruction; and (c) student attendance (or the need to re-teach for absent students).Actual instruction time was defined by the ratio between allotted classroom time and actual time the teachers and students were involved in instruction. All six schools used a 90-minute period for the intervention.

However, classes that happened during the first period tended to be curtailed by announcements, and those in the last period could be shortened for early class dismissal, assemblies, or meetings. Interruptions due to student behavior would further deplete from the allotted instruction time and may explain the difficultly that some teachers had in complete one lesson per day. During the two site visits, the evaluators were able to observe four lessons in schools 1, 3 and 4; five lessons in school 5; and six lessons in school 2. Two of the teachers were able to complete the expected number of lessons during the observation, while the others had varied success in completing their lessons.

Information on the second component – actual days of instruction – wasexamined during the monthly interviews with the teachers, who were asked to report the number of days the school was closed as well as the number of classes that were cancelled during that month. Across the year, students missed between 26 to 35 days of instruction, either due to school closures or classes cancelled. The list of closures included holidays, teacher working days, and inclement weather. Reasons for class cancellations included professional development days, testing, and school-wide events, including "pep rallies."

Student attendance was the final element considered to influence the dosage of the intervention. The evaluators had initially planned to request student attendance records from the participating schoolsat the end of the school year. Once the Striving Readers funding was cancelled, the evaluators chose not to place this additional burden on the schools and to focus resources on obtaining the student achievement data.Nevertheless, during the monthly check-ins, the evaluators asked the teachers to report on student attendance during that month. Attendance was more sporadic in school one (63% all students attended), while schools 3 and 5 reported full attendance. The main reason for sporadic attendance was student behavior. For instance, within the period of a month, one of the teachers had seven students suspended for a total of 20 days of missed classes, while another had suspensions for a total of 34 missing days.

The actual intensity of the intervention may be assessed by the number of Expeditions completed during the school year. Cambium establishes a pacing calendar for the teachers to ensure that all Expeditions are covered within the school year. However, as discussed above, a number of factors intervened with the planned calendar.

At the end of the year, the evaluators reviewed data on VPORT to assess how much of the program the teachers had covered. Of the 14 Expeditions that form the *PRJ III* curriculum, completion rates varied from 29 percent (4 Expeditions) to 78 percent (11 (Expeditions). However, the information must be viewed with caution, since data were entered into the system by the teachers. If the teacher was not up-to-date with the data input into the system, the information will be misleading.Table 4 displays the scores for fidelity of classroom implementation. Based solely on the adequacy of delivery observed during the site visits, two of the six interventionists would have been classified as attaining high fidelity of implementation (score at or above 0.90), three would attain medium fidelity (scores between 0.70 and 0.89, and one would be classified as low fidelity (score below 0.70).

Table 5: Scoring for the Classroom Observation Rubric

Components		Schools					
		1	2	3	4	5	6
Number of classes observed		4	6	4	4	4	5
Classroom time observed (minutes)		46	47	50	59	49	47
Section A: Classroom Environment							
A1 - Sufficient space		2.00	1.50	2.00	2.00	2.00	0.40
A2 - Instructional areas		2.00	1.50	2.00	2.00	2.00	0.80
A3 - Teacher resources		2.00	1.80	2.00	2.00	2.00	2.00
A4 - Student materials		2.00	2.00	2.00	2.00	2.00	2.00
Section score: $X_A/8$		1.00	0.85	1.00	1.00	1.00	0.65
Section B: Lesson Plan (General)							
B1 - Follows curriculum guide		2.00	0.50	1.00	2.00	1.50	0.60
B2 - Brisk pace		1.50	1.50	1.80	2.00	2.00	0.10
B3 - Skills modeled		2.00	1.30	2.00	2.00	1.90	0.80
B4 - Correction procedures		2.00	0.70	1.70	2.00	2.00	0.00
B5 - Students in groups		2.00	1.50	2.00	2.00	2.00	2.00
B6 - Differentiation		2.00	2.00	2.00	2.00	2.00	n/o
Section score: $X_B/(12 -$ number of n/o)		0.96	0.63	0.87	1.00	0.95	0.35
Section C: Lesson Plan (Specific)							
C1 – Components delivered in order		1.60	2.00	2.00	2.00	1.70	1.70
C2 – Steps delivered in order		0.90	1.10	0.00	2.00	1.30	0.70
C3 - Completed within suggested timeframe		0.00	2.00	n/o	2.00	n/o	0.00
Section score: $X_C/(6$-number of n/o)		0.41	0.85	0.50	1.00	0.75	0.41
Section D: Classroom Management							
D1 - % time students pay attention		1.00	1.00	1.00	1.00	1.00	0.80
D2 - % time students respond to prompts		0.80	0.70	1.00	0.90	0.80	0.70
D3 - % time students actively participate		1.00	1.00	1.00	1.00	1.00	0.70
D6 - % time students follow expectations		0.90	1.00	1.00	1.00	1.00	0.90
D4 - % time teacher addresses behavior (x2)		1.60	1.70	2.00	1.80	2.00	1.60
D5 - % time teacher engaging students (x2)		2.00	1.10	1.70	1.70	2.00	1.90
Section score: $X_D/($total time intervals $-$N/A)		0.90	0.81	0.96	0.91	0.97	0.83
Weighted scores							
Sections	**Weight**	**1**	**2**	**3**	**4**	**5**	**6**
A	.20	0.20	0.17	0.20	0.20	0.20	0.13
B	.30	0.29	0.19	0.26	0.30	0.28	0.11
C	.30	0.12	0.26	0.15	0.30	0.23	0.12
D	.20	0.18	0.16	0.19	0.18	0.19	0.17
Classroom Fidelity Index		0.79	0.78	0.80	0.98	0.90	0.53

Outcome Study: Measures

Two assessments were used to measure student performance: The Gates MacGinitie Reading Test (GMRT®) 4th edition for grade 9, and the Grade 9 EXPLORE.®GMRT,®from Riverside Publishing, is a nationally-normed reading assessment that has well-established psychometric qualities, as documented in the 2002 technical manual (MacGinitie, MacGinitie & Dryer, 2002).Alternate form reliability coefficients for test levels appropriate for grades 6-8 range from 0.82 to 0.91. K-R 20 estimates of reliability range from 0.90 to 0.95. The test was re-normed during SY 2005-2006 (Maria & Hughes, 2008). For the ISR project, GMRT®was to be applied twice a year, early in the fall and in late spring of each implementation year.The GMRT was selected due to its reputation as a sterling reading assessment with demonstrated psychometric value. The Grade 9 EXPLORE® is a component of the ACT testing system, published by ACT, Inc. This criterion-referenced test measures academic achievement in English, mathematics, reading and science using a multiple choice format. Details about the test development and psychometrics are found in the technical manual (ACT, 2011). In addition to the demonstrated psychometric value of the Grade 9 EXPLORE®, the assessment was being used by the participating high schools at the time of the ISR Project, which made it a convenient outcome. The EXPLORE Reading test is made available to Illinois schools as part of the reform agenda on moving toward a growth model of accountability and was taken by students enrolled in the ILSR schools. (see http://www.isbe.state.il.us/assessment/htmls/protocol.htm).

Outcome Study: Attrition

A randomized control trial is an effect way to minimize biased results when conducting a study. In educational studies where random assignment has been carried out, such as this one, it is important to assess attrition of the randomized sample at the time of outcome measurement. Thus, the portion of overall participating students that have left the study since the time of assignment, and the differentiation of portion leaving between treatment and control groups must both be assessed. This is because both overall and differential attrition can jeopardize the integrity of the randomized design and introduce bias in the impact estimates. Attrition breakdowns on both outcomes are presented below.

Table 6: Attrition on Outcomes

	Assigned	Attrited	Attrition Rate
Grade 9 GMRT®			
Overall	855	341	39.9%
Treatment Group	427	163	38.2%
Control Group	428	178	41.6%
Grade 9 EXPLORE®			
Overall	855	395	46.2%
Treatment Group	427	189	44.3%
Control Group	428	206	48.1%

Overall attrition levels on the Grade 9 EXPLORE® and Grade 9 GMRT®are 39.9% and 46.2%, respectively. Differential attrition levels are calculated as the difference of the attrition rates between the treatment and control groups. They are 3.4% and 3.8% for the Grade 9 EXPLORE® and Grade 9 GMRT®, respectively. Although the differential attrition rates were low, the data for students in matched pairs affected by attrition were dropped from the statistical data analyses to ensure that the balance of matched students across treatment and control groups was maintained; (the differential attrition rates in this study are considered low according to the WWC standards, as outlined in the Procedures and Standards Handbook, v. 2.1: http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19#).

This resulted in a total sample of 628 students evenly distributed across treatment and control groups in each school. The sample size, means, and standard deviation for the resulting samples of participants by treatment conditions and schools on the pretest and posttest measures of the two outcome variables used in this study (Grade 9 GMRT® and Grade 9 EXPLORE®) are provided in Table 7.

Statistical Data Analysis

The first two research questions, related to the overall average of the pretest-posttest effects of the *PRJ III* supplemental literacy intervention on reading achievement of grade 9 students at the end of the first implementation year and the variability of these effects across schools, were addressed by using a three-level HLM modelthat takes into account the multilevel structure of the data (students nested within classes and classes nested within schools). As describe earlier, theintent-to-treat (ITT) classes within schools are carefully matched grouped, *not* randomly selected intact classes, that were randomly assigned to control and treatment conditions.

Therefore, it was expected that the variation among ITT classes within schools on the dependent variable will be practically negligible, so it might be appropriate to use a more parsimonious two-level HLM model, with students nested within schools. To test this expectation, a fully unconditional three-level HLM was used first. With the notations used in Raudenbush and Bryk (2002, p.229), the analytic form of this model is

Level 1 (Student-Level) Model:

$$Y_{ijk} = \pi_{0jk} + e_{ijk}, \quad (1)$$

where

Y_{ijk} is the posttest score of student *i*from classroom *j*in school *j* on the Grade 9 GMRT®;

π_{0jk} is the mean of posttest scores of classroom *j* in school k; ande_{ijk} is a random "student effect," that is, the deviation of Y_{ijk} from the class mean; it is assumed that $e_{ijk} \sim N(0, \sigma^2)$.

Level 2 (Classroom-Level) Model:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \quad (2)$$

where

β_{00k} is the mean of students' posttest scores in school *k*; and $r_{0jk}$is a random "classroom effect," that is, the deviation of a classroom mean from therespective school mean; it is assumed that $r_{0jk} \sim N(0,\tau_\pi)$.

Level 3 (School-Level) Model:

$$\beta_{00k} = \gamma_{000} + u_{00k}, \quad (3)$$

where

γ_{000} is the grand mean; and

u_{00k} is a random "school effect," that is, the deviation of the mean of school *k* from the grand mean; it is assumed that $u_{00k} \sim N(0,\tau_\beta)$.

To address the first research question (RQ1), a three-level HLM is used that includes the students' pretest scores on the Grade 9 GMRT® as a covariate at the student level (Level 1) and the type of treatment condition (0 = control, 1 = experimental) at classroom level (Level 2). The analytic form of this model is

Level 1:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}\left(X_{ijk} - \bar{X}\mathbf{...}\right) + e_{ijk}, \quad (4)$$

where Y_{ijk} is the posttest score of student *i*from classroom *j*in school *j* on the Grade 9 GMRT®;

X_{ijk} is the pretest score of student *i*from classroom *j*in school *k* on the Grade 9 GMRT®

(centered around the grand mean, $\bar{X}\mathbf{...}$); π_{0jk} is the adjusted mean of posttest scores for class *j* in school *k*, after controlling for differences in pretest status; And π_{1jk} is the level-1 pretest effect; And e_{ijk} is a random "student effect," assuming $e_{ijk} \sim N(0, \sigma^2)$.

Level 2:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}\left(PRJ_{jk}\right) + r_{0jk} \quad (5)$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk}, \quad (6)$$

where

PRJ_{jk} is a treatment indicator variable, with PRJ_{jk} = 1 indicating that classroom *j* in school *k* is an experimental group, that is, participates in*PRJ III*, and PRJ_{jk} = 0

otherwise;

β_{00k} is the adjusted mean of posttest scores in school *k*;
β_{01k} is the treatment (*PRJ*) effect in school *k*;

r_{0jk} is a random error term, with the assumption that $r_{0jk} \sim N(0,\tau_{\pi 0})$.
β_{10k} is the average pretest effect in school *k*;
r_{1jk} is a random error term, with the assumption that $r_{1jk} \sim N(0,\tau_{\pi 1})$.

Level 3:

$$\beta_{00k} = \gamma_{000} + u_{00k} \quad (7)$$
$$\beta_{01k} = \gamma_{010} + u_{01k} \quad (8)$$
$$\beta_{10k} = \gamma_{100} + u_{10k}, \quad (9)$$

where

γ_{000} is the adjusted grand mean of posttest scores in the control groups across all schools;

u_{00k} is a random error term for β_{00k}, assuming that $u_{00k} \sim N(0,\tau_{\beta 00})$;

γ_{010} is the average treatment effect across all schools, controlling for pretest differences;

u_{01k} is a random error term for β_{01k}, assuming that $u_{01k} \sim N(0,\tau_{\beta 01})$;
γ_{100} is the average pretest slope across all schools,

u_{10k} is a random error term for β_{10k}, assuming that $u_{10k} \sim N(0,\tau_{\beta 10})$.

Under this model, the significance of the overall effect, γ_{010}, is of primary interest in addressing the first research question in this study, RQ1. The second research question, RQ2, is about the variability of the treatment effects of the *PRJ III* supplemental literacy intervention on reading achievement across schools. To address RQ2, the model presented next is a two-level HLM model, with students nested within schools.

This model takes into account the results from the preliminary analysis on the fully unconditional three-level HLM (Equations 1-3) which show that the variance among ITT classes within schools is negligible—specifically, the variance of the random "classroom effect ," r_{0jk}, in Equation 2 is not statistically significant (see Table 8).

The Gates-MacGinitie Reading Tests®(GMRT) 4thEdition and the reading component for theGrade 9 EXPLORE®were used to measure the reading achievement of treatment and control students. The analytic form of the two-level HLM design used to address RQ2 is provided next.

Level 1 (within-school) part:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\left(PRJ_{ij} - \overline{PRJ}_{.j}\right) + \beta_{2j}\left(PRE_{ij} - \overline{PRE}_{.j}\right) + \varepsilon_{ij}, \qquad (10)$$

where:

Y_{ij} = the posttest score of student *i* in school *j* on the Grade 9 GMRT®;

PRJ_{ij}= a treatment indicator variable, with PRJ_{ij} = 1 indicating that student *i* in school *j* participates in*PRJ III*, and PRJ_{ij} = 0 otherwise;

PRE_{ij} = pretest score of student *i* in school *j* [on the Grade 9 GMRT®];

$(PRJ_{ij} - \overline{PRJ}_{.j})$ and $(PRE_{ij} - \overline{PRE}_{.j})$ indicate that PRJ_{ij} and PRE_{ij}, respectively, are centered around their school means [*group centering*, Raudenbush & Bryk, 2002];

β_{0j} = the mean of school *j* on the Grade 9 GMRT®; (by virtue of the group centering);

β_{1j} = the treatment (PRJ) effect for school *j*, holding constant pretest performance;

β_{2j} = the pretest/posttest *slope* for school *j*, holding constant *PRJ* (1 or 0);

ε_{ij} = errors; $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Level 2 (between-schools) part:

$$\beta_{0j} = \gamma_{00} + u_{0j}; \quad u_{0j} \sim N(0, \tau_{00}), \qquad (11)$$

$$\beta_{1j} = \gamma_{10} + u_{1j}; \; u_{1j} \sim N(0, \tau_{11}), \qquad (12)$$

β_{2j} = γ_{20} + u_{2j}; u_{2j} ~ $N(0, \tau_{22})$, (13)

where:

γ_{00} = the grand mean [on posttest Grade 9 GMRT® scores];
γ_{10} = the overall average PRJ effect;
γ_{20} = the average pretest/posttest slope;
τ_{00} = between-school variance of mean scores;
τ_{11} = between-school variance of PRJ effects;
τ_{22} = between-school variance of pretest/posttest slopes.

Note that τ_{11}, which represents the variability of the *PRJ III* effects across schools, is of primary interest in addressing RQ2. For methodological details on a similar HLM design with matched control and experimental groups, the reader may refer to Seltzer (2004, pp. 261-266). The HLM analyses under the three models described here above were conducted using the computer program HLM, Version 6 (Raudenbush, Bryk, Cheong, & Congdon, 2004).

Alternatively, the analysis under the HLM models described here above was conducted with the dependent variable being the posttest scores of the students on the Grade 9 EXPLORE® (the findings from all analyses did not differ across these two dependent variables). Note that fidelity of implementation scores were not included as level 2 covariates. This is in accord with What Works Clearinghouse standards, which state that impact models should not correct for variation in program implementation across sites (What Works Clearinghouse, 2008). Rather, implementation findings are presented to contextualize the results of the study.

Results

The Fully Unconditional Three-Level HLM

The results from the three-level HLM described in Equations 1-3 are provided in Table 8.

The variance of the random classroom effect r_{0jk}in Equation 2, which is of primary interest in this analysis, is close to zero (τ_π= 6.05) and *not* statistically significant (*p* = .068).As a result, the proportion of variance among classes within schools is also close to zero: $\tau_\pi/(\sigma^2 + \tau_\pi + \tau_\beta)$ = 6.05/(516.41 + 6.05 + 0.60) = 0.012 (e.g., Raudenbush & Bryk, 2002, p. 230). This finding supports our expectation that, due to the careful matching of the participants in the control and experimental classes in the intent-to-treat (ITT) group of students in each school, the variability among classes within schools is practically negligible.

Table 8:A Fully Unconditional Three-Level Hierarchical Linear Analysis Under the Model Defined by Equations 1-3, with the Grade 9 GMRT® Scores as the Outcome Variable

Outcome/ School	Treatment Condition									
	Experimental (*PRJ III*)					Control				
		Pretest		Posttest			Pretest		Posttest	
	N	*M*	*SD*	*M*	*SD*	*N*	*M*	*SD*	*M*	*SD*
Grade 9 GMRT® Scores										
1	53	511.86	22.86	513.43	22.30	53	511.69	25.95	512.13	26.66
2	37	516.06	19.50	520.22	23.77	37	508.99	19.23	511.35	20.79
3	65	513.64	24.13	517.39	22.44	65	514.21	24.62	519.11	23.64
4	45	509.57	18.97	511.44	23.20	45	509.60	22.06	514.57	26.14
5	54	516.06	16.77	520.60	18.88	54	514.67	19.65	518.61	20.35
6	60	516.76	20.01	521.20	21.58	60	512.33	23.30	513.50	23.30
TOTAL	314	514.05	20.74	517.48	22.08	314	512.23	22.77	515.21	23.66
Grade 9 EXPLORE®Scores										
1	53	10.60	1.15	11.79	3.14	53	10.30	1.42	12.37	2.72
2	37	10.84	1.04	12.52	2.36	37	10.57	1.07	12.30	2.26
3	65	10.58	1.21	12.25	2.54	65	10.40	1.34	12.37	2.81
4	45	10.78	1.15	11.22	3.56	45	10.71	1.24	11.27	3.03
5	54	10.63	1.17	12.20	2.06	54	10.63	1.20	12.51	2.14
6	60	10.80	1.10	12.30	2.37	60	10.00	2.02	12.04	2.47
TOTAL	314	10.69	1.14	12.06	2.70	314	10.41	1.46	12.16	2.61

HLM Analysis Related to RQ1

As described in the Method section, the first research question (RQ1), which is about the overall treatment effect, is addressed by using the three-level HLM defined by Equations 4-9. The results from the initial analysis under this model indicated that the variances of the random effects in Equations 8 and 9 (at Level 3) are *not* statistically significant, namely (a) VAR(u_{01k}) = 0.34, χ^2(5) = 3.88, p = .57, and (b)VAR(u_{10k}) = 0.0004, χ^2(5) =5.37, p = .372.Therefore, Equations 8 and 9 were reduced to the equations $\beta_{01k} = \gamma_{010}$ and $\beta_{10k} = \gamma_{10k}$, respectively. The estimates of the parameters in the resulting three-level HLM are provided in Table 9.The results show that the overall*PRJ III*treatment effect (γ_{010} = 0.84) is *not* statistically significant (p = .45). This finding relates to the first research question (RQ1) in the study. In addition, one can see that (a) the average pretest/posttest slope, γ_{100}= 0.86, is statistically significant, p< .001, and (b) the variances of the random effects (r_{0jk}, r_{1jk}, and u_{00k}) are not statistically significant (p< .05).

HLM Analysis Related to RQ2

The second research question (RQ2), which is about the between-school variance of the *PRJ III* effects, is addressed through by using the two-level HLM model defined with Equations 10-13. As noted earlier, this model ignores the variation among the (carefully matched control and experimental) classes within schools because the results from the fully unconditional three-level HLM showed that this variation is close to zero and not statistically significant. The results are summarized in Table 10.As can be seen, the estimate of the between-school variance of PRJ effects, which is of primary interest in addressing RQ2, $\hat{\tau}_{11}$= VAR(u_{1j}) = 1.30, is*not* statistically significant, χ^2(5) = 4.76, p = .45. Thus, the GMRT®posttest performance of *PRJ III*students relative to the comparison group students, controlling for pretest differences, does not depend on school membership.

Table 9: A Three-Level Hierarchical Linear Analysis Under the Model Defined by Equations 4-9, with the Grade 9 GMRT® Scores as the Outcome Variable

Fixed Effect	Coefficient	*SE*	*t* Ratio	*p* Value
Grand mean, γ_{000}	515.87	0.80	653.65 (*df* = 5)	< .001
Average PRJ effect, γ_{010}	0.84	1.11	0.76 (*df* = 46)	.45
Average pretest effect, γ_{100}	0.86	0.02	35.00 (*df* = 47)	< .001

Random Effect	Variance Component	*df*	χ^2	
Class mean, r_{0jk}	1.822	41	52.91	.101
Pretest effect, r_{1jk}	<0.01	47	49.11	.388
School mean, u_{00k}	< 0.01	5	4.84	> .500
Level-1 error, e_{ijk}	168.35			

Note. The level-1 covariate are the pretest scores on *Grade 9 GMRT®*, centered around the grand mean, and the level-2 covariate is the classroom treatment (0 = control, 1 = experimental); (there is no level-3 covariate).

Table 10: A Two-Level Hierarchical Linear Analysis Under the Model Defined by Equations 10-13, with the Grade 9 GMRT® Scores as the Outcome Variable

Fixed Effect	Coefficient	*SE*	*t* Ratio	*p* Value
Grand mean, γ_{00}	516.17	1.15	447.91 (*df* = 5)	< .001
Average PRJ effect, γ_{10}	0.70	1.14	0.61 (*df* = 5)	0.568
Average pretest effect, γ_{20}	0.86	0.03	30.92 (*df* = 5)	< .001

Random Effect	Variance Component	*df*	χ^2	
Adjusted school mean, u_{0j}	6.30	5	24.24	< .001
School PRJ effect, u_{1j}	1.30	5	4.76	> .500
School pretest effect, u_{2j}	<0.01	5	5.63	0.343
Level-1 error, e_{ijk}	170.63			

The findings related to the two research questions (RQ1 and RQ2) were the same when using the students' performance on the Grade 9 EXPLORE®test (instead of Grade 9 GMRT®).The results are summarized in Tables 11, 12, and 13. As can be seen in Table11with the results for the fully unconditional three-level HLM, the variance of the random classroom effect $r_{0jk}$is not statistically (τ_π= 6.05, *p* = .068). As a result, the proportion of variance among classes within schools is close to zero: $\tau_\pi/(\sigma^2 + \tau_\pi + \tau_\beta)$ = 6.05/(516.41 + 6.05 + 0.60) = 0.012 (Raudenbush & Bryk, 2002, p. 230). The results from the three-level HLM used to address RQ1 (Equations 4-9), with the students' performance on the Grade 9 EXPLORE® test being the dependent variable, are summarized in Table 12. Once again, Equations 8 and 9 were reduced to $\beta_{01k} = \gamma_{010}$ and $\beta_{10k} = \gamma_{10k}$, respectively, because the results from an initial analysis under this model indicated that the variances of the random effects in Equations 8 and 9 are *not* statistically significant, namely (a) VAR(u_{01k}) = 0.34, χ^2(5) = 3.88, *p* = .57, and (b) VAR(u_{10k}) = 0.0004, χ^2(5) =5.37, *p* = .372. As shown in Table 12, the overall *PRJ III* treatment effect (γ_{010} = 0.84) is *not* statistically significant (*p* = .45). Regarding the second research question, RQ2, the results from the two-level HLM (Equations 10-13) are summarized in Table 13. As can be seen, the estimate of the between-school variance of PRJ effects, which is of primary interest in addressing RQ2, $\hat{\tau}_{11}$= VAR(u_{1j}) = 0.03, is *not* statistically significant, χ^2(5) = 2.14, *p*> .500.

Table 11: A Fully Unconditional Three-Level Hierarchical Linear Analysis Under the Model Defined by Equations 1-3, with the Grade 9 EXPLORE® Scores as the Outcome Variable

Fixed Effect	Coefficient	*SE*	*t* Ratio	*p* Value
Grand mean, γ_{000}	12.10	1.14	83.60 (*df* = 5)	< .001
Random Effect	Variance Component	*df*	χ^2	
Classroom effect, r_{0jk}	0.10	42	57.01	0.061
School effect, u_{00k}	0.04	5	9.28	0.097
Level-1 error, e_{ijk}	6.88			

Table 12: A Three-Level Hierarchical Linear Analysis Under the Model Defined by Equations 4-9, with the Grade 9 EXPLORE® Scores as the Outcome Variable

Fixed Effect	Coefficient	*SE*	*t* Ratio	*p* Value
Grand mean, γ_{000}	12.200	0.20	61.25 (*df* = 5)	< .001
Average PRJ effect, γ_{010}	0.21	0.25	0.85 (*df* = 46)	.40
Average pretest effect, γ_{100}	0.42	0.09	4.65 (*df* = 47)	< .001
Random Effect	Variance Component	*df*	χ^2	
Class mean, r_{0jk}	0.25	41	73.80	.002
Pretest effect, r_{1jk}	0.06	47	62.54	.064
School mean, u_{00k}	0.04	5	9.16	.102
Level-1 error, e_{ijk}	2.54			

Note. The level-1 covariate are the pretest scores on *Grade 9 EXPLORE*®centered around the grand mean, and the level-2 covariate is the classroom treatment (0 = control, 1 = experimental); (there is no level-3 covariate).

Table 13: A Two-Level Hierarchical Linear Analysis Under the Model Defined by Equations 10-13, with the Grade 9 EXPLORE® Scores as the Outcome Variable

Fixed Effect	Coefficient	*SE*	*t* Ratio	*p* Value
Grand mean, γ_{00}	12.10	1.17	71.30 (*df* = 5)	< .001
Average PRJ effect, γ_{10}	0.21	0.22	0.94 (*df* = 5)	0.391
Average pretest effect, γ_{20}	0.37	0.10	3.59 (*df* = 5)	< .021
Random Effect	Variance Component	*df*	χ^2	
Adjusted school mean, u_{0j}	0.11	5	12.87	.024
School PRJ effect, u_{1j}	0.03	5	2.14	> .500
School pretest effect, u_{2j}	0.02	5	7.52	0.184
Level-1 error, e_{ijk}	6.75			

Validity Control

Both sampling and statistical procedures of control were used to enhance the internal validity of the findings related to the two research questions in this study, RQ1 and RQ2. First, as described in the Method section, the control and treatment groups within schools were formed through (a) careful matching of participants on relevant variables, such as pretest scores, free and reduced meals (FARM) eligibility, Limited English Proficiency (LEP), special education status (SPED), gender, and ethnicity, (b) random assignment of the students from each matched pair to two different groups, and (c) random assignment of the resulting two groups to control and treatment conditions. Second, although the differential attrition rates were low, the data for students in matched pairs affected by attrition were dropped from the statistical data analyses to ensure that the balance of matched students across treatment and control groups was maintained. Third, the pretest scores of the participants on the outcome variable were includes as a covariate in the HLM analysis. As shown in Table 7, the pretest scores of the control and treatment groups across schools are practically almost identical thus suggesting a lack of pretest differences that may affect the reported results on RQ1 and RQ2.

Along with this descriptive information, testing for interaction between pretest and treatment condition (control vs.*PRJ III*) was also conducted. This was done by adding an interaction term (pretest x treatment) in the HLM Equation 10, but having *PRJ* (treatment), *PRE* (pretest), and their interaction (*PRJ* x *PRE*) uncentered predictors and treating the respective regression coefficients as fixed. Reported here are only results related to the interaction term *PRJ* x *PRE*, which was of primary interest in this auxiliary analysis. The regression coefficient for this interaction term was not statistically significant thus indicating a lack of interaction between the pretest scores and treatment condition for both the Grade 9 GMRT® and Grade 9 EXPLORE® scores as outcome variables:t (624) = 1.51, p = .131, and t(624) = 1.61, p =.107, respectively. Given the rigorous matching, random assignment, and attrition control for participants in the control and treatment groups, the lack of such interaction is not a surprise and provides an additional support to the validity of the findings in this study.

Discussion

Main Findings

This study investigated effects of the *PRJ III* supplemental literacy intervention on reading achievement of grade 9 students at the end of the first implementation year (SY 2010/11) under the *Illinois Striving Readers* (*ISR*) project. The results related to the first research question (RQ1) suggest that there is *no* statistically significant pretest-posttest effects of the *PRJ III* intervention on student reading performance as measured separately by the Grade 9 GMRT® and Grade 9 EXPLORE®tests, respectively.Regarding the second research question (RQ2), the results indicate that there is no between-school variation in intervention effects across schools thus indicating that the GMRT® or 9 EXPLORE®test performance of *PRJ III*students relative to the comparison group students does not depend on school membership. This finding is most likely due to (a) similarity of schools (Title-I eligible schools not making AYP) and (b) similarity in reading performance of students, with all participants scoring at the lowest quartile on ACT's EXPLORE®.

Although the findings in this study do not suggest statistical significance of the effects targeted with the two research questions, they are nevertheless results of an experimental study and thus can serve the purpose of contributing to a body of evidence that meets rigorous research standards in exploring the effectiveness of reading interventions. We argue that it is important to report these findings so that the body of evidence is not misrepresented by only those cases where statistical significance is attained.

The study also highlights a discrepancy between findings from the classroom model and the outcome study. Data generated by the observation rubric were quantified, weighted, and then used to generate a fidelity rating of inadequate, adequate, or high. Three of the six schools received an adequate fidelity rating, two received a high fidelity rating, and one received a low fidelity rating. Despite the variation of fidelity ratings, the intra-class correlation values (ICC) show that outcome variation at the school level was low. ICC values for the Gates MacGinitie Reading Test® and Grade 9 EXPLORE® are 0.045 and 0.012, respectively. These ICC values are corroborated by the results shown in Table 8, which indicate that the impact of *PRJ III* does not vary across schools.

Limitations of the Study

The primary limitation of this study is that the participating sample was drawn from volunteer schools. Therefore, because participants were not selected via a random sample from a population of interest, the results from this study cannot be generalized to outside populations (Cohen, 1994). The absence of *PRJ III* impacts can be attributed to our participating schools alone. Similar studies should be conducted in other settings to generate more evidence regarding the effectiveness of *PRJ III*.

When the outcome assessments were conducted, teachers were far from completing the program of instruction. Indeed, by the end of the school year, most had completed half or less of the program. This observation stands true to most studies that rely on statewide assessments, since these assessments tend to happen before the school year is completed and while teachers are still providing instruction on the topics that will be assessed. The influence of program completion on test results is a topic that merits further examination.

Recommendation for Future Research

The lack of external generalizability due to the volunteer sample dictates a recommendation for further studies that examine the effectiveness of *PRJ III* using similar experimental designs. Our design, consisting of a random assignment of students to experimental and control conditions, ensures the internal validity of our findings. That is, we are confident that the findings of non-significant impacts of *PRJ III* on the reading ability of 9th graders in Illinois Striving Readers high schools accurately reflect an ineffectiveness of the program with this population. This finding, however, is in contradiction with findings generated by Denton (2008) and Shneyderman (2006), as explained earlier. Findings from additional well designed studies can be used to support of refute the results of our study.

In conclusion, the findings indicate that the *PRJ III* supplemental literacy intervention had no impact on the reading achievement of grade 9 students at the end of the first implementation year (SY 2010/11). These findings contribute to a growing body of evidence for the effectiveness of these literacy interventions targeting adolescent students.

The findings cannot, however, be generalized across a broad population due to a lack of a randomly drawn sample. Therefore, the value of our study findings would be supplanted by additional studies testing the effectiveness of PRJ III, and other literacy interventions targeting adolescent students.

References

ACT, Inc. (2011). 2011/2012 Technical Manual, EXPLORE. Retrieved from http://www.act.org/explore/pdf/TechManual.pdf

Allison, P. (2001). Missing data. Thousand Oaks, CA: Sage.

Asparouhov, T. and Muthén, B. (2010a) Bayesian Analysis Using Mplus. Mplus Technical Report. Retrieved from http://nnwww.statmodel.com

Asparouhov, T. and Muthén, B. (2010b) Bayesian Analysis of Latent Variable Models using Mplus. Mplus Technical Report. Retrieved from http://nnwww.statmodel.com

Baker, S., Simmons, D.C., & Kame'enui, E.J. (2004). Vocabulary acquisition: Synthesis of the research. Retrieved from http://idea.uoregon.edu/~ncite/documents/techrep/tech13.html

Beck, I.L., McKeown, M.G., & Kucan, L. (2002). Bringing words to life: Robust vocabulary instruction. New York: The Guilford Press.

Biancarosa, G., and Snow, C.E. (2006). Reading next: a vision for action and research in middle and high school literacy. Washington, DC: Alliance for Excellent Education.

Carnevale, A. P. (2001). Help Wanted . . . College Required. Washington, DC: Educational Testing Service, Office for Public Leadership.

Denton, K. (2008). Comparison of the reading achievement of Journeys students with students receiving traditional instruction. Retrieved from http://www.voyagerlearning.com/ResearchStudyDocuments/dallas_comparison_report.pdf

Deshler, D.D., Palincsar, A.S., Biancarosa, G., & Nair, M. (2007). Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practices. Newark, DE: International Reading Association.

Gersten, R. Fuchs, L.S., Williams, J.P., Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of the research. Review of Educational Research, 71, 279-320.

Graham, S., & Perin, D. (2007). Writing next: Effective strategies to improve writing of adolescents in middle and high schools – A report to Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., and Torgesen, J. (2008). Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

MacGinitie, W.H., MacGinitie, R.K., Maria, K., & Dreyer, L.G. (2002). Gates-MacGinitie reading tests: Technical report. Rolling Meadows, IL: The Riverside Publishing Company.

Maria, K. & Hughes, K.E. (2008). Gates-MacGinitie reading tests: Technical report supplement. Rolling Meadows, IL: The Riverside Publishing Company.

Marzano, R.J. (2004). Building background knowledge for academic achievement. Alexandria, VA: ASCD.

Mastropieri, M.A., Scruggs, T.E., & Graetz, J.E. (2003). Reading comprehension instruction for secondary students: Challenges for struggling students and teachers. Learning Disability Quarterly, 26(2), 103-116.

Maxwell, J.A., & Miller, B.A. (2008). Categorizing and connecting strategies in qualitative data analysis. In S.N. Hesse-Biber and P. Leavy (Eds.). Handbook of emergent methods, pp. 461-478. New York: Guilford Press.

Muthén, L.K. and Muthén, B.O. (2010). Mplus Users Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

National Center for Education Statistics (2011). The Nation's Report Card: Reading 2011 (NCES 2012-457). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

President's Council of Advisors on Science and Technology (2010). Prepare and inspire: K-12 education in science, technology, engineering, and math (STEM) for America's future. Washington, DC: Executive Office of the President.

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

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C.K., and Torgesen, J.K. (2007). Interventions for adolescent struggling readers: a meta-analysis with implications for practice. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schatschneider, C., Buck, J., Torgesen, J.K., Wagner, R.K., Hassler, L., Hecht, S., & Powel-Smith, K. (2004). A multivariate study of factors that contribute to individual differences in performance on the Florida Comprehensive Reading Assessment Test. Technical Report # 5. Florida Center for Reading Research, Tallahassee, FL.

Shneyderman, A. (2006). Some results of the Voyager Journeys Reading Intervention system in several district schools. Unpublished Manuscript, Miami-Dade County Public Schools, Office of Program Evaluation.

Seltzer, M. (2004). The use of hierarchical models in analyzing data from experiments and quasi-experiments conducted in field settings. In D. Kaplan (Ed.), The SAGE handbook of quantitative methodology for the social sciences (pp. 259-280). Thousand Oaks, CA: Sage.

Slavin, R. E., & Cheung, A. (2003). Effective reading programs for English language learners. A best-evidence synthesis (Report No. 66). Baltimore, MD: Center for Research on the Education of Students Placed At Risk (CRESPAR).

Torgesen, J.K. (2005). Remedial Interventions for Students with Dyslexia: National Goals and Current Accomplishments. In Richardson, S., & Gilger, J. (Eds.) Research-based education and intervention: What we need to know. (pp. 103-124). Boston: International Dyslexia Association.

U.S. Department of Education, National Center for Education Statistics (2003). Remedial Education at Degree- Granting Postsecondary Institutions in Fall 2000, NCES 2004-010, by Basmat Parsad and Laurie Lewis. Project Officer: Bernard Greene. Washington, DC: 2003.

What Works Clearinghouse. (2008). WWC procedures and Version 2 standards handbook. Washington, DC: Author. Retrieved September 2, 2012, from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf.