Abstract Title Page
Not included in page count.

Title: The Effects of Collaborative Strategic Reading Instruction on the Reading Comprehension of Middle School Students: Year 2 Replication

Author(s): Elizabeth Swanson, Sarojani S. Mohammed, Alison Gould Boardman, Sharon Vaughn, Janette Klingner, Greg Roberts, Audrey Leroux, and Michael Solis

Abstract Body

Limit 5 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

Collaborative strategic reading (CSR; Klingner, Vaughn, Dimino, Schumm & Bryant, 2001) is a fully developed, feasible intervention, with evidence of its efficacy established through quasi-experimental research studies. Built on a foundation of reciprocal teaching (Palincsar & Brown, 1984) and many features associated with effective instruction (e.g., collaborative group work, interactive dialogue, procedural strategies), CSR addresses three prevailing educational challenges: (a) how to teach text comprehension strategies that improve students' reading comprehension, (b) how to adequately include struggling readers in text-related learning using grade-level text, and (c) how to provide opportunities for English language learners to interact effectively with peers and enhance their achievement.

Early studies of CSR focused on evaluating effectiveness within science and social studies content area instruction. In one such study (Klingner, Vaughn & Schumm, 1998), CSR was taught to intact, heterogeneous fourth grade classes for 45 minutes per day during an 11-day Florida state history unit. The comparison group of classes received instruction reflective of the school's typical practice. Students in the CSR group made greater gains in reading comprehension and equal gains in content knowledge. To determine whether these findings would be upheld within science instruction, 5th graders were provided with CSR instruction for 30-40 minutes per day, two to three days per week, over a four-week period during science (Klingner & Vaughn, 2000). Students frequently engaged in verbal discourse that supported vocabulary and content knowledge development; in addition, students made gains in target vocabulary over time. In a subsequent quasi-experimental study with 4th grade teachers, teachers participating in the treatment condition were provided CSR training and in-class demonstrations. A comparison group of teachers continued typical practice instruction. On a norm-referenced measure of reading comprehension, students in the CSR group outperformed students in the typical practice comparison group (Klingner, Vaughn, Argüelles, Hughes, & Ahwee, 2004) . Similar gains were demonstratedwhen third grade teachers received either CSR or partner reading training. Students in both condisions made significant gains in tests of oral reading rate and accuracy as well as reading comprehension (Vaughn et al., 2000), providing additional evidence for the use of CSR with upper elementary students.

There have also been three studies of CSR at the middle school level. In one study, researchers developed a computer adapted version of CSR (Kim et al., 2006) used with sixth through eighth grade students with LD who were randomly assigned to either a computer based CSR intervention or a typical school practice comparison group. Students in the CSR group outperformed students in the comparison group on the Gates MacGinitie. CSR was also used as one of several intervention practices designed to enhance overall school-wide reading comprehension (Bryant et al., 2000). In this study, students demonstrated gains on word identification but not reading comprehension. In the first randomized control trial of CSR (Vaughn et al, in press), the intervention was implemented by trained English/Language Arts (ELA) teachers with 7th and 8th grade students. There was a small, significant main effect of CSR on reading comprehension as measured by the Gates-MacGinitie assessment. Though the effect was not statistically significant for struggling readers, the effect size (g=.36) suggests that the influence of CSR has practical significance for this group of students

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The current study is the second in a series of multi-site, multi-year randomized control trials designed to test the efficacy of a fully developed intervention, CSR, with adolescent readers. In year 1, our research questions were: (1) Does CSR improve reading comprehension for adolescent readers attending relatively low SES schools?, and (2) Does CSR improve reading comprehension for adolescent *struggling* readers attending relatively low SES schools? During year 2, we replicated the year 1 study with a new cohort of students taught by the same teachers.

Setting:

Description of the research location.

This study was conducted in 6 middle schools in Texas and Colorado. Schools were chosen that met the following criteria: (a) students with reading difficulties were taught in English/Language Arts (E/LA) classes; (b) some of these students also received instruction in reading intervention classes for struggling 7th and 8th graders; and (c) socio-economic status of students attending the school were low to moderate.

Population / Participants / Subjects:

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

Teachers were eligible for participation if their classes met the following criteria: (a) 7th and/or 8th grade students in the class; (b) class was designated as covering primarily E/LA content; and (c) class was not specially designated as an Advanced Placement or Special Education class. Seventeen teachers agreed to participate in the study during year 1. Due to changes in job placement, several teachers' classes no longer qualified for inclusion in the study. Therefore, thirteen of these teachers continued during year 2. Students in this study were two separate cohorts of 7th and 8th graders enrolled in English/Language Arts classes in 2 schools in Texas and 4 schools in Colorado. Student demographic information is provided in Table 1. Teacher demographic information is provided in Table 2 (Appendix B).

Intervention / Program / Practice:

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

Description of Intervention: CSR helps students learn specific strategies associated with enhanced reading comprehension: (a) activating prior knowledge and predicting (*preview*), (b) monitoring understanding (*click and clunk*), (c) finding the main idea (*get the gist*), and (d) generating questions and reviewing key ideas (*wrap up*). *Preview* occurs prior to reading and consists of making predictions, connecting to students' prior knowledge and associations with the text, generating interest, and encouraging active reading of the text. *Click and clunk* occurs during reading and refers to the process of reading for meaning (clicking) and monitoring comprehension so that students notice when understanding breaks down (clunking). Students are taught to use several "fix-up" strategies (e.g., "Read the sentence before and after the clunk. Look for cues.") to figure out unknown words or concepts (i.e., words they do not know the meaning of; not word accuracy reading). Students also *get the gist* during reading by stopping after each paragraph or section to find the main idea or summarize key information. Students are taught to identify the most important who or what in the paragraph or section they have just read and then to briefly state the critical information about the who or what. *Wrap-up* takes place after reading. Students generate and answer questions about what they have read, and summrize key ideas presented in the text.

Initially, the teacher presents the strategies to the whole class using explicit instruction, modeling, and teacher think-alouds. After students have developed proficiency using the strategies, the teacher then assigns them to cooperative learning groups (approximately four students per group) in which each student plays a critical role associated with the effective functioning of the group and the implementation of strategies (e.g., leader, clunk expert, gist pro) (Johnson & Johnson, 1989). Hence, with CSR, all students are actively involved and have an opportunity to contribute to the group's understanding of text.

Administration and Duration: Teachers were asked to deliver CSR lessons 2-3 times per week over a 26-week period between September and May. During year 1, teachers reported implementing between 23 and 52 sessions total. CSR lessons lasted between 10 and 95 minutes, with a median length of 45 minutes and a mode length of 45 minutes. During year 2, teachers reported implementing between 18 and 61 sessions total. CSR lessons lasted between 5 and 90 minutes with a median length of 45 minutes and a mode length of 45 minutes.

Research Design:

Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

We conducted a randomized field trial to compare the effects of the CSR program to school-designed comparison interventions (controlled for instructional time) in 7th and 8th grade English and reading classrooms across 2 schools in TX and 4 schools in CO. Students were randomly assigned to class and then classes were randomly assigned within teacher. 61 classes during year 1 and 48 classes during year 2 were randomly assigned within teacher to either treatment or comparison condition. For teachers with an odd number of classes, the additional class was assigned to the treatment condition.

Students in the comparison group received the schools' typical instruction and intervention support. All students with permission to participate were included in pre- and post-testing. We assessed a range of word identification, fluency, and comprehension skills at pretest and immediate posttest. We also collected data on student characteristics (e.g., language and special education status, age, gender, reading proficiency) to examine comparability of groups. Students were considered "struggling" based on failure to pass the previous year's state reading test *and* a pretest standard score of less than 85 (i.e., one standard deviation below the mean) on the Test of Word Reading Efficiency (TOWRE).

Because the same teacher provided instruction for both the treatment and comparison conditions, with the students randomly assigned to condition, we controlled the effect of an individual teacher accounting for a significant amount of variance. To guard against treatment contamination into comparison classrooms, we explained to teachers the importance of their contribution to the validity of findings in an experimental study and reinforced the requirement to use the CSR practices only with the treatment group and to continue to use their usual instructional practices with the comparison classes.

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

Data Collection: Students in all classes (in years 1 and 2) were administered a battery of measures at pretest, and then provided a battery of tests at posttest. Pretest and posttest measures were administered two to three weeks prior to intervention and within two weeks post-

intervention, respectively. All assessment data were collected by trained data collectors who were blind to treatment condition, and who demonstrated at least 90% reliability on administering and scoring all measures. The reading achievement battery included the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) the Test of Sentence Reading Efficiency (TOSRE; Wagner, Torgesen, Rashotte, in press), AIMSweb Maze passages for 7th and 8th Grades (AIMSweb Maze-CBM, 2009), and the Gates-MacGinitie Reading Test (Gates & MacGinitie, 2000).

Data Analysis: Classrooms rather than students were randomly assigned to the intervention condition, threatening the assumption of independence among participants. Multilevel modeling (structural equation modeling) accounts for the data's multilevel structure while also offering the advantage of direct full information maximum likelihood (FIML) estimation of missing data, more appropriate modeling of the covariance structures of clustered data, and estimates of model fit (used to evaluate a given model's accuracy, as a tool for comparing models, and as a means of evaluating statistical significance). Multilevel modeling in Mplus 5.1 was used to estimate the effects of treatment and the moderating influence of important covariates. Teacher was treated as a *stratum* for purposes of assignment, and classes (both treatment and comparison) were randomly assigned within teachers. Analytically, this represents a randomized block design with teachers as the blocking variable (Raudenbush, 1997) and students nested in classes. A pretest score (cluster-level covariate) was included in the model, as a means of minimizing the conditional group-level variance and further increasing precision and power (Bovaird, 2007). In Mplus, this represents a two-level analysis with complex sampling. Classes were represented as *clusters*, which define levels in a multilevel model. In the unconditional (i.e., no moderating covariates) student-level model, posttest scores were regressed on the corresponding grand-mean centered pretest values. Posttest means were modeled as latent factors on the between-classes model. Treatment condition was modeled using the multiple groups option in Mplus, which allowed for formal tests of statistical significance using a nested models comparison.

Findings / Results:

Description of the main findings with specific details.

For each year, main effects were estimated for the Gates-MacGinitie, for the AIMSweb Maze, and for the Test of Sentence Reading Efficiency (TOSRE) according to the multilevel model described above. The analyses were conducted with the full sample and with the sample of students identified as struggling readers. For the full sample, an unconditional multi-group, multilevel model was fit to estimate posttest class-level means (i.e., level 2) conditioned on the student-level (level 1) model and on the earlier-described adjustments for clustering and stratification for each of the three outcomes. These models were saturated, because there were as many parameters as values to fit (accordingly, they have a χ^2 of 0 and 0 degrees of freedom).

Year 1: There were statistically significant main effects for treatment on the Gates-MacGinitie (c2=3.79, p=.05), with a one-point standard score difference between the two groups (96 for comparison 97 for treatment). The Gates had a pooled sample variance of 11.5, so a 1-point group different represents a sample derived effect size of .09. There were no differences on TOSRE (c2=2.32, p=.13) or on AIMSweb (c2=.003, p=.96). When the analysis was restricted to students who scored below passing on the state reading assessment, the effect on Gates was no longer evident (c2=.18, p=.67). The marginal standard score means were 88.8 and 88.4 for treatment and comparison, respectively. There were also no main treatment effects on TOSRE (c2=2.66, p=.10) or on AIMSweb (c2=1.94, p=1.64). There were no differences in effects due to

school or to site.

Year 2: The model-estimated group means on Gates-MacGinitie were 98.67 and 99.74 for comparison and treatment groups, respectively. This difference was not statistically significant ($\Delta\chi2$=.84, p=.36). The pooled standard deviation was 14.18, the effect size was .08 (unbiased Hedges g), and the 95% confidence interval ranged from 1.12 to -1.27. There were no differences on AIMSweb ($\Delta\chi2$=.05, p=.82) or on TOSRE ($\Delta\chi2$=.12, p=.73). When the analysis was restricted to students who scored below passing on the state reading assessment, the effect on Gates was not significant ($\Delta\chi^2$=61, p=43). There were also no main treatment effects on TOSRE ($\Delta\chi^2$=.87, p=.36) or on AIMSweb ($\Delta\chi^2$=2.97, p=.085). Descriptive statistics are presented in Tables 3 and 4 (Appendix B).

Conclusions:

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

In Year 1 of the study, there was a small, significant main effect of CSR on reading comprehension as measured by the Gates-MacGinitie assessment. We consider this finding noteworthy considering the challenge of affecting reading comprehension among older readers. Evidence indicates that CSR is effective with struggling readers. While the main effect on the Gates MacGinite was not statistically significant for struggling readers, the effect size (g = 0.36) indicates practical significance. We believe that the findings from the year 1 study suggest that CSR is a feasible and effective practice that can be readily integrated into reading and language arts instruction with positive impact. In year 2, the study teachers, now experienced users of CSR, participated in a replication trial using the same research design and CSR intervention strategies, but with a new cohort of students. Findings in year 2 indicate an overall gain in reading comprehension for students in both the comparison and intervention conditions. Several explanations shed light on this finding. During year 2, schools in both sites increased the district professional development in reading strategies and teachers were expected to teach these strategies to all students. While teachers continued to teach CSR in intervention classes, they increased their use of strategy instruction in their typical classes. Thus, intervention and comparison classes were more similar to one another in year 2 than in year 1. Second, it is possible that in addition to learning to use reading comprehension strategies, the professional development and researcher support resulted in increased pedagogical knowledge and application. That is, teachers may have improved their quality of instruction overall (e.g., providing feedback to students, selecting appropriate text, class management) and these improvements could have contributed to increased student achievement in all classes. Analyses of results is ongoing and additional findings related to the potential effect of teacher level variables such as the amount of time spent teaching CSR and the fidelity of implementation to the study results will be reported during the presentation.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

AIMSweb Maze-Curriculum Based Measurement (n.d.). Retrieved from http://www.aimsweb.com/measures-2/maze-cbm/.

Bovaird, J. A. (2007). Multilevel structural equation models for contextual factors. In T. D. Little, J. A., Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 149–182). Mahwah, NJ: Erlbaum.

Bryant, D. P., Vaughn, S., Linan-Thompson, S., Ugel, N., & Hamff, A. (2000). Reading outcomes for students with and without learning disabilities in general education middle school content area classes. *Learning Disability Quarterly, 23*(3), 24-38.

Burchinal, M. R., Bailey, D. B., Snyder, P. (1994). Using growth curve analysis to evaluate child change in longitudinal investigations. *Journal of Early Intervention*, *18*, 403-423.

Gates, A. I., & MacGinitie, W. H. (2000). *Gates-MacGinitie Reading Tests*, (4th ed.). Boston, MA: Riverside.

Johnson, D. W., & Johnson, R. T. (1989) Cooperative learning: What special educators need to know. *The Pointer*, 33, 5-10.

Kim, A., Vaughn, S., Klingner, J. K., Woodruff, A. L., Klein, C., & Kouzekanani, K. (2006). Improving the reading comprehension of middle school students with disabilities through computer-Assisted Collaborative Strategic Reading (CACSR). *Remedial and Special Education, 27,* 235-248.

Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *Elementary School Journal, 96,* 275-293.

Klingner, J. K., & Vaughn, S. (2000). The helping behaviors of fifth-graders while using collaborative strategic reading (CSR) during ESL content classes. *TESOL Quarterly, 34,* 69-98.

Klingner, J. K., Vaughn, S., Argüelles, M. E., Hughes, M. T., & Ahwee, S. (2004). Collaborative Strategic Reading: "Real world" lessons from classroom teachers. *Remedial and Special Education, 25,* 291-302.

Klingner, J. K., Vaughn, S., Dimino, J., Schumm, J. S., & Bryant, D. P. (2001). *From clunk to click: Collaborative Strategic Reading.* Longmont, CO: Sopris West.

Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *Elementary School Journal, 99,* 3-21.

Lee, Y., Nelder, J., & Pawitan, Y. (2006). *Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood.* London: Chapman & Hall/CRC

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

Palincsar, A. S., & Brown, A. L. (1984). The reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117-175.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, *2*, 173-185.

Raudenbush, S. W., & Bryk, A. S. (2002*). Hierarchical linear models: Applications and data analysis methods* (2nd edition). Thousand Oaks, CA: Sage Publications.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach*. Seventh Edition. Thousands Oaks, CA: Sage Publications, Inc.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *The Test of Word Reading Efficiency (TOWRE).* Austin, TX: Pro-Ed.

Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., & Stillman-Spisak, S. J. (in press). Efficacy of Collaborative Strategic Reading with middle school students, *American Educational Research Journal*.

Wagner, R. K., Torgesen, J. K. & Rashotte, C. A. (in press). *Test of Sentence Reading Efficiency (TOSRE).* Austin, TX: Pro-Ed.

Appendix B. Tables and Figures
Not included in page count.

Table 1. Student Demographics

	Year 1				Year 2*			
	CSR		TP		CSR		TP	
	n	%	*N*	%	*n*	%	*N*	%
Gender								
Male	199	43.4	188	46.1				
Female	170	37.1	166	40.7				
Ethnicity								
Anglo	172	37.6	136	33.3				
African American	13	2.8	13	3.2				
Hispanic	178	38.9	193	47.3				
Asian	2	0.4	9	2.2				
Native American	4	0.9	3	0.7				
Free or Reduced Lunch	194	42.4	189	46.3				
English Language Learners	10	2.2	11	2.7				
Special Education	56	12.2	28	6.9				

* At the time of proposal preparation, demographic information for year 2 student participants was not available. They will be included in the presentation.

Table 2. Teacher Demographics

	Year 1	Year 2
Male	3	3
Female	14	10
White	16	12
Asian/Pacific Islander	1	1
Average Education	16.9 years	17.38 years
Degrees Earned	9 Bachelor's, 8 Master's	4 Bachelor's, 9 Master's
Experience	10.12 years	11.62 years
Specializations	15 Eng/Lang Arts, 7 secondary, 7 elementary, 6 reading, 2 ESL, 2 Special Education, 1 speech	11 Eng/Lang Arts, 5 secondary, 6 elementary, 5 reading, 1 ESL, 1 Special Education, 1 speech

Note. Eng/Lang Arts = English/Language Arts; ESL = English as a Second Language.

Table 3. Descriptive Statistics on all Measures for the Full Analysis Sample

	Year 1				Year 2			
	Pretest mean (s.d.)		Posttest Mean (s.d.)		Pretest mean (s.d.)		Posttest Mean (s.d.)	
	TP	CSR	TP	CSR	TP	CSR	TP	CSR
Gates-MacGinitie	95.68 (13.4)	96.35 (13.7)	95.48 (13.4)	97.13 (13.6)	99.35 (14.8)	98.93 (14.0)	100.15 (14.5)	100.62 (14.4)
AIMSweb	92.64 (12.1)	91.91 (10.8)	93.46 (11.0)	92.92 (11.0)	95.02 (10.8)	94.12 (10.5)	93.03 (10.4)	92.28 (9.9)
TOSRE	23.49 (7.36)	24.67 (7.17)	29.03 (7.68)	29.22 (7.66)	26.15 (7.8)	25.68 (6.9)	30.58 (8.2)	30.47 (7.9)
TOWRE—SW	90.17 (7.40)	90.59 (8.38)	N/A	N/A	89.69 (9.1)	89.12 (9.4)	N/A	N/A
TOWRE--PD	91.53 (11.3)	91.20 (10.85)	N/A	N/A	89.73 (11.7)	89.24 (11.2)	N/A	N/A

Table 4. Descriptive Statistics on all Measures for the Struggling Readers Sample

	Year 1				Year 2			
	Pretest mean (s.d.)		Posttest Mean (s.d.)		Pretest mean (s.d.)		Posttest Mean (s.d.)	
	TP	CSR	TP	CSR	TP	CSR	TP	CSR
Gates-MacGinitie	84.37 (10.0)	85.62 (11.2)	84.25 (9.08)	87.74 (9.95)	84.32 (11.09)	83.33 (13.10)	86.93 (12.34)	87.07 (9.04)
AIMSweb	85.81 (8.75)	85.13 (8.01)	86.30 (7.97)	86.60 (7.37)	85.51 (7.66)	83.26 (6.52)	84.96 (8.12)	86.63 (8.53)
TOSRE	18.52 (4.60)	20.30 (5.89)	23.87 (6.34)	24.83 (5.76)	17.98 (4.81)	19.15 (5.62)	22.24 (6.17)	23.55 (6.10)
TOWRE—SW	84.48 (5.46)	83.76 (4.62)	N/A	N/A	82.25 (5.74)	80.86 (8.05)	N/A	N/A
TOWRE--PD	81.57 (5.52)	82.55 (5.57)	N/A	N/A	78.73 (5.94)	78.80 (7.24)	N/A	N/A