Title:

Four Years of Striving Readers Targeted Interventions: Impact Analyses and Secondary Explorations of Findings

Author(s):

William Loadman, Ohio State University; Kimberley Sprague, Brown University; Allison Meisch, Westat Inc.; Jill Feldman, Research for Better Schools; Bonnie Faddis, RMC Research Corporation

Paper 1: Springfield/Chicopee, MA
Kim Sprague, Anita Kite, Karen Hussar, Chandra Haislet, Colleen Zaller, and Joan Ford
The Education Alliance, Brown University
kim_sprague@brown.edu

Paper 2: Newark, NJ
Jennifer Hamilton and Allison D’Amour Meisch
Westat, Inc.
AllisonMeisch@westat.com

Paper 3: Mid-South
Jill Feldman, Ning Rui
Research for Better Schools

Paper 4: Ohio
William Loadman, Richard Lomax, Raenal Moore
Ohio State University
loadman.1@osu.edu
Jing Zhu
Metis Associates
zhujingosu@gmail.com

Paper 5: Portland, OR
Bonnie Faddis and Margaret Beam
RMC Research Corporation
bfaddis@rmccorp.com
Abstract Body

Background / Context:

Data from the National Assessment of Educational Progress (NAEP) show that in 8th grade, at a time when all students are expected to be able to acquire information through the reading of textbooks, 34 percent of boys and 24 percent of girls cannot read at the basic level (Grigg, Donahue, & Dion, 2007; Lee, Grigg, & Donahue, 2007). Among African American and Hispanic students, the picture is even more discouraging: 49 percent of African American and 45 percent of Hispanic 8th graders read below the basic level. Thus, by the time they reach high school, certain students have already been “left behind.” NAEP data also indicate that 27 percent of 12th-grade public school students scored below the basic level and only 35 percent scored above the proficient level. Since 1992, NAEP scores have declined significantly in all reading contexts (i.e., reading for literary experience, reading for information, reading to perform a task). These findings reflect an urgent need to identify effective scientifically based programs for improving reading in high schools in the US.

Recently there has been a noticeable increase in attention being paid to the literacy needs of students in fourth through twelfth grades—those students who are expected to be not “learning to read” but “reading to learn.” Students who struggle with reading at this level are not able to comprehend fully the texts from which they are expected to gain information. According to the National Center for Education Statistics (NCES), only 60 percent of fourth graders read at the “basic” level or above (2007). Further, according to the Alliance for Excellent Education, “69 percent of eighth graders and 65 percent of twelfth graders cannot read at or above grade level, thus making it difficult for them to comprehend complex texts and course materials” (Ayers & Miller, 2009).

In 2005, the U.S. Department of Education awarded eight Striving Readers grants to projects around the country. The goal of Striving Readers is to improve literacy skills and achievement for struggling readers in middle and high school and to increase the research base using randomized clinical research to address improvement in adolescent literacy instruction. This session will present results to date from the targeted interventions at five Striving Readers sites.

The five Striving Readers sites involved in this symposium are very different. Several of the sites are very diverse ethnically, but several others have very large majorities of a single ethnic group. Two sites are in major metropolitan areas, one is in a rust-belt city and nearby town, and another involves the high school system for incarcerated youth throughout the state. Further, the targeted interventions (READ 180 and Xtreme Reading) were adapted according to the needs and requirements of the different sites. These differences in the local context and in the adaptive use of the programs present an important opportunity to examine the different outcomes of experimental analyses. Combining experimental data with quasi-experimental statistical methods enabled the evaluators to leverage more information from their study designs than otherwise would be possible, and the presentations here include rigorous, thoughtful secondary analyses that examine how implementation might be associated with outcomes or how the intervention seems to be associated with outcomes for different subgroups.
Purpose / Objective / Research Question / Focus of Study:

Rigorous research provides information that will allow other schools and districts to select interventions that have a scientifically based track record of effectiveness. All Striving Reader grants include the mandate to evaluate literacy intervention(s) targeted to adolescents who are reading significantly below grade level. Although all studies followed the same guidelines for evaluation, each was implemented in a different context, evaluated different interventions or curricula, and utilized different evaluation designs. This symposium will bring together five evaluations to share information about the challenges in both designing, and reporting results of such studies after four years of implementation. Evaluators will provide important information to those planning similar studies in school districts and other complex social settings.

Settings:

Springfield/Chicopee: In the 2006–07 school year (the start of the grant), all five schools were eligible for Title I. In Springfield, 29 percent of the students were African American, 52 percent were Hispanic, and 14 percent were white; 71 percent were identified as low-income, and 13 percent were identified as Limited English Proficiency (LEP) students. In Chicopee, 78 percent were white, 3 percent were African American, and 18 percent were Hispanic; 35 percent were identified as low-income, and less than 2 percent were identified as LEP students.

Newark: Nineteen middle schools participated and were selected because they 1) were eligible for Title I funding, 2) served a minimum of two grades across sixth-, seventh- and eighth-grades, 3) were not already using READ 180, 4) were categorized as “in need of improvement” under No Child Left Behind, and 5) served a minimum of 25 eligible students reading at least two grades below grade level, based on the 2006 New Jersey state assessment. In these schools, 58 percent of students are African American, 41 percent are Hispanic, 88 percent are identified as low-income, and 7 percent are identified as being English Language Learners (ELL).

Mid-South: Eight middle schools serving more than 6,000 students in grades six through eight in a large city in the Mid-South are participating in the Striving Readers project. More than 90 percent of the students served by these schools are African American, while 5 percent are Hispanic; 88 percent of these students are eligible for free or reduced priced lunch, and 3 percent are identified as English Language Learners.

Ohio: The Ohio Department of Youth Services (ODYS) system is made up of seven youth detention facilities, six designated male facilities and one female facility. Students are assigned to a facility based on gender, type of offense and availability of space. High schools are located within each facility; a small number of students are middle school age. From October 2006 to August 2010, 5,512 students were enrolled. Seven percent of those students were female, 48 percent were African American, and 46 percent were white. The average length of stay for a student is about 10.5 months, although this varies substantially across students.

Portland: The Striving Readers program was implemented in 10 low-achieving Title I-eligible middle and high schools. In these schools, 65 percent of the students are identified as low income. On average, 28 percent of the students served by these schools are reading at least two years below grade level. None of the schools had achieved Adequate Yearly Progress at the
time of the Striving Readers application in 2005. Fifteen percent of the students in the 10 schools are identified as English Language Learners, and 59 percent are non-white.

**Population / Participants / Subjects:**

**Springfield/Chicopee:** Ninth-grade students are eligible and randomly assigned to one of the two targeted interventions or to the control condition. Students are screened using the Scholastic Reading Inventory. Of the 937 students eligible over the three current cohorts, only 636 were verified as eligible at final placement verification. Approximately 12 percent of the ITT group of 636 (78 students) had initially been reported inactive by the SR district team but were actually in attendance at least 75 percent of the time, based on both rosters and district attendance records (this percentage was the same in Year 2). Of these students, 27 no longer remained in the student-level district data provided ($n = 609$).

**Newark:** The New Jersey State Language Arts assessment determines eligibility; special education students who meet district’s eligibility requirements are included. Nineteen schools were randomly assigned; ten to implement READ 180 and nine to the control condition. In the third year of the study, 1,167 students participated in the evaluation.

**Mid-South:** Eligible students score in the bottom quartile of the state assessment, the Tennessee Comprehensive Assessment Program (TCAP). In the first year of the study, 1,740 struggling readers in grades six to eight were randomized to the treatment or control group; in each of the three subsequent years, approximately 550 to 600 sixth grade students were randomly assigned to treatment or control. This presentation will focus on those students who were still in the study during the fourth year (2009–2010).

**Ohio:** All students incarcerated in the seven ODYS facilities who score below reading level, as measured by the Scholastic Reading Inventory (SRI), are eligible. In the first three years of the study, 1,527 eligible students were randomly placed, but only 979 were in the intent to treat analysis. The remaining students were most likely released early and therefore there was no intent to treat these youth for at least two quarters.

**Portland:** The Striving Readers targeted intervention was provided to students in grades 7–10 who were reading two or more years below grade level. This presentation focuses on three cohorts of students who participated in one of the first 3 years of the program. The combined sample for Years 1 through 3 included 599 treatment group and 674 control group students who had posttest scores.

**Intervention / Program / Practice:**

The programs presented include versions of READ 180, developed by Scholastic, and Xtreme Reading, developed by the University of Kansas Center for Research on Learning. In Springfield-Chicopee and Mid-South, these programs were supplemental to the regular language arts curriculum; in the Newark, Ohio, and Portland they were incorporated into the regular arts curriculum as well as offered as supplements.

The READ 180 program is used to address the individual needs of struggling adolescent readers who are reading below grade level through adaptive and instructional software, teacher-directed instructional rotations, and the use of tailored textbooks and independent or modeled
reading of high interest literature. The program focuses on elements of phonics, fluency, vocabulary, comprehension, spelling, writing and grammar, and promoting self-directed learning. The **Xtreme Reading program** represents one of the levels of instruction in the Content Literacy Continuum (CLC), the framework within the Strategic Instruction Model (SIM). The Xtreme Reading program (level three of the framework) focuses on strategy instruction, particularly reading instruction that helps struggling middle and high school students to develop accurate word recognition and increased fluency and comprehension. Xtreme Reading’s core instructional approaches include direct instruction, teacher modeling, paired student practice, and independent practice.

**Springfield/Chicopee:** Teachers receive training and coaching visits, generally one per month for their first and second year teaching the interventions at a minimum. READ 180 teachers were to receive a two-day initial training, a one-day follow-up training and an additional 8 seminars. Xtreme teachers were to receive a three-day initial training and four days of follow-up training during the year are provided.

**Newark:** The Newark Striving Readers project modified **READ 180** to include some supplemental instruction aligned with the state assessment. For eligible students, the **READ 180** curriculum replaces the district’s regular core language arts curriculum. Teachers receive 20 hours of group training, five hours of training on using student assessment data to differentiate instruction, plus in-school coaching visits on an as-needed basis.

**Mid-South:** Students can receive the targeted intervention, in addition to their regular language arts classes, for two to years. Teachers were to participate in several all-day training sessions, and up to seven two-hour training sessions during the year. Teachers also had access to online professional development material and ongoing classroom support.

**Ohio:** New teachers were hired to teach the **READ 180** program; they were expected to attend 20 hours of professional development. In ODYS, **READ 180** replaces the regular 45-minute literacy class for eligible students randomized to the treatment group. In addition, these treatment group students also receive 45 minutes of additional instruction using the model, for a total of a 90-minute instructional block each day.

**Portland:** Professional development is offered to both participating teachers and literacy coaches for seven days during the summer and monthly half-days during the academic year. In the first two years of the project, two teachers co-taught the Xtreme Reading class along with a content class such as language arts or social studies. Each school had 1 Xtreme Reading class per grade level, taught by the same pair of teachers (1 reading specialist and 1 content specialist).

**Research Design:**

Each study is a randomized controlled field trial; Newark is a cluster-randomized field trial. The primary research question is whether or not the interventions or curricula implemented in these settings is effective in improving adolescent reading achievement for the targeted group as compared to the control group.

**Springfield/Chicopee:** Eligible ninth grade students receive the standard ELA curriculum. Students are screened using the SRI as per district plans at the end of eighth grade, and criteria for inclusion are reviewed at each phase (e.g., students are ineligible for several reasons as
identified by the districts and developers to avoid misplacement). Eligible students are randomly assigned to participate in one of the two supplemental programs (READ 180 or Xtreme Reading) or to “business-as-usual,” which may consist of tutoring or an elective in addition to the “business-as-usual” standard ELA curriculum. Eligible teachers are also randomly assigned to teach READ 180, Xtreme Reading, or the control group.

Newark: A cluster randomized design in which schools were randomly assigned to either implement READ 180 or to continue to provide only the regular language arts curriculum. Schools were blocked on four school-level variables prior to randomization: 1) school size, 2) proportion of ELL students, 3) proportion of students with special needs, and 4) AYP status.

Mid-South: Students scoring in the lowest quartile on the English/Language Arts component of the TCAP are randomly assigned to receive the supplemental READ 180 class or to a control group. Students continue to receive instruction in the regular language arts curriculum. Students in the treatment group receive READ 180 in addition to the regular language arts instruction.

Ohio: Incoming students scoring below reading level, as measured by the Scholastic Reading Inventory (SRI), are randomly assigned to either the READ 180 class or to a control group. Students maintain their assignment until released from the facility.

Portland: Prior to the start of each school year, eligible students in Grades 7–10 are randomly assigned to either the Xtreme Reading group or the control group. For Grades 7–8, the control group participates in the regular language arts curriculum; for Grades 9–10, the control group students participate in another elective course of their choosing.

Data Collection and Analysis:

Data are collected in all sites by the program which follows specified testing protocols. All sites have specified the use of multilevel modeling to account for the clustering of data with the exception of Springfield/Chicopee and Portland given those designs (they model the impacts of interventions on student outcomes using ANCOVA). All models will test intervention effects on achievement scores (students random assignment to group), controlling for pre-intervention reading levels and other covariates including school and cohort. A variety of covariates will be tested for inclusion in the analytical models. Where appropriate, analyses will also be run to disaggregate effects by grade level, gender, ELL, and special education. The outcome measures are as follows:

Springfield/Chicopee: Stanford Diagnostic Reading Test 4
Newark: Stanford Achievement Test-10, New Jersey State Language Arts assessment
Mid-South: Iowa Test of Basic Skills, Tennessee Comprehensive Assessment Program
Ohio: California Assessment Test, Scholastic Reading Inventory
Portland: Oregon State Assessment Test, Group Reading Assessment Diagnostic Evaluation (GRADE) normal curve equivalent (NCE) scores
Findings / Results:

Springfield/Chicopee In the initial years, the results were less conclusive given the sample size and the design includes power derived from the combination of all ninth grade cohorts over the entire period of the grant (five cohorts in sum). In Springfield/Chicopee, both interventions currently include four of the five cohorts to assess effectiveness. The results presented will include those based on an ITT analysis fitting a linear two-level model with student and teacher/school as the two levels. At the first level of the model, achievement for students within schools is predicted by a series of student characteristics.

Based on Year 3 results currently pending release, there were observed and significant effects of one of the interventions as compared to the control group. READ 180 students scored significantly higher as compared to Control students (2.70 points on average unadjusted and 5.28 adjusted). Unadjusted means represent the true difference between groups in a random assignment study, but the adjusted means were calculated in the event random assignment did not yield equivalent groups due to the smaller sample sizes. In Year 2, combined analysis had not yielded significant intervention effects, and intervention results were more similar as compared to the Control group. As current research indicates, when achievement gains are assessed across grade level, effect sizes decrease in the upper grades (Bloom, Hill, Rebeck Black, & Lipsey, 2006). Therefore, striving readers in the high schools would generally be expected to gain less than those in the lower grades simply as a result of the trajectory of student growth or development of reading skills.

Despite challenges in the implementation specifications and monitoring, patterns emerged in the treatment-only or Treatment-of-the-Treated (TOT) group and implementation levels. That is, moderate or higher levels of implementation were associated with higher outcome scores. Results are to be presented for the Intent-to-Treat (ITT) group and for the TOT group in the context of the implementation study results. In addition, multilevel models will be presented for both the ITT and TOT analyses, with implementation scores included for the treatment-only group analyses as a predictor. A final set of analyses to be presented include the trajectories for students participating in more than one year, though these analyses are non-experimental and descriptive as the RCT participation occurs in only one year (if participating students are eligible after that year, they receive another year of treatment including those in the control group). The discussion will be focused on the implications of implementation on the results of all analyses presented.

Newark: In Newark, the impact of READ 180 on eligible students was estimated for one, two, and three years of program implementation. The results are based on an ITT analysis fitting a linear two-level model with student and school as the two levels. At the first level of the model, achievement for students within schools is predicted by a series of student characteristics. At the second level of the model, performance of treatment and control schools is compared. Additional analyses are conducted for subgroups of students: African-American students, Hispanic students, and special education students. Based on analyses from the first 3 years of NPS Striving Readers data, READ 180 had an overall significant effect on some literacy skills. Overall, students in treatment schools exhibited the same level of achievement as students in control schools after they had 1 year of exposure to READ 180. However, after 2 and 3 years of exposure to READ 180, significant findings were observed for students in the treatment group. Our evaluation findings suggest that struggling adolescent readers may need at least two years of...
exposure to the intervention before significant impacts can be observed. A discussion on intervention fidelity and implementation will provide context to the evaluation findings.

**Mid-South:** The ITT results from years one through three showed no differences between treatment and control groups either as a whole or when analyzed by grade level (analyses from year four will be conducted after receipt of spring 2010 test data). Discussion with the district revealed a desire to understand whether *READ 180* was related to differential achievement in students in different subgroups. The secondary analyses presented in this symposia will use propensity score matching to identify the type of students in the treatment group who showed improvements on reading outcomes measures.

**Ohio:** The results from a series of HLM models found that the *READ 180* program had a significant overall impact on the low-achieving youth’s Lexile outcome. Youth in the *READ 180* group on average performed 57.34 points higher that their comparison counterparts after being offered two quarters of treatment. The effect size measured by Glass’s delta was fairly substantial given the huge variability of Lexile scores.

**Portland:** The analysis revealed a significant intervention effect for the overall sample: the treatment group students had significantly higher posttest reading achievement than the control group students. A significant treatment effect was also present for the middle school sample (Grades 7–8) but not the high school sample (Grades 9–10). The standardized effect size (SES) for the overall sample was .18; at the middle school level the SES was .27, and at the high school level the SES was .09. The same multilevel model was used to estimate the impact of the intervention on spring Oregon State Assessment Test (OSAT) reading scores. A significant treatment effect was present only for the middle school sample.

Because the intervention was implemented differently across the 10 Striving Readers schools, a nonexperimental analysis was conducted to examine between-school variability in treatment effects. Fidelity of implementation was measure by data collected through classroom visits, teacher surveys, and teacher logs of classroom activities. Additional data such as teacher characteristics, group professional development hours, school literacy coach assistance, and professional developer assistance for each teacher were also used. These analyses revealed that for Years 1 and 3 (but not Year 2), between-school variability in treatment effect was significant. Additional analyses that explored the possible effect of implementation of the impact of the intervention revealed that teacher fidelity of implementation, but not number of years of teacher experience or amount of professional development attended, significantly contributed to the between-school variability in treatment effects. This finding was true for the Year 1 sample, but in the Year 3 sample none of the teacher variables accounted for the between-school variability in treatment effect.

**Conclusions:**

Many schools are struggling to meet the needs of students who lack the reading skills required to succeed in high school and beyond. Because schools have finite resources, it is crucial that administrators allocate financial and human capital to literacy interventions promising the greatest potential for success. These studies are considered critical by the Department of Education to contribute to the field of adolescent literacy by addressing the primary research question of what interventions are effective. In addition, these studies are also some of the first to address what methods should be employed in such RCT trials including what
is required in terms of policy for implementation and collaboration to conduct such efforts. Given the steady push for more rigorous research and a corresponding dearth of information on this particular group of students, these papers add significantly to our understanding of both the benefits of an experimental research design and the latest research in the field of adolescent literacy. The presenters argue that while there are substantial challenges to conducting rigorous experiments in a school setting, the results can advance our understanding of both practical research methodology and the issue studied. The goal of this symposium is to present information regarding the benefits of such studies in the context of their complications.
Appendix A. References


