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 1

Author(s): Sarojani S. Mohammed, Elizabeth Swanson, Greg Roberts, Sharon Vaughn,
Janette K. Klingner & Alison Gould Boardman.

Abstract Body
Limit 5 pages single spaced.

Background/context:
Description of prior research, its intellectual context and its policy 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. CSR was designed to improve the reading comprehension of struggling readers, including English language learners and students with reading disabilities (Klingner et al., 2001). 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.

The following studies provide initial support for the use of CSR.

Study 1. In this experimental study, students were assigned randomly to one of two treatment conditions; there was no comparison condition that involved "business as usual" or school practice. Twenty-six 8th grade Hispanic English language learners with learning disabilities participated in the study (Klingner & Vaughn, 1996). Students made significant improvements in reading comprehension on the *Gates MacGinitie Reading Test*: $F(1, 22) = 77.14$, $p = .0001$.

Study 2. This quasi-experimental study (Klingner, Vaughn, & Schumm, 1998) was conducted in diverse, inclusive 4th grade classrooms that included struggling readers. The intervention was implemented by the research team rather than the classroom teachers. Students were taught how to use CSR while reading social studies text. Control students received typical teacher-directed instruction in the same content. CSR students made statistically significant greater gains than students in the control condition on the *Gates MacGinitie Reading Test*: $F(1,138) = 10.68$, $p = .001$ ($ES = .44$) and demonstrated equal proficiency in their knowledge of the social studies content.

Studies 3 and 4. Case study approaches were used to examine the efficacy of CSR with English language learners. Results indicated that students demonstrated high levels of academic engagement and skillfully assisted each other with word meanings, main idea, and understanding text (Klingner & Vaughn, 2000). In Study 4, Bryant, Vaughn, Linan-Thompson, Ugel, and Hamff (2000) implemented CSR in an inclusive middle school program where gains for students with and without disabilities were demonstrated (Bryant et al., 2000).

Study 5. In this quasi-experimental study, researchers compared the findings for diverse students from 5 CSR and 5 comparison teachers from 5 schools (Klingner, Vaughn, Argüelles, Hughes, & Ahwee, 2004). Students in CSR classrooms improved significantly in reading comprehension when compared with comparison students. On the Gates-MacGinitie, posttest differences were statistically significant in favor of the CSR classes (with pretest scores used as the covariate), $F(1, 208) = 6.39$, $p = .01$, $\eta^2 = .03$, $d = .19$. Effect sizes were: .25 for high/average achieving students, .51 for struggling readers, and .38 for students with learning disabilities in reading. Teacher case studies revealed that, with the exception of one teacher, students' comprehension gains were associated with the quality and quantity of CSR implementation.

Study 6. Most recently, a computer adapted version of CSR was implemented with middle school students with reading difficulties (Kim et al., 2006). The experimental group outperformed the comparison group on the Woodcock Johnson Passage Comprehension test, $F(1, 31) = 4.75$, $p < .05$, with an effect size of .50.

In summary, over the last ten years we have studied the effects of CSR either separately or as part of a package of reading interventions in culturally and linguistically diverse elementary and middle school classrooms that included struggling readers. We have found that CSR is a feasible, effective model that can be implemented in authentic education settings. The current study is the first true randomized control trial investigating the efficacy of CSR.

Purpose / objective / research question / focus of study:
Description of what the research focused on and why.

This project is a multi-site, multi-year study designed to test the efficacy of a fully developed intervention, Collaborative Strategic Reading (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?*

Setting:
Description of where the research took place.

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 within English/Language Arts classes; (b) the school offered separate reading intervention classes for 7th and/or 8th graders; and (c) socio-economic status of students attending the school were low to moderate.

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

Students in this study were 7th and 8th graders enrolled in English/Language Arts and/or reading intervention classes in 2 schools in Texas and 4 schools in Colorado. Demographic information is provided in Table 1 (Appendix B).

Seventeen teachers (7 in Texas, 10 in Colorado) participated in the study, for a total of 61 classes, of which 34 were randomly assigned to the CSR condition (treatment), and 27 assigned to "business as usual" (comparison). Demographic information about the teachers is provided in Table 2 (Appendix B).

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.

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 and teaches students to identify the most important information in an entire passage. Wrap-up includes two components. First, students generate and answer their own questions about what they have read, and second, students review what they have learned by summarizing the key ideas presented in the text.

Initially, the teacher presents the strategies to the whole class using explicit instruction, modeling, and teacher think-alouds. The teacher provides guided practice with multiple opportunities for feedback. 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 everyone has the opportunity to contribute to the group's learning from and understanding of text.

CSR includes various materials that enhance the effectiveness and feasibility of the model: (a) learning logs that enable students to keep track of learning and provide a record that teachers can check and students can use for review; (b) cue cards for every role, with prompts that remind students what to do; and (c) clunk cards, with a list of "fix up" strategies for figuring out a clunk word or concept.

Once students are able to apply CSR's reading comprehension strategies and help each other while working in their cooperative groups, the teacher's role is to circulate among groups and provide ongoing assistance. Teachers help by actively listening to students' discussions and providing feedback, clarifying difficult words, modeling strategy usage, encouraging students to participate, and providing positive reinforcement. The focus of students' work should be on learning the material and helping their classmates learn it as well, not merely going through the steps of a given strategy. Also, teachers should regularly conduct whole-class previews and wrap ups to introduce new material, facilitate follow up activities, and reinforce learning.

Teachers were asked to deliver CSR lessons 2-3 times per week over a 26-week period between October and May. 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.

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. The 61 classes were randomly assigned within teacher to either treatment or comparison condition. In other words,

students were randomly assigned to class and then classes were randomly assigned within teacher. 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 were included in all pre- and post-testing. We assessed a range of word identification, fluency, and comprehension skills at pretest and immediate posttest.

Teachers were considered novice implementers of CSR. 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 continuing to use their usual instructional practices with the comparison classes. Because the same teacher provided intervention 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. We also collected data on student characteristics (e.g., language and special education status, age, gender) 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).

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

All students in all classes 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).

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 main findings with specific details.

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 at pretest 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). The model-estimated (level-2 latent) standard score average on the Gates-MacGinitie was 97.04 for treatment classes and 95.87 for the comparison conditions; and on AIMSweb were 93.42 and 92.53 for the comparison and treatment conditions, respectively. The estimated posttest (raw) scores on the TOSRE for comparison and treatment were 28.75 and 29.27. Descriptive statistics are presented in Table 3 and 4 (Appendix B).

To evaluate the statistical significance of these group differences, the fit of each of the three unconditional models was compared to its respective conditional model with posttest estimates constrained as equal across conditions. For Gates-MacGinitie, $\Delta\chi^2$ was 9.91 (p < .01), suggesting that participants in CSR outperformed non-participants when the effects of clustering and pre-treatment differences were explicitly modeled. Differences on AIMSweb ($\Delta\chi^2$=1.13, Δdf = 1) and TOSRE ($\Delta\chi^2$=.41, Δdf = 1) were not statistically significant. Results for the sub-sample of struggling readers were similar to those for the total sample. The model-derived posttest score on the Gates-MacGinitie was 87.66 for CSR participants, about 3.14 standard score points greater than initially struggling students in the comparison. Though not statistically significant (p = .066), the difference represents about 21% of the 15-point standard deviation used by the Gates-MacGinitie, an effect with considerable practical significance (Rossi, Lipsey, & Freeman, 2004).

Conclusions:
Description of conclusions and recommendations based on findings and overall study.

Based on the current analysis, we conclude that there is a small, significant main effect of CSR on reading comprehension as measured by the Gates-MacGinitie assessment; but that this effect is not statistically significant for struggling readers. We believe that the findings from this study suggest that CSR is a feasible and effective practice that can be readily integrated into reading and language arts instruction with positive impact. We are encouraged about the potential effectiveness of this practice because the positive findings from this efficacy study resulted from treatment implementation conditions that are readily replicable.

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.

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. Demographic information for student participants

Category	Treatment Group	Comparison Group
Male	54%	53%
Female	46%	47%
White	47%	38%
Hispanic	48%	55%
African American	4%	4%
Asian	1%	3%
Average Age	13.9 years	13.7 years
Average attendance	6.4 days absent	6.5 days absent
Economically Disadvantaged	53%	53%
Special Education Status	15%	8%
Limited English Proficient	3%	4%

Table 2. Demographic information for teacher participants

Category	Texas (N=7)	Colorado (N=10)	Total
Male	0	3	3
Female	7	7	14
White	7	9	16
Asian/Pacific Islander	0	1	1
Average Education	16 years	17.6 years	16.9 years
Degrees Earned	7 Bachelor's	2 Bachelor's, 8 Master's	9 Bachelor's, 8 Master's
Experience	8.9 years	11 years	10.12 years
Specializations	7 Eng/Lang Arts, 4 secondary, 4 reading, 2 ESL, 1 elementary, 1 Special Education	8 Eng/Lang Arts[a], 6 elementary, 3 secondary, 2 reading, 1 Special Education, 1 speech	15 Eng/Lang Arts, 7 secondary, 7 elementary, 6 reading, 2 ESL, 2 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

Measure	Pretest Mean (s.d.)		Posttest Mean (s.d.)	
	Comparison	Treatment	Comparison	Treatment
Gates-MacGinitie	95.68 (13.4)	96.35 (13.7)	95.48 (13.4)	97.13 (13.6)
AIMSweb	92.64 (12.1)	91.91 (10.8)	93.46 (11.0)	92.92 (11.0)
Student Engagement Index	3.18 (0.35)	3.12 (0.36)	3.05 (0.43)	3.01 (0.41)
Meta-comprehension Strategy Index	9.80 (3.56)	10.15 (3.69)	10.01 (3.62)	10.48 (3.91)
Test of Sentence Reading Efficiency (TOSRE)	23.49 (7.36)	24.67 (7.17)	29.03 (7.68)	29.22 (7.66)
TOWRE - SW	90.17 (7.40)	90.59 (8.38)	N/A	N/A
TOWRE - PD	91.53 (11.3)	91.20 (10.85)	N/A	N/A

Table 4. Descriptive statistics on all measures for the struggling readers sample

Measure	Pretest Mean (s.d.)		Posttest Mean (s.d.)	
	Comparison	Treatment	Comparison	Treatment
Gates-MacGinitie	84.37 (10.0)	85.62 (11.2)	84.25 (9.08)	87.74 (9.95)
AIMSweb	85.81 (8.75)	85.13 (8.01)	86.30 (7.97)	86.60 (7.37)
Student Engagement Index	3.12 (0.38)	3.03 (0.32)	2.95 (0.59)	2.95 (0.45)
Meta-comprehension Strategy Index	9.30 (3.56)	9.06 (3.58)	9.63 (3.51)	9.73 (4.00)
Test of Sentence Reading Efficiency (TOSRE)	18.52 (4.60)	20.30 (5.89)	23.87 (6.34)	24.83 (5.76)
TOWRE - SW	84.48 (5.46)	83.76 (4.62)	N/A	N/A
TOWRE - PD	81.57 (5.52)	82.55 (5.57)	N/A	N/A