

1 **A Randomized Evaluation of the Success for All Middle School Reading Program**

Anne Chamberlain

*Success for All Foundation
Baltimore, Maryland*

Cecelia Daniels

*Success for All Foundation
Baltimore, Maryland*

Nancy A. Madden

*Success for All Foundation
Baltimore, Maryland*

Robert E. Slavin

*Johns Hopkins University
Baltimore, Maryland*

ABSTRACT

This article describes a randomized evaluation of The Reading Edge, a reading program for middle school students. The Reading Edge was designed to integrate findings of research on cooperative learning and metacognitive reading strategies into a replicable reading instructional package that could be implemented effectively in Title I middle schools. In this study, 405 sixth graders in two high-poverty, rural middle schools previously unfamiliar with the program were randomly assigned to participate in The Reading Edge or to continue with their existing reading programs. After one year of instruction, observations of classroom use of metacognitive strategies, cooperative learning, goal setting/feedback, and classroom management, showed moderate levels of implementation in Reading Edge classes but little use of metacognitive strategies, cooperative learning, or goal setting/feedback in control classes. Statistically significant differences in student scores on the Vocabulary subscale of the Gates-MacGinitie Reading Test, and marginally significant scores on the Total Achievement score, provide support for the basic reading model, but larger and longer studies are needed to establish the full effects of this approach.

INTRODUCTION

The reading performance of students in many high-poverty middle schools is very low. On the National Assessment of Educational Progress (NCES, 2005), 43% of eighth graders who qualified for free lunch scored below the basic reading level, compared to only 19% of their non-qualifying peers. This situation is not new, but by any standard, the reading scores of disadvantaged and minority students at the end of the middle grades is unacceptable (see Alvermann, 2001; Cooney, 1998; Jackson & Davis, 2000). Clearly, there is a need for well-evaluated programs capable of significantly accelerating the achievement of young adolescents.

While the poor reading skills of students in high-poverty middle schools have long been recognized as a problem, there are very few replicable interventions available to improve the reading achievement of students in these grades, and fewer still that have evidence of effectiveness from experimental-control comparisons. No comprehensive reading program for middle grades students currently in use has ever been evaluated in a randomized experiment, and very few have even been evaluated in matched quasi-experiments. The two key exceptions, programs that have at least been evaluated in matched experiments, are reciprocal teaching (Palincsar & Brown, 1984; Rosenshine & Meister, 1994) and transactional strategy instruction (Brown et al., 1993). Unfortunately, it is rare to encounter either of these approaches in middle school reading instruction. It is rarer still to see them in a high-poverty middle school.

The research reported here presents a randomized evaluation of The Reading Edge, a literacy program for middle school students used (as of the 2006-07 school year) in 144 schools nationwide. Although the Success for All Middle School, in which The Reading Edge is central, produced positive effects on the reading achievement of middle school students in a matched experiment (Slavin, Daniels, & Madden, 2005), a randomized experiment of The Reading Edge is an important progression in the research. Since many schools adopt The Reading Edge as a “stand-alone” program, it is important to understand its impact as such. A randomized methodology ensures the highest level of scientific rigor by ruling out selection bias (Mosteller & Boruch, 2002). Further, randomized designs fulfill current standards for rigorous evaluations, such as those applied by the federal What Works Clearinghouse.

MIDDLE SCHOOL REFORM AND THE SUCCESS FOR ALL MIDDLE SCHOOL

The middle school movement has dominated the national reform conversation regarding the education of young adolescents.

This movement has focused on such structural elements as grade span (6-8 instead of 7-9), interdisciplinary teaming, site-based decision making, flexible scheduling, and reducing ability grouping. These were the focus of *Turning Points* (Carnegie Council on Adolescent Development, 1989), the influential touchstone of the middle school movement. These structural changes are probably good practice, but 18 years after *Turning Points*, evidence that they make a systematic difference in student achievement remains elusive (see Balfanz et al., 2002). The Carnegie Corporation itself, in *Turning Points 2000* (Jackson & Davis, 2000), recognized the problem: Structural changes create the potential for change in student outcomes, but genuine change requires reforming the classroom instruction that takes place in the restructured middle school (see Lipsitz et al., 1997; Newmann, 1993; Moore et al., 1999). Today, any reform in classroom instruction takes place in the context of standards-based accountability systems that are linked to serious consequences for both educators and students. Under No Child Left Behind (NCLB), middle school test scores are just as important as elementary ones, yet the teacher shortage falls most heavily on high-poverty middle schools. This presents an additional challenge to instructional reform, as young adolescents in disadvantaged areas are less likely to have teachers certified in their content area (McEwin & Dickinson, 1997).

In response to the widely perceived need for more focus on curriculum and instruction, especially related to literacy, middle school reformers are increasingly providing teachers with specific materials, professional development, and assistance in order to implement research-based instruction. Early successes have been achieved in comprehensive middle school reform models, notably the Talent Development Middle School (Balfanz & MacIver, 2000) and the Success for All Middle School (SFAMS) (Slavin, Daniels, & Madden, 2005).

The Reading Edge evaluation builds upon studies of SFAMS and its predecessor, Cooperative Integrated Reading and Composition (CIRC), as well as numerous studies of the Success for All elementary reading program. A matched experimental study of SFAMS took place in 7 experimental and 7 control schools in 6 states (Slavin, Daniels, & Madden, 2005). On state assessments, six of the seven SFA Middle Schools gained more than their matched comparison schools in reading. Combining data across schools, the SFA Middle School students' scores were significantly higher than those of their matched controls. Studies of CIRC in upper elementary and middle schools also produced significant positive reading outcomes (Stevens, Madden, Slavin, & Farnish, 1987; Stevens & Durkin, 1992; Stevens, Slavin, & Farnish, 1991; Stevens & Slavin, 1995; Calderón, Hertz-Lazarowitz, & Slavin, 1998). A national randomized evaluation of the SFA elementary model (Borman et al., in press) also found positive effects.

THE READING EDGE – PROGRAM DESCRIPTION

The following description of The Reading Edge explains the structure and rationale for each component of the program. Because of the nature of a randomized trial, there were two key differences between the implementation of the Reading Edge in most schools and the implementation of the program in this study's two research sites.

First, in most schools using The Reading Edge, a full-time, school-based facilitator partners with the school's instructional leaders, coaches teachers, and communicates regularly with SFAF coaches. Secondly, Reading Edge schools typically involve a large percentage of the teaching staff in the program. This allows schools to schedule reading during one to three common time periods each day. Peer support for teachers has the potential to increase when a large cadre of teachers works together. The research sites did not have school facilitators, the number of participating teachers was limited, and schools did not have common reading periods for regrouping.

The Reading Edge is a comprehensive literacy program for all middle grades students, from struggling to advanced, with the goal of preparing them to be strategic, independent, and motivated readers and learners. The Reading Edge combines effective instructional practices, a coherent curriculum, and frequent informal assessment and feedback to students, with extensive professional development for teachers and instructional leaders. This enables middle school teachers to provide intensive reading instruction that emphasizes four key areas: metacognitive strategy use, cooperative learning, goal-setting and feedback, and classroom management.

Since reading levels vary widely for this age group, students are assessed and grouped for instruction in classes at their reading level. All reading classes meet during the same 60 minute period each day. Students are re-assessed quarterly to determine their current level of performance. These assessment results and other indicators allow teachers to move students immediately to a more challenging group. This structure allows for both differentiated, targeted instruction, and rapid acceleration. It allows students who are struggling with decoding and phonics to work through well-structured texts with a high proportion of phonetically decodable words, while in other classes, students use novels or expository texts appropriate to their reading level to improve fluency and comprehension. As soon as students are ready, they are moved to the next level. Extensive training and ongoing coaching support enable teachers and facilitators to implement this aspect of the program efficiently and with integrity.

Every Reading Edge lesson follows a regular pattern, but the instruction itself varies according to the nature of the text and the challenges it presents. Program developers carefully choose readings that lead the students to relevant discoveries and practice opportunities.

Combining explicit instruction in metacognitive reading strategies with peer collaboration and regular assessment and feedback builds students' confidence as they try out their new tactics. Although lessons vary significantly in level of difficulty and purpose, all follow this structure: Setting the Stage, Active Instruction, Teamwork, and Time for Reflection.

The four key areas of metacognitive strategy use, cooperative learning, goal-setting and feedback, and classroom management techniques that engage students, are embedded in the everyday classroom routines that teachers use for instruction, and that students use to become better readers (Slavin, Daniels, & Madden, 2005). The following sections describe how each of the key program areas identified above are manifested in The Reading Edge.

METACOGNITION IN THE READING EDGE

Numerous studies show that metacognition plays an important role in both learning to read and in learning new content through reading (Armbruster et al, 1987; Gertz, 1994; Langer & Neal, 1987). Research suggests that students who can assess their own knowledge and interest in a topic and choose and apply effective strategies to understand unfamiliar text on that topic can control their learning environment and successfully make meaning from what they read (Tei & Stewart, 1985).

Whether or not students use metacognitive strategies, and how they use them, seems directly related to age and experience. By middle and high school, successful readers will have developed enough background knowledge, vocabulary, and experience with both narrative and expository text to take the next step in their reading development. Students can move on to develop the thinking skills necessary to grapple with the difficult text presented in content-area classes, and to acquire the study skills needed to store, retrieve, and use information after they achieve comprehension (Gertz, 1994; Langer & Neal, 1987).

However, it is not the case that all students enter middle school reading fluently or widely, nor do they leave middle school ready for the challenges of high school. If early adolescence is the time during which students lay the groundwork for more mature ways of thinking and rigorous coursework in high school, then it is essential that students do more than just "catch up". Students need the opportunity to set goals, experiment with new skills and strategies, and learn that persistence leads to success. All students, including those who are reading on or above grade level by middle school, can benefit from instruction in recognizing various text structures, practicing strategies to help them make meaning from text, and experiences that provide them a greater awareness of their own characteristics as learners (Muth, 1987).

Early in every Reading Edge lesson, students are presented with “The Big Question”, a provocative question that students ponder and discuss over the course of the day’s activities. With stems such as “Have you ever...” or “How would you handle this character’s challenge...” or “Based on what you know about...”, students must draw upon their own experiences and beliefs as well as details from the text they are reading in order to formulate their answer. This question often leads to student-generated questions of the same kind, and enriches team and partner discussions about the text. It also requires students to paraphrase or summarize what they have read, and encourages students to monitor their comprehension as they read.

Also in the beginning of every lesson, teachers complete a “Building Background” segment in which they activate prior knowledge with “KWL” charts that ask students to summarize what they know about the topic (K), what they want to know (W), and then what they learned (L). Teachers make connections to student interests, and hold conversations in which they discuss vocabulary important to the understanding of the text. They preview the text, and discuss text features that prompt students to make predictions about the main idea, topic, or theme, depending on the nature of the text.

During the “Active Instruction” portion of the lesson, the teacher engages students in targeted instruction on how to use a particular strategy or skill. For example, the teacher will read aloud a passage from the text and stop and “think aloud” about something significant, perhaps an example of foreshadowing, and how it influences her thinking as she reads. As she does this, she is breaking down a larger strategy into smaller steps so students can understand the otherwise invisible tools that good readers use. Now that this strategy has been modeled and discussed, students practice using it as they read, and later reflect upon whether or not it helped them.

During Teamwork, students read some text silently and some aloud with a partner. They stop regularly to paraphrase what they just read, to share insights, and to clarify understanding. After reading, each team of 4 or 5 students uses another routine to discuss what they have read, and “Discuss and Defend” their answers to prepared “Team Talk Questions”. Students have the opportunity to clarify their thinking, try out new ideas, use new vocabulary, and help someone else understand the text, all in the safety of a small collaborative group. During Teamwork time, the teacher circulates around the room to check for understanding, prompts and reinforces positive behavior, and conducts quick, informal conferences called “One-to-One’s” with partners or teams. During these interactions, the teacher can informally assess the students’ grasp of the targeted strategy or skill.

COOPERATIVE LEARNING IN THE READING EDGE

Cooperative learning refers to a wide variety of methods in which students work in small groups (usually, about four members) to help one another learn. Extensive experimental research on cooperative learning methods has found that these strategies generally increase student achievement if they incorporate group goals and individual accountability (Slavin, 1995; Slavin et al., 2003; Webb & Palincsar, 1996). Cooperative learning has been particularly effective and extensively researched in middle schools (Slavin et al., 2003). Cooperative learning methods in which students have regular opportunities to discuss ideas with partners, to help each other study, and to provide formative feedback to each other have been successfully evaluated in several randomized evaluations in elementary and middle schools (Slavin, 1977, 1978, 1979, 1995; Slavin & Karweit, 1984). Cooperative learning has also been successfully evaluated many times as a means of introducing and practicing metacognitive skills. For example, Meloth & Deering (1992, 1994) found that peers could help each other acquire metacognitive strategies. Dansereau (1988) and O'Donnell (1996, 2000) studied "cooperative scripts," in which students took turns summarizing and evaluating each other's summaries. Fantuzzo, Polite, & Grayson (1990) developed and evaluated reciprocal peer tutoring strategies to help students study complex material.

In The Reading Edge, students work in teacher-assigned, cooperative teams of four or five students. Each student is individually accountable for his or her own learning and there are no group grades. However, teams are rewarded when all members improve their performance compared to their past performance. Students set goals together, and with the support of practiced routines, they help each other learn new content, use, evaluate, and personalize strategies, and remain attentive to the task at hand. Teams also provide a safe environment for individuals to try out new ideas or admit when they are confused. Each team works to help each member achieve his or her "personal best" so there is always an equal opportunity for success – teams are not in competition with each other. Teams strive to improve upon past performance. Cooperative learning allows students to get to know one another as individuals as they work together toward common goals (Berkowitz & Bier, 2005).

STUDENT GOAL-SETTING AND FEEDBACK IN THE READING EDGE

Research demonstrates that students who perceive a classroom emphasis on meeting mastery goals rather than on ability and performance goals are more likely to take on challenging tasks, use

effective strategies, and have a positive attitude in class. These students are also more likely to believe that effort brings success (Ames & Archer, 1988). A study of goal-setting and math achievement for African American students transitioning to high school showed that focusing on mastery goals produced positive outcomes in motivation and higher grades (Gutman, 2006). A large body of research has shown the achievement benefits of specific strategies such as goal setting (Schunk & Schwartz, 1993) and self-monitoring by recording one's progress (Zimmerman, Bonner, & Kovach, 1996).

Built into The Reading Edge are several unique components that engage students in the learning process: goal-setting, peer support, and frequent assessment and feedback. These powerful motivators provide the impetus for increased metacognitive strategy use and student achievement.

Goal-Setting

The most important aspect of the first three weeks of reading instruction in The Reading Edge is to engage students in the learning process. During both the Getting Along Together unit and the first Reading Edge Foundation Unit, teachers engage students in straight talk about reading difficulties they may have had in the past, especially related to work in their content area classes. Students engage in activities that help them identify goals that are important to them, and to learn how to work collaboratively to reach those goals. During the first two weeks of school, students take an initial or baseline reading assessment. Students learn how they scored on the initial reading assessment and are placed in an appropriate reading class. All reading levels begin with a Foundation Unit designed to demonstrate how strategies students already use in everyday situations (making predictions, asking questions, etc.) can also be applied to learning problems, such as understanding one's science textbook. Connecting with students' prior knowledge and experience with problem-solving strategies launches the over-arching investigation into how to choose and apply multiple, effective strategies for success in reading.

Peer Support

Placing students in reading classes by instructional level gives students a safe place to work on the issues that are most difficult for them. By the middle grades, most students who are reading below grade level are painfully aware of their deficiencies, and most have learned how to hide them. Early in The Reading Edge, students learn that everyone has something to learn, including their teacher, and that in this reading class they will work with their teammates to fill in gaps

and build on their strengths. Not only is it acceptable to talk about what stumps them when they are reading, it is part of the everyday routine.

Frequent Assessment and Feedback

At the beginning of every cycle (six days of instruction ending with a curriculum-based assessment), students receive a Learning Guide. On it, students record reading goals for each lesson and points they earn for classwork, homework, and tests that make up their grade for the cycle. The Learning Guide is a self-assessment tool that helps the students identify strengths and weaknesses and set personal goals. It also helps students stay organized. Students keep track of their work as they complete it, and identify the work they still need to master. As the teacher visits each team during one-to-one conferences, she initials work that students have done well, and helps them identify next steps.

Teams can earn bonus points for demonstrating extra effort by participating in activities such as Book Club, in which students share their critiques of books they have read outside of school. In “Word Wise”, students share how they have used their vocabulary words in new contexts. Students also learn to use a Team Mastery Process to prepare for the end of cycle test. This gives students time to try out new study skills in a personally meaningful context.

At the end of each cycle, students compare their current total score to their previous work. Teams receive “improvement points” for each member who increases his or her score, and the team is recognized in a whole-class celebration. This is also a time for the whole class to review their reading goals for the cycle and discuss which strategies worked well for them, which ones did not, and why.

Every quarter, students take a standardized reading assessment. This information is added to the growing set of data about how students are performing. Students know that as soon as they show sufficient improvement in reading, they will move up to the next level, or even skip a level. Regular and frequent assessment and feedback routines include the student as an equal partner in the process. Students learn that hard work and persistence will pay off.

CLASSROOM MANAGEMENT IN THE READING EDGE

Research indicates that a rapid pace of instruction, consistent with high student comprehension, both maintains students’ attention and increases students’ achievement (e.g., Barr, 1987; Good, Grouws, & Ebmeier, 1983). In addition, classroom management methods based on cooperative learning have both immediate and lasting impact on students’ behavior and achievement. For example, Hawkins, Doueck, & Lishner (1988) used preventive classroom management methods emphasizing cooperative learning and interactive teaching with low-

achieving seventh graders. In comparison with control group students, those involved in the program were suspended and expelled less often, had better attitudes toward school, and had higher achievement. These effects were substantially maintained in a long-term follow-up assessment into high school (Hawkins, Guo, Hill, Battin-Pearson, & Abbott, 2001). Other longitudinal studies have also shown immediate and long-term positive effects of classroom management programs that emphasize cooperative learning and student engagement (O'Donnell et al., 1995; Freiberg, Connell, & Lorentz, 2001; Dolan et al., 1993). The Reading Edge teachers begin each school year with a conflict-resolution and academic problem-solving unit called "Getting Along Together" (GAT). In this unit, students learn how to use effective social, communication, and problem-solving skills to maintain positive social and academic relationships. These skills are then referenced and practiced in reading lessons. During Setting the Stage, the teacher and students discuss and set a cooperation goal for the day. Teams can earn bonus points and recognition for meeting goals such as active listening, helping and encouraging each other, or 100% team participation.

During Active Instruction, teachers use techniques for keeping the lesson moving at a brisk pace by posing important questions to the entire class and randomly calling on students to answer. Teachers use routines such as "Think-Pair-Share" and "Numbered Heads" to allow students time to examine their thinking and then try it out in a brief team discussion. Students learn that any team member could be called upon to answer for their team, so they must help each other prepare to answer every question (Kagan, 2001). For quick checks for understanding, teachers use whole-class responses such as "Thumbs-Up/Thumbs Down" to keep the pace of instruction brisk, and every child involved (Emmer, Evertson, & Worsham, 2003).

THE READING EDGE EVALUATION

In 2005, the Success for All Foundation received a grant from the William T. Grant Foundation to study the mechanisms and impact of The Reading Edge. This was a critical step for the program. Although a study of the Success for All Middle School in which The Reading Edge is embedded had shown significant positive results (Slavin et al., 2005), most schools adopt The Reading Edge as a stand-alone intervention, so it is important to evaluate this key component on its own. The following sections describe a mixed-methods evaluation of The Reading Edge.

METHODS

Sample

The Reading Edge program was implemented in two majority-White, high-poverty, rural middle schools, one in Ronceverte, West Virginia and one in Interlachen, Florida. The West Virginia school served a student population in which 50% of students qualified for free lunch, and in which 95% of students were White, 4% African-American, and 1% Native. At the Florida school, 69% of students were eligible for free lunch, and the student population was 75% White, 8% African-American, 13% Hispanic, and 1% Asian or Native.

Design

The study used a mixed methods approach, including a randomized experimental design complemented by intensive classroom observations. This combination of methods allowed researchers to understand achievement within specific school contexts.

Schools were recruited in spring, 2005. In exchange for their participation, they received all Reading Edge materials and training at no cost. Those teachers and students assigned to the treatment group received materials and support immediately, while those in the control group will receive the same materials and training in fall, 2008 (a delayed treatment design). In the study year, control groups were provided with books or assessment materials related to programs they were already using. Prior to the 2005-06 school year, both schools allowed researchers to randomly assign 6th grade teachers to treatment (The Reading Edge) or control conditions. Incoming 6th graders were stratified by state reading assessment levels, then gender and ethnicity. Within strata, students were randomly assigned to treatment or control conditions. Students were pretested within the first four weeks of school, and then posttested in spring, 2006.

Control Condition

Students randomly assigned to the control group received daily reading instruction using whatever methods and materials they had used previously. Both schools grouped control students into classes based on 5th grade state reading scores. At the WV school, students with low reading skills used Scholastic Read 180 (a computerized reading program), and the Scholastic Read Aloud Anthology. Larry Bell's 'Twelve Powerful Words' were used in all reading classes and levels. Students at more proficient reading levels participated in a teacher-designed 'Literacy Circle' where, in groups of 5-6, they read and discussed novels. Scott Foresman reading materials were used for

non-fiction content. In the Florida school, low readers used Corrective Reading, while their peers used Scholastic Read XL.

Measures

The Scholastic Reading Inventory (SRI), a widely used measure of reading, was used as an achievement pretest in fall, 2005. Since Reading Edge classrooms then administer the SRI quarterly, it was necessary to select a different posttest, to avoid a practice effect. The Gates-MacGinitie Reading Test (MacGinitie, MacGinitie, Katherine, & Dreyer, 2000) was used as a posttest. Its total score can be broken into two subscales, comprehension and vocabulary. Two-day observations were conducted twice per year at each site, using a structured observation protocol based on the CLASS instrument (LaParo & Pianta, 2003). The observation tool noted and characterized the 4 instructional elements identified as key by program developers: metacognitive strategy use, cooperative learning, goal setting and feedback, and classroom management/student engagement. The resulting instrument was designed to register many ways of implementing the four elements, so that their use in control classrooms would not be overlooked.

RESULTS

Reading Achievement

Initial analyses showed that random assignment produced equivalent groups. There were no differences at pretest on the SRI ($F(1, 403) < 1$, n.s.), and chi square analyses showed no experimental-control differences in race, free lunch eligibility, or special education status. Posttest Gates achievement data were analyzed using analyses of covariance (ANCOVA), with SRI pretest scores as covariates. Table 1 summarizes the findings.

As shown in Table 1, there were significant experimental-control differences at posttest on the Vocabulary subscale of the Gates ($F(1, 402)=3.95$, $p=.048$), controlling for pretests. The Comprehension subscale did not show significant differences ($F(1, 402)=1.99$, $p=.16$), and the Total score indicated a difference in favor of the Reading Edge with marginal statistical significance $F(1, 402)=3.55$, $p=.06$. Effect sizes (differences in adjusted means divided by the unadjusted control group standard deviations) were modest: Gates Total ($ES=+0.14$), Gates Comprehension ($ES=+0.11$), and Gates Vocabulary ($ES=+0.14$). A difference of $+0.14$ would be equivalent to increasing a test score from the 50th to the 56th percentile.

Table 1
Achievement Outcomes and ANCOVA Results

	N	Mean	Standard Deviation	Adjusted Mean	Effect Size
SRI-Pretest					
Reading Edge	203	885.3	(171.7)		
Control	202	880.9	(173.8)		
Gates Total					
Reading Edge	203	58.07	(16.15)	57.93	+0.14 ^a
Control	202	55.61	(16.10)	55.76	
Gates					
Comprehension					
Reading Edge	203	29.95	(9.14)	29.87	+0.11
Control	202	28.83	(8.92)	28.90	
Gates Vocabulary					
Reading Edge	203	28.12	(7.92)	28.05	+0.14*
Control	202	26.78	(8.35)	26.85	

Note. ^a $p = .06$; * $p = .05$

Observation Results

Several classes from each treatment, representing a variety of reading levels, were each observed for a full class period (approximately 50 minutes), twice during the school year. Spring observations revealed differences between treatment and control classrooms in terms of the quantity and quality of the four key areas observed. In addition to a checklist and open-ended notes, each classroom was rated in each of these areas, on a Likert-type scale of 1-7, 1 being 'low', based on the CLASS instrument.

Metacognitive Strategy Use

The biggest differences between conditions appeared in the Metacognitive Strategy Use category (experimental $M=2.38$, control $M=1.00$). This is not surprising, given that daily instruction on metacognitive strategy use is built into The Reading Edge curriculum. However, although the Metacognitive mean classroom rating was greater for Reading Edge, the score was still low. Observer notes indicate that implementation of this key program factor remained at a mechanical stage. Specifically, although some direct instruction and teacher modeling of metacognitive strategy use was observed in treatment classrooms, students in these classrooms rarely engaged in discussions about them. Any discussion that surfaced was centered on

what strategy was used, rather than why or how. Strategies were presented as isolated tools for particular problems, rather than as replicable strategies. This was most evident when teachers failed to take advantage of ‘teachable moments’; unscripted opportunities to illustrate, suggest, or discuss a strategy use. For example, during silent reading in one class, the teacher noticed a student writing down words on a scrap of paper. At the end of silent reading, he was praised, and the teacher said to the class, “Josh, would you tell them what you did while you were reading?” The teacher then answered her own question, telling the class that Josh had written down words he didn’t understand. She asked Josh to read those words aloud, and praised him before the class, for “...doing a good job.” There was no discussion about when these words would be clarified, why this is a helpful practice, or why Josh chose this particular strategy.

One indication that students had absorbed at least the vocabulary of metacognitive strategy use was seen in a different Reading Edge classroom, as a teacher attempted to read a student’s handwriting. “Does this say ‘entertain’ or ‘entrance’?” The student responded with mock patience, “Well, you could read the rest of the sentence to find out.”

In control classrooms, some strategies, such as graphic organizers, were used; however, the presentations were ‘static’. In one class, for example, students were each provided a worksheet with a blank graphic organizer on it, and specific instructions for what the contents of each circle should be. The graphic organizer exercise was just that—an exercise. The only reference to using this skill beyond the task at hand was to remind students that graphic organizers would be on the state test.

Cooperative Learning

Cooperative learning was also used differently in treatment and control groups. Experimental classrooms used cooperative learning somewhat more effectively (experimental $M=3.50$, control $M=2.50$). Although one control class was divided into small groups, virtually all of the work was individual seatwork. There were no group products or processes. In contrast, groups in the treatment classes had student-generated team names, and earned team rewards both for team products and the quality of their teamwork. These processes included Team Talk Questions, idea generation, and consensus building. Perhaps the best example of this was seen in an activity during which teams were asked to categorize a list of vocabulary words from a novel they’d recently completed. Students decided how many categories to have, and how to classify their words. As the time for this activity drew to a close, the teacher exclaimed, “OK! It’s decision time! Consensus time!” then, “First of all, let’s talk about process. How did

your group work together? How did you split this task? You have one minute to talk about it [as a team], and I need a spokesman for each group.” “I want you to discuss how you talked about [the category]. Were there some words that some people knew in your group and others didn’t?”

Goal Setting and Feedback

Goal setting and feedback varied somewhat across conditions (experimental $M=3.30$, control $M=2.25$), and also varied within the treatment condition. In control classrooms, charts of student progress were often posted, though never referred to during any observation visit. Goal setting usually referred to a task-oriented ‘Day’s Objective’ or to progress on the state test. Informal feedback was frequent, behavior-oriented, both negative and positive, and usually delivered to the class as a whole, whether directed at the entire class or an individual.

There was more strategic use of goal setting and feedback in Reading Edge classrooms. Task-oriented ‘Day’s Objectives’ were posted here, too, and referred to just as infrequently as they were in control classes. However, goals were also evident and tied to feedback in the consistent use of ‘team score sheets’, which provided team-level goals for each week, and were maintained by the teams themselves. Similarly, one teacher in particular made thoughtful use of the ‘learning guide’, which students checked for assignment criteria with which they self-monitored their individual progress, as a matter of course. Informal feedback on behavior in this particular classroom was also consistently positive, with smaller group and one-on-one interaction, as teachers facilitated group activities. Research notes indicate that while providing feedback including rewards is a familiar routine with teachers, tying that feedback to attainable, short and intermediate goals, is less so. Reading Edge teachers seemed to be discovering a greater potential for goal setting, but improving state test scores and avoiding retention were the most frequently mentioned “goals” among all teachers.

Classroom Management/Student Engagement

On the classroom management factor, treatment and control conditions shared the same means ($M=3.75$). Although there was the same degree of classroom order, management styles varied markedly, particularly at the site where the control teacher used traditional methods for ‘keeping order’ (desks in rows, silence during activities, and calling only on raised hands). The other control teacher seemed to have developed a strong bond with her students, eliciting such respect from them that little active management was necessary. Reading Edge

teachers, to varying degrees, used fast pacing, “OK! We’ve got a two-minute idea bank!”, and verbal signals to keep students on task. In one classroom, when cooperative ‘buzz’ spilled over into off-task behavior, the teacher sang fast and loud, to the tune of ‘shave and a haircut, two bits,’ “Da na-na na na,” to which the students replied with equal enthusiasm, “NA NA!” The room fell silent, then an acceptable buzz returned. Accountability measures (such as team score sheets) were also used to keep students on task in Reading Edge classrooms. At both schools, students were generally attentive and respectful, so that they did not particularly ‘test’ the efficacy of the different management approaches.

DISCUSSION

While effect sizes were modest, the findings of this randomized experiment show promise for the Reading Edge program. These findings are particularly encouraging in that they were detected *despite* a within-school randomized study—emergence of the four key implementation areas was inhibited due to the isolation of the treatment teachers as a result of a within-school randomized design. That is, these teachers lacked the support of other implementors or a full-time facilitator, as would be standard in a typical implementation. Additionally, the within-school randomization divided students in such a way that prohibited regrouping during common reading periods. Despite these implementation constraints, observation data suggest that while implementation of four key, research-based components of reading instruction was generally mechanical in Reading Edge classrooms, it differed substantively from that in the control condition.

A second cohort of sixth-graders, entering these schools in fall, 2006, has been randomly assigned, and researchers will have the opportunity to follow the progress of these students as they are taught by teachers now more experienced and confident. Clearly, more research is needed to better understand program impact and to increase the effectiveness of the intervention, but much can be learned from the process and results of this study.

Future research on The Reading Edge should address many issues beyond the basic findings reported here. First, research with a much larger sample of schools could allow for random assignment of schools, rather than students within schools. This would create conditions more like those in practice, where The Reading Edge is invariably used as a schoolwide intervention. A larger sample would also allow for determination of correlations between teachers’ quality and completeness of implementation of each program element and students’ achievement growth, as well as other mixed-method research to explain how and under what conditions the program produces positive effects.

Future studies should also continue the interventions over a longer time period, both to determine long-term impacts and to assess program outcomes with teachers who are experienced with the interventions, beyond the turbulent early stages of a new program. Multi-year studies of the Success for All elementary program (e.g., Madden et al., 1993; Borman et al., in press) typically find much better outcomes with each successive year of program implementation (see Slavin & Madden, 2001, 2006). The same would probably be true of The Reading Edge.

Schools serving disadvantaged young adolescents need replicable, reliable interventions capable of helping students become capable and strategic readers. Larger impacts may be seen in schoolwide implementations, in implementations with experienced teachers, and in longer implementations, and much work remains to be done to continuously improve the model itself, but the research reported here gives hope that these efforts will be worthwhile.

AUTHORS' NOTE

This research was funded by the W.T. Grant Foundation. However, any opinions expressed are those of the authors and do not represent the positions or policies of the W.T. Grant Foundation.

REFERENCES

- Alvermann, D.E. (2001). *Effective literacy instruction for adolescents*. Chicago: National Reading Conference.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267.
- Armbruster, B.B., Anderson, T.H., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly, 20*, 93-115.
- Balfanz, R., & Maciver, D.J. (2000). Transforming high-poverty urban middle schools into strong learning institutions: Lessons from the first 5 years of the Talent Development Middle School. *Journal of Education for Students Placed at Risk, 5* (1 & 2), 137-158.
- Balfanz, R., Ruby, A., & MacIver, D. (2002). Essential components and next steps for comprehensive whole-school reform in high poverty middle schools. In S. Stringfield and D. Land (Eds.), *Educating at-risk students*. Chicago: National Society for the Study of Education.
- Barr, R. (1987). Content coverage. In M. J. Dunkin (Ed.), *International encyclopedia of teaching and teacher education*. New York: Pergamon.

- Berkowitz, M., Bier, M.C. (2005). *What works in character education: A research-driven guide for educators*. Washington, D.C.: Character Education Partnership.
- Borman, G., Slavin, R.E., Cheung, A., Chamberlain, A., Madden, N.A., & Chambers, B. (in press). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal*.
- Brown, A.L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J. C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations*. New York, NY: Cambridge University Press.
- Calderón, M., Hertz-Lazarowitz, R., & Slavin, R.E. (1998). Effects of Bilingual Cooperative Integrated Reading and Composition on students making the transition from Spanish to English reading. *Elementary School Journal*, 99, (2), 153-165.
- Carnegie Council on Adolescent Development (1989). *Turning points: Preparing American youth for the 21st Century*. New York: Carnegie Corporation.
- Cooney, S. (1998). *Education's weak link: Student performance in the middle grades*. Atlanta: Southern Regional Education Board.
- Dansereau, D.F. (1988). Cooperative learning strategies. In C.E. Weinstein, E.T. Goetz, & P.A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 103-120). Orlando, FL: Academic Press.
- Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolff, J., Turkkan, J. S., Ford, C., & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology*, 4, 317-345.
- Emmer, E., Evertson, C., & Worsham, M. (2003). *Classroom management for secondary teachers (6th ed.)*. Boston, MA: Allyn & Bacon.
- Fantuzzo, J.W., Polite, K., & Grayson, N. (1990). An evaluation of reciprocal peer tutoring across elementary school settings. *Journal of School Psychology*, 28, 309-323.
- Freiberg, H. J., Connell, M. L., & Lorentz, J. (2001). Effects of Consistency Management on student mathematics achievement on seven Chapter I elementary schools. *Journal of Education for Students Placed at Risk*, 6 (3), 249-270.
- Gutman, L. M. (2006). How student and parent goal orientations and classroom goal structures influence the math achievement of African Americans during the high school transition. *Contemporary Educational Psychology*, 31 (1), 43-63.

- Gertz, E. A. (1994). *Enhancing motivation and reading achievement: intervention strategies for the underachieving middle school student*. Ed.D. Practicum, Nova Southeastern University.
- Good, T., Grouws, D., & Ebmeir, H. (1983). *Active mathematics teaching*. New York: Longman.
- Hawkins, J.D., Doueck, H.J., & Lishner, D.M. (1988). Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. *American Educational Research Journal*, 25, 31-50.
- Hawkins, J. D., Guo, J., Hill, K., Battin-Pearson, S., & Abbott, R. (2001). Long-term effects of the Seattle social development intervention on school bonding trajectories. *Applied Developmental Sciences*, 5, 225-236.
- Jackson, A.W., & Davis, G.A. (2000). *Turning points 2000: Educating adolescents in the 21st century*. New York: Teacher's College Press.
- Kagan, S. (2001). Teaching for character and community. *Educational Leadership*, 59 (2), 50-55.
- Langer, M. A., and Neal, J. C. (1987). Strategies for Learning: An Adjunct Study Skills Model. *Journal of Reading*, 31 (2), 134-39.
- LaParo, K.M. & Piata, R.C. (2003). *CLASS: Classroom Assessment Scoring System*. Charlottesville, VA: University of Virginia
- Lipsitz, J., Jackson, A., Austin, L. (1997). What works in middle-grades school reform. *Phi Delta Kappan*, 78 (7), 517-531.
- MacGinitie, W.H., MacGinitie, R.K., Katherine, M., & Dreyer, L.G. (2000). *Gates MacGinitie Tests of Reading*. Itasca, IL: Riverside.
- Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1993). Success for All: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal*, 30, 123-148.
- McEwin, C.K. & Dickinson, T.S. (1997). Middle-level teacher preparation and licensure. In J.L. Irvin (Ed.), *What research says to the middle-level practitioner* (pp. 223-229). Columbus, OH: National Middle School association.
- Meloth, M.S., & Deering, P.D. (1992). The effects of two cooperative conditions on peer group discussions, reading comprehension, and metacognition. *Contemporary Educational Psychology*, 17, 175-193.
- Meloth, M.S., & Deering, P.D. (1994). Task talk and task awareness under different cooperative learning conditions. *American Educational Research Journal*, 31 (1), 138-166.
- Moore, D.W., Bean, T.W., Birdyshaw, D., & Rycik, J.A. (1999). *Adolescent literacy: A position statement*. Newark, DE: International Reading Association.

- Mosteller, F., & Boruch, R. (Eds.). (2002). *Evidence matters: Randomized trials in educational research*. Washington, DC: Brookings.
- Muth, K. D. (1987). Structure strategies for comprehending expository text. *Reading Research and Instruction*, 27 (1), 66-72.
- National Center for Education Statistics (2005). *The nation's report card: Reading 2005*. Washington, DC: Author.
- Newmann, F.M. (1993). Beyond common sense in educational restructuring: The issues of content and linkage. *Educational Researcher*, 22 (2), 4-22.
- O'Donnell, J., Hawkins, J. D., Catalano, R. F., Abbott, R. D., & Day, L. E. (1995). Preventing school failure, drug use, and delinquency among low-income children: Long-term intervention in elementary schools. *American Journal of Orthopsychiatry*, 65 (1), 87-100.
- O'Donnell, A.M. (1996). Effects of explicit incentives on scripted and unscripted cooperation. *Journal of Educational Psychology*, 88 (1), 74-86.
- O'Donnell, A.M. (2000). The effects of explicit incentives on scripted and unscripted cooperation. *Journal of Educational Psychology*, 88 (1), 74-86.
- Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of research. *Review of Educational Research*, 64, 479-530.
- Schunk, D.H., & Swartz, C.W. (1993). Goals and progress feedback: Effects of self-efficacy and writing instruction. *Contemporary Educational Psychology*, 18, 337-354.
- Slavin, R.E. (1977). Classroom reward structure: An analytic and practical review. *Review of Educational Research*, 47, 633-650.
- Slavin, R. E. (1978). Student teams and achievement divisions. *Journal of Research and Development in Education*, 12, 39-49.
- Slavin, R. E. (1979). Effects of biracial learning teams on cross-racial friendships. *Journal of Educational Psychology*, 71, 381-387.
- Slavin, R.E. (1995). *Cooperative learning: Theory, research, and practice* (2nd Ed.). Boston: Allyn & Bacon.
- Slavin, R.E., Hurley, E.A., & Chamberlain, A.M. (2003). Cooperative learning and achievement: Theory and research. In W.M. Reynolds & G.E. Miller (Eds.), *Handbook of Psychology*, Volume 7 (pp. 177-198). Hoboken, NJ: Wiley.
- Slavin, R.E., Daniels, C., & Madden, N.A. (2005). The Success for All Middle School: Adding content to middle grades reform. *Middle School Journal*, 36 (5), 4-8.

- Slavin, R.E., & Karweit, N. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal*, 21, 725-736.
- Slavin, R.E., & Madden, N.A. (2001). *One million children: Success for All*. Thousand Oaks, CA: Corwin.
- Slavin, R.E., & Madden, N.A. (2006). *Success for All: Summary of research on achievement outcomes*. Baltimore, MD: Success for All Foundation.
- Stevens, R.J., & Durkin, S. (1992). *Using Student Team Reading and Student Team Writing in middle schools: Two evaluations*. (Tech. Rep. No. 36). Baltimore: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.
- Stevens, R.J., Madden, N.A., Slavin, R.E., & Farnish, A.M. (1987). Cooperative Integrated Reading and Composition: Two field experiments. *Reading Research Quarterly*, 22, 433-454.
- Stevens, R.J., & Slavin, R.E. (1995). The effects of Cooperative Integrated Reading and Composition (CIRC) on academically handicapped and non-handicapped students' achievement, attitudes, and metacognition in reading and writing. *Elementary School Journal*, 95 (3), 241-262.
- Stevens, R.J., Slavin, R.E., & Farnish, A.M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology*, 83, 8-16.
- Tei, E., and Stewart, O. (1985). Effective studying from text. *Forum for Reading*, 16 (2), 46-55.
- Webb, N.M., & Palincsar, A.S. (1996). Group processes in the classroom. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of Educational Psychology*. New York: Simon & Schuster Macmillan.
- Zimmerman, B.J., Bonner, S., & Kovach, R. (1996). *Developing self-regulated learners: Beyond achievement to self-efficacy*. Washington, DC: American Psychological Association.

