Abstract Title Page

Title: Using Randomized Clinical Trials to Determine the Impact of Reading Intervention on Struggling Adolescent Readers: Reports of Research from Five Nationally Funded Striving Readers Grants

Author(s):

William Loadman, Ohio State University, Kim Sprague, Brown University, Jennifer Hamilton, Westat Inc., Deb Coffey, Research for Better Schools, Bonnie Faddis, RMC Research Corporation

Paper 1: Springfield/Chicopee, MA
Kim Sprague
The Education Alliance, Brown University
kim_sprague@brown.edu

Paper 2: Newark, NJ
Jennifer Hamilton
Westat, Inc.
jenniferhamilton@westat.com

Paper 3: Memphis, TN
Deb Coffey
Research for Better Schools
coffey@rbs.org

Paper 4: Ohio
William Loadman, Richard Lomax, Raeal Moore,
Ohio State University
loadman.1@osu.edu
Jing Zhu
Metis and 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 Black, and Hispanic students, the picture is even more discouraging: 49 percent of Black, 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 no longer "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 five Striving Readers sites that use READ 180.

The five sites in which the program has been implemented 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 includes the high school system for incarcerated youth throughout the state. Further, the READ 180 program was adapted according to the needs and requirements of the different sites. These differences in the local context and in the adaptive use of the program present an important opportunity to examine the different outcomes of experimental analyses as well as rigorous, thoughtful secondary analyses that examine how implementation might be associated with outcomes. Combining experimental data with quasi-experimental statistical methods enabled the evaluators to leverage more information from their study designs than otherwise would be possible.

Presenters will discuss the causal evidence as well as important secondary analyses including techniques developed and/or used during data analyses to correct for potential biases related to implementation fidelity and to explore potential relationships between implementation fidelity and student outcomes.

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 three 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, all five schools were eligible for Title I. In Springfield, twenty-nine percent of the students were African American, 52 percent

Population / Participants / Subjects: Springfield subjects 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).

Memphis: Eight middle schools serving over 6000 students in grades six through eight in Memphis, Tennessee are participating in the Striving Readers project. More than ninety percent of the students served by these schools are African American and 5 percent are Hispanic. Eighty-eight 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 a female facility, all participating. 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. In May 2006, 1,628 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: Four high schools, three middle schools, and three K–8 schools currently participate in the Portland Striving Readers grant. All 10 of these schools receive Title I funding; 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.

All Striving Reader grants include students who are reading two levels below their grade level (in some cases students who are not below a fourth grade reading level).

<u>Springfield/Chicopee:</u> Ninth-grade students are eligible and randomly assigned to one of the two targeted interventions or to the control condition: equally distributed among groups. Students are screened using the Scholastic Reading Inventory. Of the 334 students eligible in 2006-07, 285 were placed in the targeted interventions or in the control group in the fall.

<u>Newark:</u> 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 first year of the study, 1,371 students participated in the evaluation.

<u>Memphis:</u> Eligible students score in the bottom quartile of the state assessment, the Tennessee Comprehensive Assessment Program (TCAP). In the first year of the study, 1740 struggling readers were randomized to the treatment or control group.

<u>Ohio:</u> 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 year of the study, 795 eligible students were randomly placed but only 461 receive instruction due intake complications. The remaining students were either released early, were on lock down, or did not participate.

<u>Portland:</u> Eligible students' grades 7 through 10 were placed using the Oregon State Assessment Test or the Group Reading Assessment and Diagnostic Evaluation. In the first year of the study, 1,311 students were randomly assigned and 586 students were included (had both pre- and post-test scores).

Intervention / Program / Practice:

The programs presented include versions of READ 180, developed by Scholastic Inc., and Xtreme Reading, developed by the University of Kansas Center for Research on Learning. In Springfield-Chicopee and Memphis, 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.

READ 180 is a commercial program from Scholastic that is designed to improve the reading skills of students in fourth through twelfth grades. The program model includes "differentiated instruction, adaptive and instructional software, high-interest literature, and direct instruction in reading, writing, and vocabulary skills" (Scholastic, n.d.). The program calls for 90-minute classes with rotations between whole-group instruction, small-group instruction, individual use of *READ 180* software, and independent reading of leveled fiction and nonfiction texts.

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 approximately 42 hours of training and coaching visits, generally one per month, for each intervention. 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.

Memphis: Teachers were to participate in two all-day summer 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 as needed. Students can receive the targeted intervention for two years.

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 will 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 7 days during the summer and monthly half-days during the academic year. In the first 2 years of the project, 2 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. 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: 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.

Memphis: 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
Memphis: 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

Results

Exhibits 1 & 2 in the appendix depict the cross-site impact results in standard deviation units for middle school and high school students respectively based on two years of data. The presentations will present data from three full years of implementation from five of the eight sites (those five using the Read 180 program).

Springfield/Chicopee:

Findings / Re **Setting:**

Specific description of where the research took place.

Nineteen middle schools participate 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).

Population/Participants/Subjects:

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

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.

Intervention/Program/Practice:

Specific description of the intervention, including what it was, how it was administered, and its duration.

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.

Research Design:
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.

Data Collection and Analysis:
Stanford Achievement Test-10, New Jersey State Language Arts assessment. The data will be analyzed using Hierarchical Linear Model (HLM) software fitting a linear two-level model with student and school as the two levels. The dependent variable of the model is student achievement scores. The primary independent variable is an indicator for treatment or control status at the school level. Other variables are included for student gender, ethnicity, special needs status, socio-economic status, whether supplemental reading services were received, and limited English proficiency status. Additional school controls included the number of students eligible for the program and the school's SINI status.

Findings/Results:
Newark: The impact of READ 180 on eligible students will be estimated for one, two, and three years after program implementation. The results that we propose to present will be 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: female students, male students, African-American students, Hispanic students, and special education students. Results from analyses of year two data suggested program effects vary considerably among demographic groups and may be obscured by aggregate impact estimates.

Memphis: The Mid-south Striving Readers Project (MSRP) includes eight middle schools in a large urban school district. All eight schools have large majorities of African American students; only one school has a substantial minority of ELL, mostly Spanish-speaking, students. All eight are Title I schools, and all eight scored lower than 95 percent of U.S. middle schools on the 2005 Iowa Test of Basic Skills (ITBS) reading test. The purpose of this paper is two-fold: 1) to present findings of secondary analyses of the second year of data, and 2) to relate preliminary findings from the third year of MSRP.

The intent-to-treat (ITT) evaluation is a true experimental design that employs student-level randomization into control (students who receive "business-as-usual") and treatment (students in *READ 180*) groups. The outcome measures are the state assessment given to all students and the ITBS reading subtest. Evaluation findings are based on cross-sectional and long-term hierarchical linear modeling (HLM) analyses of control and treatment students overall and in several subgroups (e.g., race/ethnicity, sex, grade).

Results from years one and two indicate that there are no statistically significant differences between control and treatment student outcomes either as a group or in the various subgroups. However, these findings do not address in any way why *READ 180* did not improve reading outcomes treatment students more than the "business-as-usual" that control students experienced. There are many potential reasons for this; therefore, two avenues of secondary analyses were explored.

    A. To explore the potential effects of cross-contamination, evaluators used Bloom's equation to convert the intent-to-treat (ITT) estimates into treatment-on-the-treated

(TOT) estimates. These adjustments, while arguably more accurate, are very small and do not change the statistical signficance of the findings.

B. *READ 180* is a supplemental program in MSRP, and the configurations of classes in the eight treatment schools resulted in treatment students receiving a total between 110 and 140 minutes of ELA and *READ 180*. Control students receive their regular ELA classes and may also receive instruction in related classes such as creative writing or reading. To explore the potential impact of the sheer amount of literacy-related instruction, evaluators regressed the total number of minutes against student gain scores. These findings showed no statistically significant relationship between the amount of time spent in ELA and/or reading classes and outcomes measures (multiple regression analyses that included student demographic variables also found no statistically significant relationship).

The second subject of this paper will be to present findings of true experimental analyses of year three data of the Mid-south Striving Readers Project. Evaluators will again use HLM in cross-sectional and long-term analyses, which will again include examination of treatment and control students overall and in several subgroups. The seven outcome measures are again, the four content area subtests (language arts, social studies, science, and mathematics) of the state assessment and the three portions of the ITBS reading assessments.

Ohio:

Portland:

The Portland Public Schools (Oregon) Striving Readers grant serves 4 high schools and 6 middle schools using the Strategic Instruction Model—Content Literacy Continuum developed by the University of Kansas' Center for Research on Learning. The targeted intervention is the Xtreme Reading curriculum, which is designed for middle and high school students who need explicit strategy instruction. Xtreme Reading focuses on 6 strategies: word mapping, word identification, self-questioning, visual imagery, paraphrasing, and inference. The whole school intervention uses a set of Content Enhancement Routines (CERs) designed to help all students understand the key content in their courses. Frequently used CERs included the Course Organizer, Unit Organizer, Frame, Concept Mastery, and Question Exploration routines.

Students eligible for the targeted intervention were randomly assigned to treatment and control conditions within each school and grade level. This presentation focuses on the impact of the targeted and whole school interventions over the past 3 years. A multilevel model was used to estimate the impact of the targeted intervention on spring GRADE normal curve equivalent (NCE) scores, controlling for fall GRADE scores, grade level, ethnicity, and English proficiency. The analysis revealed significant intervention effects for middle school students but not high school students. (Total sample sizes included 582 treatment group and 649 control group students in Grades 7–10). In addition, students in the treatment group showed significant improvement in their reading motivation compared to students in the control group (Year 3 only). The presentation will include a discussion of implementation data (professional development participation, classroom observations, teacher instruction logs, teacher surveys) and their relationships to student outcomes.

An Interrupted Time Series (ITS) analysis was used to evaluate the effect of the whole school intervention on state reading achievement test scores for students in Grades 6–10, beginning 3 years prior to the Striving Readers grant through the 3rd year of implementation to date (Spring 2009). (Analysis findings will not be complete until late fall 2009).
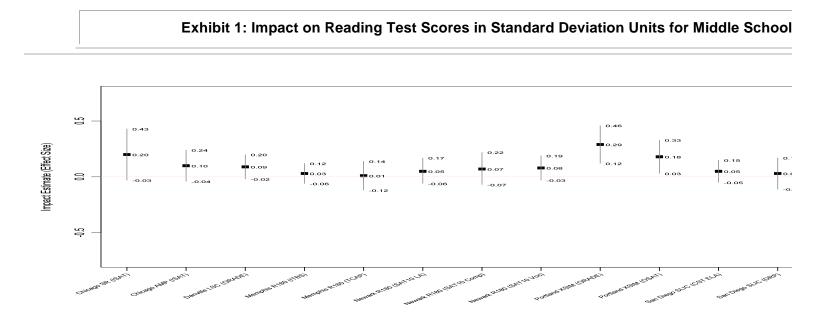
Conclusions:

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.
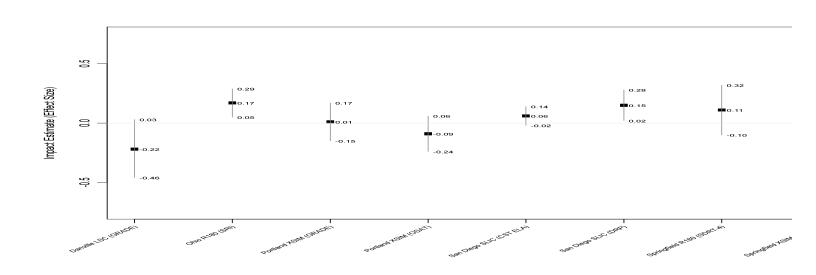
Appendices.

Appendix A. References

References

Ayers, J., & Miller, M. (2009, June). Informing Adolescent Literacy Policy and Practice: Lessons Learned from the Striving Readers Program. Washington, DC: Alliance for Excellent Education. Retrieved July 12, 2009, from http://www.all4ed.org/publication_material/issue_policy_briefs

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York,* 2nd ed. Washington, DC: Alliance for Excellent Education. Retrieved April 7, 2009, from http://www.all4ed.org/publication_material/adlit

Scholastic. (n.d.) *READ 180* Reading Intervention Program: A Comprehensive Reading Intervention Solution. Retrieved July 13, 2009, from http://teacher.scholastic.com/products/read180/overview/

U. S. Dept. of Education. National Center for Education Statistics. Institute of Education Sciences. 2007. The Nation's Report Card: Reading 2007. Retrieved July 12, 2009, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007496

Summary of Striving Readers Projects:
Targeted Interventions for Struggling Readers

Exhibit 1: Impact on Reading Test Scores in Standard Deviation Units for Middle School

e: Bold square is the estimated impact on reading test scores in standard deviation units. The line represents the confidence interval of the impac
rlaps with zero, the pink line, the estimated impact is not significant.

Impact Estimate (Effect Size)

Data point	Values
Danville LSC (GRADE)	0.03 / -0.22 / -0.46
Ohio R180 (SRI)	0.29 / 0.17 / 0.05
Portland XSIM (GRADE)	0.17 / 0.01 / -0.15
Portland XSIM (OSAT)	0.06 / -0.09 / -0.24
San Diego SLIC (CST ELA)	0.14 / 0.06 / -0.02
San Diego SLIC (DRP)	0.28 / 0.15 / 0.02
Springfield R180 (GDRT-4)	0.32 / 0.11 / -0.10
Springfield XSIM	

Source: Striving Readers Year 2 Evaluation Reports (available at: *http://www.ed.gov/programs/strivingreaders/performance.html*)
Note: Bold square is the estimated impact on reading test scores in standard deviation units. The line represents the confidence interval of th
overlaps with zero, the pink line, the estimated impact is not significant.

Exhibit 2: Targeted Interventions: Impact on Reading Test Scores in Standard Deviation Units for High School