Abstract 1 Title Page
Not included in page count.

Title: Experimental Effects of the Strategic Adolescent Reading Intervention on Reading Performance in High Poverty Middle Schools

Authors and Affiliations:
Lowry Hemphill, Wheelock College
James Kim, Harvard University
Monica Yudron, Harvard University
Maria LaRusso, Harvard University
Suzanne Donovan, SERP
John Sabatini, ETS
Tenaha O'Reilly, ETS

Abstract 1 Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

The roughly one-quarter of U. S. eighth graders who score below basic on national assessments of reading are poorly equipped for the reading demands of secondary school. They struggle with summarizing and making text-based inferences (NCES, 2013). Weak decoding, word knowledge, and fluency may limit their ability to process text efficiently (Verhoeven & van Leeuwe, 2008).

Intervention in middle school needs to be comprehensive, both because of the heterogeneity of the population of struggling readers and because of the interdependence of components of reading. Effective interventions for adolescents address gaps in basic reading skills while also promoting the analytic and interpretive reading skills important in upper grades (Faggella-Luby, Graner, Deshler, & Drew, 2012; Lee & Spratley, 2012).

Designing interventions for older readers that are effective at scale has proven difficult. The Enhanced Reading Opportunities study, for example, found only small improvement in comprehension of ninth graders in two literacy interventions (Somers, Corrin, Sepanik, Salinger, Levin, & Zmach, 2010). Investigation of a sixth grade strategy intervention found only minimal impacts for participants' reading comprehension, despite gains in reported strategy use (Cantrell, Almasi, Rintmaa, Carter, Pennington, & Buckman, 2014). Recent meta-analyses of adolescent reading interventions have concluded that effect sizes are typically small, especially for reading comprehension (Edmonds, Vaughn, Wexler, Reutebuch, Cable, & Tackett, et al., 2009; Flynn, Zheng, & Swanson, 2012).

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

We present preliminary results of a clinical trial of a new multi-component program, the Strategic Adolescent Reading Intervention (STARI). Questions guiding the analysis were:

(1) What is the intention-to-treat (ITT) estimate of STARI on the Reading Inventory and Scholastic Evaluation (RISE) for a sample of struggling readers in Grades 6 to 8?
(2) What is the treatment-on-the treated (TOT) estimate of STARI on the RISE subtests?

Setting:

Description of the research location.

Four Massachusetts school districts served as research sites: two large urban districts, District A and B, and two rural/suburban districts, District C and D. Districts volunteered to be part of the study and solicited schools to participate (in the case of the larger districts) or had all their middle schools participate in the two smaller districts.

In District A, percentages of students in the four participating schools scoring below proficient on the 2013 state ELA assessment ranged from 55 to 79. Between 79 and 91 percent of the student population was African American or Hispanic in these schools and between 83 and 92 percent were eligible for free or reduced price lunch, a common measure of family poverty.

In District B, percentages of students in the two participating schools scoring below proficient in ELA were 34 and 41 percent. Forty-one to 58 percent of the student population was African American or Hispanic and 66 to 87 percent were eligible for subsidized lunch.

In District C, percentages of students in the two middle schools below proficient in ELA were 30 and 39. Thirty-three to 36 percent of students were African American or Hispanic; 49 percent of students at both schools were eligible for subsidized lunch.

In District D, 38% of the students scored below proficient in ELA. Seventeen percent were African American or Hispanic and 44 percent were subsidized lunch eligible.

Population / Participants / Subjects:

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

In each of the participating schools, students scoring below proficient on the state ELA assessment were eligible to participate in STARI. Students in substantially separate special education classes, students who were level 1 or 2 English language learners, and students whose special education plan included intensive, rules-based phonics intervention, were excluded from study participation. The remaining students were randomized into treatment or control conditions. Students assigned to the treatment condition were rank ordered as part of the randomization and schools were asked to place students into the available seats in intervention classes, following their rank orders.

Intervention / Program / Practice:

Description of the intervention, program, or practice, including details of administration and duration.

The Strategic Adolescent Reading Intervention (STARI) is a comprehension-focused middle school reading intervention developed by author and the Strategic Education Research Partnership Institute (SERP). STARI is delivered in one teaching block for 3-5 days per week. Teachers receive project-authored student workbooks for both fluency and comprehension practice, unit novels and non-fiction books, slides, and detailed daily lesson plans. In each block, students worked for 15 minutes on leveled fluency passages, engaging in timed reading, repeated reading, phrase-cued reading, and practice with isolated words and phrases. Fluency work was carried out in partners and each two-day fluency cycle included a discussion with the partner about a controversy or personal response to the fluency passage, as well as checks on literal comprehension.

Comprehension strategies were introduced in teacher mini-lessons and practiced and developed through 30-40 minutes daily spent reading program novels and nonfiction books. Work with program novels and non-fiction was organized into three thematic units. Students alternated reading in small groups, guided by the teacher, and reading in partners, with workbook prompts to engage in comprehension strategies such as questioning, summarizing, and predicting. Each unit included one or more debates about a controversial issue raised in the unit novels or nonfiction. Students gathered information from the unit texts to support positions in the debate and prepared debate presentations working in teams.

Key features of the STARI intervention were extensive reading of text on unit themes of importance in students' lives (bullying, the war in Iraq, diverse families), multiple opportunities for students to talk about text, and a focus on developing personal stances on text content.

Teachers were introduced to the program through a three-day summer institute and received regular in-class guidance from a project literacy coach. STARI teachers met regularly as part of district-based professional learning communities and participated in three, day-long workshops during the school year, focused on talk about text, comprehension strategies, and feedback on the curriculum.

Research Design:

Description of the research design.

We implemented within-school lotteries in each of the 8 schools across the 4 districts to estimate impacts on student reading outcomes. Eligible students were assigned lottery numbers, ranked order from the lowest to the highest number, and then assigned to STARI classes. Overall, among the 390 students with fall baseline RISE reading scores, 55% (n = 213) were assigned to STARI classes and 45% (n = 177) were assigned to a counterfactual condition, which ranged from enrichment classes that did not include English language arts instruction to supplemental tutoring in a range of subjects. Because random assignment to STARI or control is conditional on school blocks, we included school fixed effects in all analytic models.

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

In this study, we present treatment effects on posttest reading measures, using pretest and grade level as covariates, and school fixed effects to account for the lottery design. In fall 2013, children were pretested on the Reading Inventory and Scholastic Evaluation (RISE), which was developed by a team of researchers at ETS (Sabatini, O'Reilly, Halderman, & Bruce, 2014; Sabatini, Bruce, & Steinberg, 2013).The RISE is a 45 to 60 minute web-administered assessment that assesses six subtests that underlie reading proficiency in the middle grades. The subtests include: (1) Word Recognition & Decoding (50 items, 6 minutes, $\alpha = .91$), (2) Vocabulary, (38 items, 6 minutes, $\alpha = .86$), (3) Morphological Awareness, (32 items, 7 minutes, $\alpha = .90$), (4) Sentence Processing, (26 items, 7 minutes, $\alpha = .81$), (5) Efficiency of Basic Reading Comprehension (36 items, 9 minutes, $\alpha = .90$), and Reading Comprehension (22 items, 20 minutes, $\alpha = .76$). The RISE is designed for students in grades 6 to 8 and provides data that can inform decisions about literacy instruction at the district, school, and classroom levels. Since we use the pretest as a covariate to improve the precision of the estimated treatment effects in all analytic models, we included students with both a pretest and posttest score. The analytic sample ranges from 388 to 390 students. Students in the STARI and control conditions had statistically equivalent scores on each of the six RISE pretests.

For STARI students, we created an exposure measured based on the % of the curriculum that children completed. There were three units in the STARI curriculum; therefore, the exposure rates ranged from .33 (1 unit completed) to .67 (2 units completed) to 1.00 (3 units completed). The unit completion rate for each student was used as the exposure measure and a proxy for whether students completed the entire curriculum (units 1, 2, and 3).

Findings / Results:

Description of the main findings with specific details.

Question #1: Intention-to-treat effects

To address the first question, we use ordinary least squares (OLS) regression models to identify the causal effect of being randomly assigned via a lottery process to STARI regardless of the number of units that students completed from fall 2013 to spring 2014. Thus, the ITT estimates offer an unbiased estimate of the treatment effect if students were offered the opportunity to participate in the STARI intervention. Using OLS regression, we used the following model to generate an unbiased intention-to-treat estimate of STARI on each of the 6 RISE subtests:

(1) $Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \beta_3 RB_i + \varepsilon_i$

where Y_i represents the RISE subtest score outcome for student i, T_i represents whether the student was randomly assigned to STARI, X_i is a vector of student background variables (i.e., the pretest score, grade), RB_i represents the school fixed effect, and ε_i represents the error term. The coefficient β_1 is the estimated difference in posttest scores between treatment and control students and represents the intention-to-treat estimate on each RISE subtest. As shown in Table 1, the ITT effects were positive and statistically significant for 5 out of the 6 outcomes. Using the pooled within group standard deviation for each posttest to compute Cohen's *d*, the effect size for the STARI impact was d = .23 for word recognition and decoding, d = .20 for vocabulary, d = .16 for morphology, d = .23 for sentence processing, d = .22 for efficiency of basic reading, and d = .05 (n.s.) for reading comprehension.

Question #2: Treatment-on-the treated effects

To address the second question, we estimated the treatment-on-the-treated (TOT) estimate, using initial random assignment status as an instrument for percentage of STARI units that students in the treatment actually completed during the fall 2013 to spring 2014 school year. Initial random assignment is likely to be a valid instrumental variable if it predicts STARI curriculum completion rates, is uncorrelated with the residuals in the second stage test score equation, and influences test scores exclusively through a student's participation in STARI.

We used instrumental variables analysis in two stages. The fully specified first stage model takes the form:

(3) $Z_i = \pi_0 + \pi_1 X_i + \pi_2 T_i + \pi_3 RB_i + \delta_i$

where T_i represents initial random assignment status and serves as the instrument for STARI curriculum completion rates, X_i is the pretest score, RB_i denotes the school by grade randomization blocks, and δ_i is the error term. Using equation 3, we estimated Z'_i, which is the predicted value in the STARI curriculum completion rates based on initial random assignment status. The second stage model is written as:

(4) $Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z'_i + \beta_3 RB_i + \varepsilon_i$

where the posttest reading score is predicted by Z'_i and the same independent variables that were included in the first stage model. In the second stage model, the coefficient Z'_i yields a treatment-on-the-treated estimate that indicates the impact of completing the full STARI curriculum on the posttest RISE reading outcomes. The treatment-on-the treated effects were slightly larger than the ITT effects, suggesting that full exposure to the STARI curriculum improved student outcomes. Effects for the TOT estimates were the following: d = .26 for word recognition and decoding, d = .24 for vocabulary, d = .22 for morphology, d = .28 for sentence processing, d = .25 for efficiency of basic reading, and d = .07 (n.s.) for reading comprehension

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

Our evaluation results of STARI suggest that a multi-component Tier-2 intervention for struggling adolescent readers can improve a range of reading outcomes including students' ability to decode words, to read connected text with speed and accuracy, to improve their knowledge of word parts, and to improve sentence and text-level comprehension outcomes. By the middle grades, it may be critical to engage struggling readers with curriculum units that tap student interests and provide opportunities to talk about text with teachers and peers. In keeping with the conference theme, we will also present results on the extension of treatment effects over time on other reading comprehension measures that were administered at posttest, including a state standardized reading test.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Cantrell, S.C., Almasi, J., Rintmaa, M., Carter, J. C., Pennington, J., & Buckman, D. M. (2014). The impact of supplemental reading instruction on low-achieving adolescents' reading engagement. *Journal of Educational Research, 107,* 26-58.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79,* 262–300.

Faggella-Luby, M. N., Graner, P. S., Deshler, D. D., & Drew, S. V. (2012). Building a house on sand: Why disciplinary literacy is not sufficient to replace general strategies for students who struggle. *Topics in Language Disorders, 32*, 69-84.

Flynn, L. J., Zheng, X., & Swanson, H. L. (2012). Instructing older struggling readers: A selective meta-analysis of intervention research. *Learning Disabilities Research and Practice, 27,* 21-32.

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenge of adolescent literacy.* New York: Carnegie Corporation of New York.

National Center for Education Statistics. (2013). *The nation's report card: A first look: 2013 mathematics and reading.* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Sabatini, J. P., Bruce, K., & Steinberg, J. (2013). *SARA reading components tests, RISE form: Test design and technical adequacy* (ETS RR-13-08). Princeton, NJ: ETS.

Sabatini, J., O'Reilly, T., Halderman, L. K., & Bruce, K. (2014). Integrating scenario-based and component skill reading skill measures to understand the reading behavior of struggling readers. *Learning Disabilities Research & Practice, 29*(1), 36-43.

Somers, M.-A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (2010). The Enhanced Reading Opportunities study final report (NCEE 2010-4021). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Verhouven, L. & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22,* 407-423.

Appendix B. Tables and Figures

Not included in page count.

Table 1: Intention to treat impacts of the STARI intervention on RISE subtests

	Word Recognition		Vocabulary		Morphology		Sentence Processing		Basic Reading		Reading Comprehension	
	β (SE)		β (SE)		β (SE)		β (SE)		β (SE)		β (SE)	
Intercept	129.60	***	170.66	***	92.41	***	197.22	***	106.79	***	159.03	***
	(14.48)		(15.34)		(17.17)		(17.63)		(17.69)		(19.47)	
Fall 2013 Score	0.62	***	0.54	***	0.72	***	0.41	***	0.70	***	0.55	***
	(0.04)		(0.04)		(0.05)		(0.05)		(0.05)		(0.05)	
STARI	**5.93**	***	**4.21**	*	**5.23**	*	**5.93**	*	**6.48**	***	**1.40**	
	(2.07)		**(1.81)**		**(2.20)**		**(2.42)**		**(2.35)**		**(2.53)**	
N	390		389		388		389		389		386	
R^2	0.44		0.37		0.44		0.21		0.41		0.27	

Table 2: Treatment on the treated impacts of the STARI intervention on RISE subtests

	Word Recognition		Vocabulary		Morphology		Sentence Processing		Basic Reading		Reading Comprehension	
	β (SE)		β (SE)		β (SE)		B (SE)		β (SE)		β (SE)	
Intercept	137.80	***	172.82	***	101.02	***	201.60	***	117.64	***	157.22	***
	(13.73)		(14.15)		(16.11)		(16.29)		(16.30)		(18.22)	
Fall 2013 Score	0.62	***	0.54	***	0.71	***	0.40	***	0.68	***	0.55	***
	(0.04)		(0.04)		(0.05)		(0.05)		(0.05)		(0.05)	
STARI exposure	**6.69**	***	**5.09**	*	**6.03**	*	**7.09**	*	**7.29**	**	**1.88**	
	(2.49)		**(2.17)**		**(2.64)**		**(2.88)**		**(2.81)**		**(3.02)**	
N	390		389		388		389		389		386	
R^2	0.43		0.36		0.43		0.22		0.40		0.27	