EFFECT OF SELF-REGULATED STRATEGY DEVELOPMENT STORY-WRITING INSTRUCTION: ADULT SCHOOL VOLUNTEERS IN ACTION

By

LESLIE ANN ROGERS

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

August, 2010

Nashville, Tennessee

Approved:

Steve Graham
Karen Harris
Deborah Rowe
Joseph Wehby

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A primary goal of education is to teach students how to read and write. Students with strong literacy skills are better equipped to succeed academically and fully participate in today's world. Although reading well obviously aides in knowledge acquisition, being able to write well is often the primary means by which knowledge acquisition is measured (see Graham, 2006b). Being able to write well also affords individuals a powerful social outlet. With the advent of electronic forms of communication, there are seemingly infinite opportunities for students to engage in online social networking forums, all of which require at least minimal writing skills. Such writing opportunities can provide multiple benefits, such as increased opportunities for self-expression, self-reflection, and communication (Graham, 2006b; Salahu-Din, Persky, & Mille, 2008).

The inclusion of writing in national educational assessments and the formation of National Panels focused specifically on writing further emphasize the increased importance of writing in today's world. Since 1998, the National Assessment of Educational Progress (NAEP, 2009) has assessed the long-term trend of students' writing performance at the state level (writing assessments at the state level began in 1992). The results from these assessments indicate that although there has been increased attention given to writing, the majority of students still do not write well enough to meet grade level demands (Salahu-Din et al., 2008). For example, in the most recent NAEP writing

assessment with 8th and 12th grade students, between 67% (grade 8) and 76% (grade 12) of students wrote *below* proficiency levels (Schneider, 2007). The results were similar with fourth-grade students in an earlier assessment (NAEP; Persky, Daane, & Jin, 2003). Approximately 67% of grade 4 students wrote below grade-level proficiency levels. These national assessments indicate that the majority of U.S. students have not acquired needed writing skills at their grade levels.

Based in part on the relatively low writing performance of American students on the NAEP, a national panel in 2003 called for a writing revolution, recommending that schools increase their attention to teaching writing and applying effective teaching procedures (National Commission on Writing, 2003, 2004; Persky et al., 2003). Effective, research-based writing procedures have been identified in a number of recent reviews (e.g., Graham & Perin, 2007a, 2007b, Rogers & Graham, 2008). These syntheses have provided valuable recommendations for improving writing practices, such as teaching students how to plan before writing.

Struggling Writers

Why are there so many struggling writers? First, writing is a difficult task, requiring considerable effort and foresight (Graham, 2006b; Hayes & Flower, 1980). Hayes and Flower (1980) studied the verbal protocols provided by skilled adult writers as they wrote, and found their writing behavior was constrained by a variety of factors, placing considerable demands on their ability to manage the writing task. Constraints included: the task environment (e.g., the topic, the intended audience), long-term memory (what is known about the topic), and key cognitive processes. One of the cognitive

processes they identified as essential to successful writing was the ability to plan (i.e., set goals and strategies to meet those goals). Essentially, Hayes and Flower found that skilled writing was a goal-directed activity that involved considerable planning and drafting (80% of skilled writers' verbal statements while composing during early writing stages were related to planning and drafting text).

The importance of planning in writing is further evident in the amount of time skilled writers devote to this activity. For example, skilled business writers spend over 66% of their time planning (Gould, 1980). Others have suggested that advanced planning makes the process of writing easier for younger writers, as it allows them to direct attention to other important processes, such as sentence construction, thinking about the needs of the reader, and so forth (Scardamalia & Bereiter, 1986). In summary, writing is challenging and planning provides one mechanism for making it more manageable (Graham, 2006b).

A second possible reason why many students struggle with writing is that they use an approach to writing that minimizes planning and other self-regulatory processes such as monitoring and evaluation (Scardamalia & Bereiter, 1986; see Graham, 2006b). Scardamalia and Bereiter (1986) found that novice writers simply convert the task of writing into telling what one knows, resulting in little overall planning and drafting or goal setting. Students first identify a relevant idea, write it down, and then use this idea to stimulate their next idea (Bereiter & Scardamalia, 1987). This differs considerably from the thoughtful and planful approach to writing applied by the more skilled writers observed by Hayes and Flowers (1980).

Why do struggling writers use the knowledge telling approach? Such an approach makes the writing process easier and less effortful, and this may serve a compensatory purpose as almost all aspects of writing, including handwriting, spelling, sentence construction, and so forth require a great deal of effort for these beginners. However, such an approach leads to inadequate text generation and places students at the lowest end of the writing expert continuum (Alamargot & Chanquoy, 2001). Pashler's (1994) bottleneck theory can be used to explain why many young students often write the first thing that comes to mind, rather than plan out a thoughtful text. The basic tenet of the bottleneck theory is that young students' attention and cognitive resources when writing are overtaxed with the basic processes of getting words onto paper, leaving few remaining resources to allocate to high-level skills such as planning and drafting (McCutchen, 1988).

One means of addressing this bottleneck is to have students plan in advance of writing. At this point, fewer resources need to be allocated to lower-level demands. By having a plan available prior to writing, the planning is not constantly competing with other writing skills, such as spelling, handwriting, or sentence construction, for scarce cognitive resources. Advanced planning can take many forms at this point: setting goals, planning what to say, thinking about text organization, and considering the audience. Given that planning plays such an important role in skilled writing, it is reasonable to assume that teaching young struggling writers to plan in advance will have a positive impact on their writing performance.

Current Study

Struggling writers tend to do little systematic and overt planning and drafting while writing. My study tackles this problem head on by teaching students how to plan in advance of writing and by encouraging them to continue the process of planning and drafting. The validated Self-Regulated Strategy Development (SRSD; Harris, Graham, and Mason, 2003) model was applied to teach students a general and more specific strategy for planning and drafting a story. Prior studies have validated the teaching of the target strategies via SRSD with young, struggling writers. I briefly describe this strategy and the SRSD model next.

SRSD story-writing instruction. In the current study, students were taught two writing strategies for planning and drafting. The first strategy was a general planning and drafting strategy for writing. The strategy processes were captured in the mnemonic POW. The three steps in POW were (a) *Pick* an idea for writing a story, (b) *Organize* notes for writing a good story (students were taught how to organize their notes using processes captured in another mnemonic which is described next), and (c) *Write* and say more (this served as a reminder for students to continue to plan while writing).

The story-specific planning and drafting strategy was captured in the mnemonic *WWW, What = 2, How = 2*. Students were taught how to generate ideas and write notes related to each of the basic parts of a story (Harris et al., 2003). Students were taught to ask themselves the following questions (not in any particular order): ***W***ho is the main character? ***W***hen does the story take place? ***W***here does the story take place? ***What*** does the main character do or want to do; what do other characters do? ***What*** happens when the main character tries to do it; what happens with other characters? ***How*** does the story

end? ***How*** does the main character feel; how do other characters feel? The questions provided a guide (rather than a formula) for planning and writing better stories. Each student's plan resulted in a unique story based on the student's specific interests. Both the POW and WWW strategy were taught within the context of the well-validated SRSD model of instruction (Harris et al., 2003).

SRSD provides a structure for teaching students both the academic and self-regulation strategies needed for completing a task (c.f., Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006). In this study, SRSD instruction was used to teach students to plan and write better stories (POW and WWW). SRSD is composed of six instructional stages: *Develop Background Knowledge; Discuss It; Model It; Memorize It; Support It; Independent Performance* (see *Chapter 2* for a complete description of the specific goals for each instructional stage). The stages can be applied in a non-linear fashion and can be revisited when necessary. Throughout each stage, students and teachers collaborate and record progress made (e.g., students record the number of story elements included in their story). The instruction is scaffolded with more explicit instruction and teacher-led support (e.g., modeling how to plan a story that includes the basic story parts) provided in the early stages. The recursive nature of SRSD stages also allows for increased levels of support in the later stages if needed. Instruction continues until students can successfully and independently apply the targeted strategies (i.e., POW and WWW, What = 2, How = 2; Harris et al., 2006).

The self-regulatory components of SRSD are also important instructional elements (Harris et al., 2003). Students are taught (and given support with) how to regulate their own behavior. For example, students are taught how to use self-talk (e.g.,

"This is hard, but I know there's a strategy I can use.") and other self-regulation strategies (e.g., self-monitoring or goal setting). Also, as noted earlier, enhanced motivation is targeted through students' use of positive self-statements and obtaining visible evidence of their progress. The self-regulation procedures (goal setting, self-assessment, self-instructions, and self-reinforcement) are developed across the same six-stage instructional continuum (e.g., Develop Background Knowledge) with the ultimate goal of independent use of said strategies.

SRSD effectiveness. The effectiveness of SRSD for improving writing outcomes has been repeatedly demonstrated in the extant literature (c.f., Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009; Graham & Perin, 2007b; Rogers & Graham, 2008; see *Chapter 2* for a more complete description of these meta-analyses).

Graham and Perin (2007b). For example, Graham and Perin (2007b) evaluated the effects of writing instruction delivered to adolescents in grades 4 to 12 using a meta-analytic approach (Lipsey & Wilson, 2001). A meta-analysis was is a method used to summarize statistics across similar studies (have all used similar research designs). Such procedures allow one to show the direction and size of a treatment's effect (see Lipsey & Wilson, 2001, for a more complete description). In Graham and Perin's (2007b) meta-analysis of the writing intervention research, they identified a total of eight experimental or quasi-experimental studies that evaluated the effectiveness of SRSD instruction for improving adolescent students' writing. First, the SRSD studies were of overall high quality (see *Chapter 2* for a more complete description of the nine quality indicators used). The average quality score among the 8 SRSD studies evaluated was 7.0 (range of 5.0 to 9.0). Second, the SRSD studies had a large overall effect on improving writing.

The average weighted effect size for SRSD was 1.14. This effect size was almost twice as large as the average weighted effect size for non-SRSD interventions where writing strategies were taught (.62). Third, all SRSD instruction was delivered by teachers or researchers.

Rogers and Graham (2008). Rogers and Graham evaluated the effectiveness of single subject writing intervention research (studies were conducted using reversal or multiple baseline designs). Out of the 88 total single subject designs reviewed, 27 (31%) were SRSD studies. The findings for the SRSD studies were similar to those reported by Graham and Perin (2007b). First, the studies that evaluated the effects of teaching students strategies for planning and drafting narrative and expository text (all SRSD) had overwhelmingly positive results. The percentage of post-treatment data points that represented an improvement over the strongest baseline data point exceeded 90% (calculated using Percentage of Non-Overlapping Data [PND], Scruggs & Mastropieri, 2001; see *Chapter 2* for a complete description) and maintained three weeks or more after the treatment had ended (mean PND range at maintenance was 86% to 90%).

Second, confidence can be placed in the findings as more than 74% (20/27) of the SRSD studies met at least 8/11 (72%) quality indicators (quality indicators included such criteria as establishing reliability for dependent variables, providing adequate participant information, and collecting treatment fidelity and social validity data; see Rogers & Graham, 2008). The final reported outcome related to the type of individual who led the SRSD instruction. In all of the 27 SRSD studies, interventions were only delivered by professionals, such as teachers and researchers.

Baker et al., (2009). An independent evaluation of SRSD's effectiveness was also

conducted. Baker et al. evaluated the quality of evidence for SRSD instruction delivered to kindergarten through twelfth grade students with or at risk for learning disabilities (LD). They reviewed experimental (n = 5) and single subject (n = 16) studies published in peer-refereed journals. The findings for these SRSD studies were similar to those reported by Graham and his colleagues (Graham & Perin, 2007b; Rogers & Graham, 2008). First, the effect sizes for the experimental studies were high, ranging from 0.80 to 1.85 (average weighted effect size was 1.22). Second, experimental control was established across all 16 single subject studies. Experimental control was established when (a) the SRSD treatment was introduced (across participants) in a systematic time-lagged fashion and (b) documented positive changes occurred (trend and level) only after SRSD treatment introduced. Third, Baker et al. reported that confidence could be placed in these findings as the research reported in the 21 SRSD studies was of high quality. They concluded that SRSD met the criteria (see Gersten, Fuchs, Compton, Coyne, Greenwood, & Innocenti, 2005; Horner, Carr, Halle, McGee, Odom, & Wolery, 2005) for being evidence based for students with and at risk for LD.

In conclusion, the three reviews provide evidence (over 35 SRSD studies evaluated) that the SRSD model is effective at improving writing performance. Studies examining the effectiveness of SRSD have only occurred with teachers or researchers conducting the intervention (Rogers & Graham, 2008) and have primarily been delivered to upper elementary and middle school students (see Harris et al., 2006).

There is evidence that SRSD (focused primarily on teaching students how to plan and write) is also effective when delivered to struggling young writers in second and third grade (Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006). The studies

used group design procedures to evaluate the effectiveness of researcher-led SRSD paired-instruction on several writing outcomes. In both studies, the comparison groups were teacher-led classroom-delivered Writer's Workshop lessons. Additionally, for each of the group studies, the number of story elements (scores ranged from 0 to 7) and story quality (scores ranged from 0 to 8) were calculated using the same methods.

Graham and his colleagues (2005) identified 73 struggling third-grade writers. Students were separated into two types of SRSD planning and drafting treatments (with and without peer support) and a comparison group. Third-grade struggling writers who were taught story planning and drafting with SRSD wrote longer stories, included more basic story elements and wrote stories of higher quality. These effects also maintained after the treatment had ended. For example, students taught how to plan using SRSD on average wrote 76.96 words (SD = 24.48) at post-test and 83.38 words (SD = 55.06) at maintenance. Students in the comparison group on average wrote 35.46 words (SD = 12.83) at post-test and 43.92 words (SD = 25.52) at maintenance. The associated effect sizes (Cohen's d) were 3.23 at post-test and 1.55 at maintenance. The results were as impressive for the effect of SRSD instruction on inclusion of basic story elements at post-test (d = 1.79) and maintenance (d = 1.81) and overall writing quality (post-test d = 2.42; maintenance d = 1.60).

Harris and her colleagues (2006) identified 63 struggling second-grade writers. Students taught to plan using the SRSD model included more basic story elements. The SRSD only group included more basic story elements at post-test (M = 5.77; SD = 1.35) than students in the comparison condition (M = 3.14; SD = 0.81). These effects maintained over time with the SRSD only group (M = 5.55; SD = .91) outperforming the

comparison group (M = 3.23; SD = 1.23). These differences were statistically significant (both p's $< .01$, d = 1.52 at post-test and d = 1.46 at maintenance).

The current study replicated and extended these studies. It replicated by re-examining the effectiveness of teaching story planning and drafting with SRSD to third-grade struggling writers. It evaluated the effectiveness of such a treatment on improving (a) the number of essential story elements, (b) holistic writing quality, and (c) the length of stories. As in previously highlighted studies, results were evaluated at post-test and maintenance. The current study also extended previous reviews of SRSD instruction by evaluating elementary students' progress at regular intervals *during treatment.* Kiuhara (2009) evaluated students' progress during SRSD instruction, but her research was conducted with high school students. The current study is the first to evaluate treatment data with elementary students. This information aided the understanding of young struggling writers' progression during SRSD instruction. These data also provided information useful to improving SRSD tutor preparation.

The current study more importantly extended these results by evaluating the effectiveness of using nonprofessional adult *volunteers* to teach students how to plan and write better stories using the SRSD model. Such a study has yet to be conducted although recent reviews of nonprofessional adult-led literacy interventions have shown that such treatments can be effective at improving young students' literacy skills (Ritter, Barnett, Denny, & Albin, 2009; see Chapter 2 for a complete review of this literature). To date, the majority of evaluations of nonprofessional adult-led literacy interventions have only evaluated the effectiveness of reading interventions. One recent exception is a study conducted by Reid, Luschen, and Lienemann (2009).

The Reid et al. study (2009) evaluated the effectiveness of an SRSD planning and drafting intervention using a multiple baseline design across participants with multiple probes collected during baseline. One paraeducator delivered instruction to three struggling second-grade writers. The paraeducator had a Bachelor of Science in Communication Studies, and had two years experience as a paraeducator, and was employed by the school district. The paraeducator had experience tutoring students in skill development, reading, and math, but did not have prior knowledge of SRSD instruction.

The paraeductor-delivered SRSD planning and drafting intervention resulted in improved writing outcomes at post-test for the number of basic story elements included in the writing. The number of sessions required for students to reach criterion was 8, 10, and 13. The mean number of story elements included during the baseline phase was 3.0, 2.3, and 4.5. The mean number of story elements included during post-treatment (or the independent performance stage) was 6.3, 3.7, and 6.3 (respectively). Two of the three students had 100% nonoverlapping data points and one student had 66% nonoverlapping data points. The student with a 66% PND showed a decreasing trend during the Independent Performance stage, but scores improved after receiving booster sessions. Students also improved the quality of their writing. Students mean scores went from 2.3, 1.0, and 3.8 during baseline to 3.5, 2.0, and 5.3 during post-treatment. Such a study supports the need to further investigate the effectiveness of teaching planning and drafting strategies to young struggling writers using individuals without professional teaching experience, such as nonprofessional adult *volunteers*, as Reid et al. found that

this instruction was effective when delivered by a paraeducator who had little formal teacher preparation.

In conclusion, the current study had both practical and theoretical importance. If planning and drafting is an important element of writing, then teaching students who do not typically plan or write well how to plan and draft should have a positive impact on their writing performance. This study tested this theoretical proposition, with the expectation that planning and drafting would improve struggling writers performance, as this was the case in previous studies (Graham et al., 2005; Harris et al., 2006). Results from similar reviews (i.e., reviews of nonprofessional adult-led literacy interventions, Ritter et al., 2009) suggest that interventions delivered by individuals other than researchers or teachers are effective, yet the majority of this research has only assessed effectiveness of reading interventions.

Objectives

While multiple studies have explored the effectiveness of using nonprofessional adult *volunteers* to deliver reading interventions, none have examined the potential effectiveness of using this group of volunteers to deliver a writing intervention. The purpose of the current investigation was to assess the effectiveness of using nonprofessional (individuals with a high school degree or above, without an education major, and without writing tutor experience) adult volunteers to deliver supplemental writing instruction as a means for improving struggling third-grade writers' story writing performance. This study added to the extant research literature in several important ways.

First, the current study extended what is known about the effectiveness of using the SRSD model to improve struggling writers' performance. The SRSD model has been well validated in the literature, with researchers and teachers primarily delivering the instruction (c.f., Baker et al., 2009; Graham & Perin, 2007a; Rogers & Graham, 2008). There has also been documented interest in examining the effectiveness of SRSD delivered by adults other than teachers or researchers. As noted earlier, Reid et al. (2009) recently examined the effects of using a paraprofessional to deliver SRSD story-writing instruction to struggling second-grade students. This study extended this line of inquiry by examining the potential effectiveness of using nonprofessional *volunteers* to deliver the writing instruction.

Second, this study extended the literature related to supplemental, nonprofessional adult-led literacy interventions in two ways. First, *nonprofessional adult volunteers* are more clearly defined than in prior investigations. A clearer definition allows researchers and schools to better understand what type of volunteers can be drawn upon to deliver instruction so that similar outcomes can be expected. This is especially important in more rural areas where college students (or adults with a more extensive knowledge of teaching) are often not available to deliver such instruction. Second, the current investigation involved nonprofessional adults implementing a *writing* intervention. To date, researchers have only examined the effectiveness of nonprofessional adult-delivered *reading* interventions.

Third, the current examination was designed to address methodological weaknesses found in previous nonprofessional adult-led literacy interventions. First, procedural fidelity was examined to ascertain the extent to which the treatment was

delivered as intended (extent to which each lesson component was applied and teaching quality). Second, social validity was assessed, both from the nonprofessional adults' and students' perspectives. Third, the procedures were adequately described. These practices were not common in prior research with nonprofessional adult-led instruction (see Chapter 2).

Research Questions

The following research questions guide this investigation:

Research Question 1: Did nonprofessional adults effectively deliver an SRSD story-writing intervention to third-grade students identified as at-risk writers? This question was answered in two ways. First, did nonprofessional adults demonstrate mastery of each lesson during training sessions? Second, did nonprofessional adults complete all intervention sessions with a high degree of fidelity? Based on previous research and training provided to nonprofessional adults (Reid et al., 2009), I expected that the nonprofessional adults would demonstrate mastery of the material during training sessions and would complete at least 90% of lesson components (for each lesson) during the intervention.

Research Question 2: Did the nonprofessional adult-led SRSD story-writing instruction result in improved writing output for struggling third-grade writers? More specifically, did treated students compose longer stories, include more essential story elements in their writing, and write stories of better overall quality? Based on previous research (Reid et al., 2009), I anticipated that the current SRSD writing treatment would result in increased

writing output (total length and number of story elements included) and overall writing quality.

Research Question 3: Did nonprofessional adult-led SRSD story-writing instruction result in maintained improvements for struggling third-grade writers? Did students maintain improved writing performance as measured across all three writing outcome measures (essential story elements, writing quality, and writing productivity,) at post-treatment (within four weeks after treatment ended) and maintenance (six weeks after treatment ended)? Based on recent meta-analyses (e.g., Rogers & Graham, 2008), I anticipated positive effects at both post-treatment and maintenance. First, I anticipated that each student's post-treatment data (mean and median levels) would be higher than their baseline data (mean and median). Second, I anticipated that each student's maintenance data (mean and median) would decrease only slightly when compared to their post-treatment data (mean and median; no more than 20%).

Research Question 4: Did students and/or nonprofessional adults report high social validity for the current SRSD story-writing instruction? Previous research (e.g., Reid et al., 2009) showed that both nonprofessional adults and students found value in the SRSD story-writing instruction and would recommend its continued use. Thus, I anticipated that students and nonprofessional adults would report high social validity.

Definition of Terms

Nonprofessional Adult Volunteers. Nonprofessional adult volunteers refers to adults (a) with a High School degree or above, (b) without an education major, and (c) without writing tutor experience. In this study, nonprofessional adult volunteers included parents as long as a parent did not instruct his/her own child. Individuals volunteered to participate, but were paid a small stipend for completing the training ($50) and intervention phases ($100 per tutee).

Structured Literacy Interventions. This involved reading and/or writing interventions with a set of specific activities that were used in a systematic fashion. Structured literacy interventions were those in which tutoring manuals and student materials were provided and at least minimal training was offered (see Wasik, 1998).

SRSD Instruction. Self-Regulated Strategy Development (Harris et al., 2003) is a six-stage model that provides a structured framework for individuals to teach academic and self-regulation strategies needed to improve students' writing (SRSD has also been used to teach math and reading, see Harris et al., 2003). The six SRSD stages are recursive, stress good teaching, and require active student collaboration. The six stages are: *Develop Background Knowledge*, *Discuss It*, *Model It*, *Memorize It*, *Support It*, and *Independent Performance*. In this study, nonprofessional adults delivered an SRSD story-writing intervention.

At-risk Writers. Students were considered at-risk writers if they scored at or below the 16th percentile on the Test of Written Language-3 (TOWL-3, Hammill & Larsen, 1996). This test assessed student's ability to write a complete and interesting story.

Oral Reading Fluency (ORF). Oral reading fluency (ORF) was determined from the scores obtained on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2007) assessment. Students read three passages out loud to a school examiner. The middle ORF score from these measures was used to assess each student's ORF level. The descriptive levels of performance for the middle of third grade were as follows: students reading at or below 67 words per minute were considered *at-risk* for reading difficulties; students reading between 67 and 91 words per minute were considered *at some risk* for reading difficulties; students reading 92 words per minute or more were considered at *low risk* for reading difficulties.

Writing Academy. The nonprofessional adults, struggling third-grade writers, and classroom teachers referred to the current SRSD story-writing intervention as the Writing Academy.

CHAPTER II

LITERATURE REVIEW

The purpose of this study was to assess the effectiveness of using adult nonprofessionals (individuals with a High School degree or above, without an education major, and without writing tutor experience.) to deliver supplemental writing instruction as a means for improving struggling third grade writers' story-writing performance. In an attempt to support this line of inquiry, I first present a review of the literature related to *all* supplemental adult-led supplemental literacy instruction and its effects at improving academic outcomes in reading (no such information is available in writing). This summary provides the context in which the nonprofessional adult-led literacy intervention was conducted. Next, I review the literature related to *nonprofessional* adult-led literacy interventions. Such a review has not been conducted. This information is useful in understanding the design of the proposed study (e.g., amount of training/feedback that should be provided). Finally, I describe the Self-Regulated meta-analyses) on its effectiveness. As part of the SRSD review, I examine one particularly relevant SRSD study (Reid, Luschen, & Lienemann, in 2009), as the story treatment used in this investigation is similar to what was done with third-grade struggling writers in this study.

Despite conducting a comprehensive search, I was unable to locate any studies that focused on adult-led literacy instruction in writing. Consequently, I reviewed *all* adult-led nonprofessional instruction in the area of reading, a similar but not identical

skill, to determine issues that should be attended to when providing such instruction in the area of writing. Before describing the methods and results for each literature review (*all* adult-led supplemental literacy interventions, *nonprofessional* adult-led literacy interventions, and SRSD instruction), I first provide the rationale for providing nonprofessional adult-led literacy instruction in reading and define the purpose and operational definitions related to the first two reviews.

Rationale for Providing Nonprofessional Adult-led Literacy Instruction in Reading

Reading performance is not what it should be (America Reads Challenge, ARC, White House, 1997; National Assessment of Educational Progress, NAEP, Lee, Grigg, & Donahue, 2007). An alarming number of students continue to fall below nationally designated grade-level reading proficiency levels (Lee et al., 2007). For example, on the most recent NAEP reading assessment, approximately 7 out of every 10 fourth-grade students were only at the basic or below basic reading levels (Lee et al., 2007). That most young students are only partially mastering the prerequisite knowledge and skills considered fundamental for reading proficiently at their grade level is unacceptable.

As a consequence of these relatively weak reading skills, there has been an increased emphasis on raising reading performance. The focus of such improvements has occurred through a greater emphasis on scientific-based instruction (NCLB, 2001; Partnership for Reading, 2000) and the use of peer review or peer tutoring procedures (Cohen, Kulik, & Kulik, 1982; Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006; Topping & Hill, 1995). Schools have also attempted to improve classroom instruction by eliciting the help of adult nonprofessional adults to teach struggling readers (Erlbaum, Vaughn,

Hughes, & Moody, 2000; Ritter, Barnett, Denny, & Albin, 2009; Wasik, 1998). The last of these, programs that rely heavily on adult volunteer support, are often associated with lower costs, and can often be supported through government funding (e.g., ARC, 1997; Serve America Act, 2009). However, it is important to determine if these programs are effective.

Over the past decade, there has been increased interest in examining the effectiveness of adult volunteer-led reading programs and increased attention to the specific impact tutor qualification has on overall results (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). Wasik (1998) was among the first to examine the effectiveness of studies using adult nonprofessional adults on improving young students' reading outcomes, citing the importance of knowing whether a growing number of federally funded programs were in fact effective. Erlbaum et al. (2000) more specifically investigated the relation between tutor expertise and overall effect sizes, claiming that schools seeking more cost-effective approaches (not having to pay professional tutors) needed to know if such programs were effective. Ritter et al. (2009) conducted a more in-depth systematic evaluation of adult volunteer-led supplemental reading programs, citing the greater need for such an evaluation to occur due to increased school accountability.

In these previous reviews, nonprofessional adults have not been clearly defined. The evaluations thus far have included a mix of tutors identified loosely as *community members* or *nonprofessional* adults. For example, Ritter et al. (2009) included all randomized field trials in which adult tutors were identified as nonprofessional adults, and stated that such individuals were nonprofessionals. In fact, this group of nonprofessional adults included a variety of individuals from very experienced

individuals (e.g., retired teachers) to those without any teaching experience. It seems plausible that results would vary (or need for more training would exist) based on such distinctions. To date, no evaluation has occurred that assesses the effectiveness of supplemental reading programs on improving students' reading outcomes, when such programs are delivered by a specific subset of the community volunteer population: nonprofessional adults (adults without formal teaching experience).

Purpose and Operational Definitions

The purpose of this literature review is to evaluate the effectiveness of nonprofessional adult-led programs on improving students' literacy outcomes. A nonprofessional adult is defined as adults (a) with a High School degree or above, (b) without an education major, and (c) without writing tutor experience. For the purpose of this literature review, this could include parents as long as those parents were providing remediation to more than their own child (e.g., Neuman & Roskos, 1993).

College students are excluded from this evaluation for several reasons. First, such evaluations of this group have occurred elsewhere (Erlbaum et al., 2000; Ritter et al., 2009). Additionally, the effects of programs using college students have typically been separated from that of other adult-delivered tutoring programs, with average effects sizes in the moderate to large range (see Erlbaum et al. and Ritter et al.). Finally, the elimination of college students in the current evaluation will allow for clearer generalizations, especially for schools unable to draw from such populations (not located near universities or colleges).

Effectiveness of all Adult-Led Supplemental Literacy Interventions

Previous reviews (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998) of supplemental adult-led reading instruction (including instruction led by teachers and paraprofessionals) support two main conclusions. First, although results must be tempered, the overall consensus is that such programs have positive effects on improving reading outcomes (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). Second, although positive, the results vary considerably from study to study (Erlbaum et al.; Wasik). In the next section, these two main conclusions are examined in greater detail.

Synthesis of Previous Reviews

Overall positive effects. There have been three frequently cited reviews on the effects of supplemental adult-led reading interventions for struggling readers (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). Wasik was among the first to review the effectiveness of supplemental adult-led reading instruction for struggling readers. Wasik conducted a best-evidence synthesis of 17 supplemental reading programs delivered to students in kindergarten to third grade. She concluded that there was an overall lack of experimental research to support individual treatments, but found that more successful programs were more structured and provided more tutor training (especially if the treatment was more involved).

Erlbaum et al. (2000) conducted a meta-analysis of 29 experimental studies that investigated the effectiveness of supplemental reading programs delivered to struggling students in grades one to six. Erlbaum et al. reported that there was much variation across

the experimental studies, with a small overall effect on improving students' reading outcomes (ES = .41). Finally, Ritter et al. (2009) conducted a meta-analysis of 21 randomized field trials that investigated the effectiveness of supplemental reading programs delivered to struggling young students in kindergarten to eighth grade. The effects were small overall with improvements on global reading (ES = .41) and oral reading fluency (ES = .30) measures.

Although the effects from these reviews are positive and are drawn from a growing body of experimental research, some caution must be used when interpreting the findings. One issue is the overall small sample sizes in individual studies. In the most recent review of this literature, Ritter et al. (2009) identified 11 samples in 9 experimental studies in which community nonprofessional adults delivered the instruction. In all but one of the studies (Baker, 2000), the treatment groups numbered less than 30. The average number of students in the tutored groups was 20 (range from 11 to 43). Other issues are inadequate treatment descriptions (Erlbaum et al., 2000) and omission of treatment fidelity data (Ritter et al., 2009). Despite these weaknesses, researchers have been able to draw several specific conclusions about adult-led supplemental reading instruction. Although this review specifically addresses nonprofessional adult-led instruction, it is important to examine these previous reviews as they provide important information on factors that may influence outcomes when just nonprofessional adult-led reading instruction is considered.

Variability in outcomes. There are several common threads that tie the literature on adult-led supplemental reading instruction together. Such programs have been identified as effective, but effectiveness varied depending on: (a) program structure, (b)

tutee characteristics, (c) tutor characteristics/training, (d) treatment duration, and (e) outcome measures (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998).

It should be noted that unless specifically stated, Erlbaum et al.'s (2000) results represent an average across studies that used paid/professional adults (teachers, paraprofessionals) and volunteer adults (college students and community members). The other two reviews represent conclusions about all adult *volunteer*-led remedial reading instruction, however these tutors were often paid (Wasik, 1998; Ritter et al., 2009). Additionally, Ritter et al.'s review included studies that used students' own parents to deliver the instruction.

Program structure. The following conclusions are drawn from the literature related to adult-led supplemental reading instruction (to include those delivered by teachers and paraprofessionals). Effect sizes varied based on program structure, with more structured programs surpassing those that were less structured (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). Wasik reported that programs that were more structured/comprehensive (e.g., included explicit modeling and scaffolded instruction) had more positive outcomes than those that were not or those that did not provide treatment descriptions. Ritter et al. also reported that highly structured programs were superior (ES = .59) to less structured programs (ES = .14) in improving global reading outcomes.

Others attributed the variation in effect sizes to instructional focus. For example, Erlbaum et al. (2000) found that programs that focused on mixed skills (focus on decoding, comprehension, and word recognition; $n = 30$; $d = .50$) were superior to those that focused on either visual-perceptual skills ($n = 2$; $d = .03$) or did not provide adequate

information to determine focus (*n* = 4; *d* = -.07). Most of the reviews on the effects of adult-led supplemental reading instruction have assessed effectiveness of programs on young students (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). There has been a clear interest in the relation between tutee characteristics and overall reading outcomes.

Tutee characteristics. These conclusions are based on the literature related to adult-led supplemental reading instruction (to include those delivered by teachers and paraprofessionals). Research on the effectiveness of such programs has included studies conducted with students in grades kindergarten through third grade (Wasik, 1998), first through sixth grade (Erlbaum et al., 2000) and kindergarten through eighth grade (Ritter et al., 2009). All of the researchers have examined variability in student responsivity. There have been mixed results when examining the effects of grade level on improved reading outcomes.

Some researchers have found statistically significant differences in outcomes, with larger effects observed for younger students than for older students (Erlbaum et al. 2000). For example, Erlbaum et al. found statistically significant differences between effects for students in grades 1 through 3 (*n* = 28; *d* = .46) and grades 2 through 3 (*n* = 8; *d* = .37) and those for students in grades 4 through 6 (*n* = 5, *d* = .06). Others have reported that no statistically significant differences related to tutee grade level exist or have suggested differences were due to other tutee characteristics (Ritter et al., 2009; Wasik, 1998).

However, Ritter et al. (2009) reported that programs delivered to first grade students were not superior to those delivered to second through sixth grade students (outcome only described in text). Wasik did not investigate the differences in outcomes

related to grade level, but suggested that there were apparent differences in students' responsivity. Wasik (1998) concluded that tutoring appears to be more effective for certain students and called for more research in this area. Unlike the mixed outcomes related to tutee grade level and overall outcomes, there has been more agreement on the impact of tutor characteristics and importance of training on overall reading results.

Tutor characteristics and training. These conclusions are based more specifically on the literature related to adult-led supplemental reading instruction delivered by nonprofessional adults (community nonprofessional adults) as opposed to instruction delivered by other adults (teachers, paraprofessionals, college students, and students' own parents). Research has shown that effect sizes vary across several tutor-level variables such as (a) tutor training/support, (b) overall consistency of services provided, and (c) incentives provided for participation (Erlbaum et al., 2000; Meyer, 2008; Wasik, 1998, Wasik et al., 2002).

First, effect sizes were of moderate size in studies that described tutor training. Erlbaum et al. (2000) investigated the effects of tutor training on a small sample of studies that used community nonprofessional adults as tutors (8 samples found in 6 studies). They found that reporting tutor training was associated with ES variability. The 6 comparisons (found in 5 studies) that described community volunteer training procedures reported a larger overall effect than the 2 comparisons (found in 1 study) that did not describe training (d = .59 vs. d = -.17).

Wasik (1998) also reported that tutor training and support were related to improved reading outcomes. Wasik reported that programs that emphasized adequate training (i.e., included an understanding of basic reading principles) and regular feedback

were superior to those that provided only minimal training and support. However, Wasik cited cases in which extensive training was not provided, yet success was still observed (Helping One Student to Succeed; HOSTS) and suggested training is most needed for programs that have an emphasis on students engaged in higher level reading and writing programs and requires "informed judgment on the part of the tutor" (Wasik, p. 283).

Erlbaum et al. (2000) also found that the consistency of the services provided impacted overall program success. Erlbaum et al. reported that effect sizes were larger when tutors came to each tutoring session and tutored for the full amount of time. Erlbaum et al. cited Vadasy, Jenkins, Antil, Way, and O'Connor's (1997a) findings to support this conclusion. Vadasy et al. reported that effects were larger for students tutored by more consistent tutors (ES = .85) versus those tutored by less consistent tutors (ES = .06).

Researchers have also suggested that tutor compensation may be important to a program's overall success (Wasik et al., 2002; Meyer, 2008). For example, Wasik et al. (2002) reflected on the reviewed research, as well as their own experiences, and suggested that using incentives to maintain tutor participation is beneficial to the success and longevity of a tutoring program. Another leading researcher in the area of volunteer-led instruction (Meyer, 2008, personal communication), stated that providing perks for involvement is important to a programs' overall success; including providing a small stipend or paying for parking. There has also been much reported on a program's treatment length and its relation to a program's overall success.

Treatment duration. The following conclusions are drawn from the literature related to adult-led supplemental reading instruction (to include those programs delivered

by teachers and paraprofessionals). It turns out that more is not always better, in fact, more may lead to lower outcomes, depending on how treatment duration is measured (Erlbaum et al., 2000). Erlbaum et al. found that studies that lasted 20 weeks or less *(n* = 14) had a larger mean weighted ES than those that lasted over 20 weeks *(n* = 16; *d* = .65 vs. *d* = .37). However, Erlbaum et al. did not find that total time in treatment was associated with variability in treatment outcomes. Those treatments that lasted less than 50 hours *(n* = 15) were not statistically different than those that lasted between 50 and 150 hours *(n* = 12; *d* = .38 vs. *d* = .32).

Wasik (1998) reported that the results related to this topic were complicated and were more likely due to treatment quality rather than quantity. Wasik found that some studies showed statistically significant differences for students who received more treatment (e.g., Book Buddies), whereas others did not (Juel, 1996; see Wasik, 1998). She concluded that it was the quality of the program, rather than the length of the treatment, that most influenced a program's overall success. The final variable found to impact the overall effect of adult-led supplemental reading instruction programs on improved reading outcomes is type of outcome measures.

***Outcome measures*.** Again, the following conclusions are drawn from the literature related to adult-led supplemental reading instruction (to include those programs delivered by teachers and paraprofessionals). The type of outcome measure has been associated with differential effect sizes (Erlbaum et al., 2000; Ritter et al., 2009). Erlbaum et al. reported that the *type* of outcome measure was significantly associated with variation in effect sizes. Erlbaum et al. disaggregated outcomes into the following categories: comprehension (reading and listening), fluency (words and passages), word

recognition (decoding, composite reading), spelling, and writing (writing vocabulary and writing). The majority of the outcome measures had moderate effects *(n* = 91; *d* = .51, range .41 to .68). The strongest effects were observed with the writing vocabulary outcome measures *(n* = 10; *d* = 94), although the researchers noted that these measures had weak psychometric properties. The weakest effects were observed with spelling *(n* = 11; *d* = .14) and reading comprehension measures *(n* = 19; *d* = .28).

Ritter et al. (2009) also reported that the weakest effects were observed on reading comprehension measures. Ritter et al. separated reading outcomes into five categories (overall, global, letters and words, comprehension, and oral reading fluency). The effects were mostly small, ranging from .26 (n = 13; reading global) to .41 (n = 15; reading letters and words). The only category in which statistically significant differences were not found was reading comprehension (n = 8; ES = .18).

Summary of Effectiveness of all Adult-Led Supplemental Literacy Interventions

In conclusion, much can be summarized from the research related to all adult-led supplemental reading programs. The overall results are positive, but must be explored tentatively as there are several methodological weaknesses present in many of the studies. Although positive, results vary across five main variables: program structure, tutee characteristics, tutor characteristics/training, treatment duration, and outcome measures. To date, these conclusions cannot be generalized to programs that use nonprofessional adults (those without formal teaching experience or training), as such a review has not been undertaken.

Based on the aforementioned reviews and recommendations, a systematic review was completed to assess the effects of using nonprofessional adults to deliver supplemental reading instruction to struggling pre-school and school-aged children. The literature strongly supports the need for such a review (i.e., the growing need for more 1:1 programs to increase overall literacy rates, need to assess cost-effectiveness of such programs, and era of greater school accountability). The literature also supports the direction for how such a review should be accomplished (i.e., variability be examined related to five categories of study characteristics just reviewed).

The next section describes the methods used to conduct this review. Using the key variables identified in previous reviews of adult-led supplemental literacy programs (see Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998), five variables were identified: program structure, tutee grade level, tutor training and level of support, treatment duration, and outcome measures. The conclusions will offer a clear, and much needed, extension of the literature.

Current Review: Effectiveness of *Nonprofessional* Adult-Led Literacy Interventions

Method

Location and selection of studies. The strategies used to select studies were influenced by two factors. First, the current review was aimed at a specific type of tutor participants. Studies were included if the intervention agent(s) were clearly identified as nonprofessional adults. Nonprofessional adult was operationally defined as a post-high school student who did not have any teaching experience or formal training. As stated

earlier, this could include parents if those parents were teaching multiple students (not just teaching their own child). This did not include college students as (a) such evaluations have been conducted and (b) including such a group would limit the generalizability of the data. To date, such a review has not been accomplished.

Second, the current review also drew from a wider selection of available studies than did previous reviews. Studies were not limited to elementary-aged tutees as in previous reviews (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998; Wasik et al., 2002). Studies were included if they occurred in a pre-school through twelfth grade classroom. Additionally, qualitative and single subject studies were not excluded as in other reviews (Erlbaum et al., Ritter et al.).

A systematic literature review was conducted to identify intervention articles published between 1975 and 2008 that focused on school-based academic interventions delivered by nonprofessional adults. Electronic searches were conducted in psychology and education databases (PsycINFO, ProQuest, Education full-text, and Dissertation Abstracts) using the following combinations of terms: (a) one-to-one, academic, literacy, school-based, or writing; (b) intervention, study, training, instruction, tutoring, or treatment; and (c) volunteer, community, nonprofessional, or non-parent. Abstracts were read and, when the study seemed applicable, full articles were retrieved for further review.

Another source for studies was the reference lists in the obtained articles. All reference lists were checked and applicable *(n* = 105) studies were collected.

Exclusion of studies. Studies were omitted if (a) tutor qualifications could not be ascertained or if tutors were other than nonprofessionals teaching children other than their

own; (b) they were experimental or quasi-experimental designs but did not provide the information needed to establish an effect size; or (c) they were qualitative but did not use academic-related outcome measures to assess overall effects. The most common reason for excluding a study related to the tutor *type*. Two tutor types frequently used in studies were students' own parents and college students. Several of the parent-led interventions were found in a recent meta-analysis (Erion, 2006). Erion reported that parent-led interventions had positive effects on student achievement. There were also many studies conducted by college students whose participation in some way related to specific course requirements (e.g., Edl, 2007, Woo, 2005). In an attempt to disentangle possible confounding factors (e.g., relationship with tutees, degree to which sustained tutor participation was effected by a course grade), these studies were omitted.

The second most common reason for omission was that the study did not provide sufficient data. Qualitative designs were included, but only when effects on academic outcome measures were included. Qualitative studies that did not meet this requirement were excluded (e.g., Allen & Chavkin, 2004; McLurkin, 2006). Experimental and quasi-experimental studies needed to provide sufficient data to calculate effect sizes (include a control group and report standard deviations). Those studies that did not include a control group (e.g., Smetana, 2005) or used a control group from a previous study (Bell, 2001) were excluded. Studies that did not report standard deviations were also excluded (e.g., Meier & Invernizzi, 2001).

Inter-rater reliability on inclusion/exclusion of studies was conducted. The researcher randomly selected 25% of the 105 articles (n = 24; 10/19 included studies and 14/86 rejected studies) to evaluate reliability. A graduate student, blind to the purpose of

the study, made the same decision as the researcher 83% (20/24) of the time. Most of the disagreements (3/4) occurred because there was disagreement on whether it was necessary for a comparison to be made to a control group that received no treatment. Consensus was reached before final analyses occurred. Those studies that did not include a control group that received no treatment were excluded.

Interrater reliability was also performed for all other areas in which studies were coded. These areas (e.g., structured versus unstructured programs) are described in the next section. The 10 quality indicators are presented further in text. In all, 26% (5/19) of included studies were randomly selected and coded by another graduate student. The overall reliability for each of the 13 categories (11 quality indicators as well as structure and focus of program) was good with an overall agreement of 82% (range 60% to 100%). When disagreement occurred, discussion occurred until a mutual consensus was reached.

Grouping of studies. In all, there were 19 experimental and quasi-experimental studies and 1 qualitative study identified and analyzed for the purpose of this review. All of the interventions were reading interventions. In all, there were 12 different treatments (as well as 3 unspecified treatments) examined in the 19 studies (each identified as they are presented in Table 2): Tutor Assisted Intensive Learning Strategies (TAILS; $n = 1$); Helping One Student to Succeed (HOSTS; $n = 3$); Houghton Mifflin Tutorials ($n = 1$); Howard Street Tutoring Program ($n = 1$); Ginn Reading Series ($n = 1$); unspecified ($n = 3$); Intergenerational Literacy Project (ILP; $n = 1$); Sound Partners (SP; $n = 3$); Start Making a Reader Today (SMART; $n = 1$); Intergenerational Tutoring Program ($n = 1$); Thinking Partners (TP; $n = 1$); Paired Reading (n =1); and Poetry Academy ($n = 1$). In order to better describe these treatments and their effects, the studies were separated into

specific categories based on the intervention's degree of structure and the focus of the intervention.

Structured versus unstructured programs. Programs were first separated into two groups, depending on the amount of structure in each tutoring program. Structured programs were defined as those with a set of specific activities that volunteer tutors had to use in a systematic fashion. Programs identified as structured had to provide evidence that tutors were provided with: (a) manuals, (b) student materials, and (c) at least minimal training (see Wasik, 1998). Unstructured programs were those that clearly stated they were unstructured (e.g., Baker et al., 2000) or did not meet one of the three mentioned criteria.

The majority of the studies provided evidence of structure within text (14/19; 74%). For example, Al Otaiba et al. (2005) described a treatment (TAILS) delivered by nonprofessional adults who had received 13 hours of training. The nonprofessional adults were provided manuals that included detailed information on activities (e.g., phonological activities, building words using manipulatives, decoding activities, and/or reading) as well as teaching strategies (model, lead, and test format).

Other program descriptions were obtained from reputable sources. For example, Burns (2004) provided a thorough description of the HOSTS program whereas Ramey (1990, 1991) did not. As both were evaluating the effectiveness of the HOSTS program, Burns description was used to place the Ramey studies in the appropriate category). The majority of the studies reviewed were structured (17/19; 89%).

Multiple-foci versus single-focus programs. After separating programs based on their structure, they were further separated based on the comprehensiveness of the

treatment (see Erlbaum et al., 2000). Most of the treatments were highly involved programs with multiple foci (14/19; 74%). These treatments focused on reading comprehension, word recognition, and reading fluency. In an attempt to draw conclusions about similar types of programs, the 14 studies with multiple foci and high levels of structure were separated and reported first in Table 2. The other 5 studies with only a single focus (structured or unstructured) were separated and described next in Table 2.

The singularly-focused programs concentrated on reading comprehension or reading fluency. The majority of these treatments were structured programs, which focused on reading comprehension *(n* = 3). The remaining were unstructured programs with a focus on reading comprehension *(n* = 1) and reading fluency *(n* = 1).

The decision to create categories based on the structure and foci of the treatment differed from previous reviews in that there was an attempt to compare data across similar types of programs. This has not been done in previous reviews (Erlbaum et al., 2000; Wasik, 1998; Wasik et al., 2002). This was facilitated by computing an average effect size for studies for a specific type of treatment (e.g., HOSTS) where the outcomes were similar across studies. Additionally, average effect sizes were computed across all studies in a category (e.g., structured programs with multiple foci).

Calculation of effect sizes. Effect sizes were calculated for the 19 experimental and quasi-experimental studies (see Table 2). Cohen's *d* was used to compute the effect sizes. This involved subtracting the control group's mean posttest score from the treatment group's mean posttest score and dividing that number by the pooled standard deviation (SD) of the two groups. When possible, raw scores were used to make these

calculations. When standardized norm-referenced scores were the only data provided, these data were used and documented as such (see Table 2).

Four different effect sizes were calculated. First, the average effect size for like measures (e.g., two word recognition measures) within a study was calculated. Additionally, if a study included both post-treatment and maintenance data, those effect sizes were reported separately (e.g., effect sizes for word recognition at post-test and maintenance). Second, for each type of treatment (e.g., HOSTS) within a category (e.g., structured programs with multiple foci) that contained two or more studies, average effect sizes for specific measures (e.g., word recognition at post-test and maintenance) were computed when possible.

Third, effect sizes were averaged across studies within a category (e.g., structured programs with multiple foci) using two procedures. The first procedure involved averaging effect sizes across like measures (e.g., word recognition at post-test), presenting averages for standard scores and raw scores separately. The second procedure involved averaging effect sizes for a measure (e.g., word recognition at post-test and maintenance) across all studies (regardless of type of score) in a category (e.g., structured programs with multiple foci).

Fourth, at the end of the quantitative analysis, the average effect size for specific measures (e.g., word recognition at post-test and maintenance) was computed for all studies (see Table 2). The format for the final summary section was similar to category (e.g., structured programs with multiple foci) summaries. The average effect sizes across all like measures (e.g., word recognition at post-test and maintenance) in Table 2 were first separated by type of score used (standard or raw). Finally, The average effect sizes

across all like measures (e.g., word recognition at post-test and maintenance) were reported (regardless of type of score).

The procedures for calculating effect sizes differed from previous reviews. To illustrate how these procedures differed from previous reviews, take Erlbaum et al.'s (2000) analysis of Vadasy et al.'s (2000) study. Erlbaum et al. averaged effect sizes for nine different measures into one mean within-sample ES of .98. When these ES were calculated using the current procedures, two post-treatment ES were identified (fluency = .49; word recognition = 1.00) as well as two maintenance ES (fluency =-.09; word recognition = .52). Such breakdowns of data are more informative and can improve the precision of analysis.

Coding of quality indicators. The quality of experimental and quasi-experimental designs was evaluated based on recommendations made by Gersten et al. (2005). There were 10 quality indicators that were chosen (see Table 1 for complete definitions of all quality indicators). The first three quality indicators related to *student participants*: (a) equivalent mortality for experimental and control groups; (b) pretest equivalence of experimental and control groups, and (c) description of and equivalence of intervention agents. Studies received a score of 0 or 1 for providing information related to the first two criteria and a 0, .5, or 1 for the last criterion. More specifically, studies received only partial credit (.5) for tutor description if only one of the following was provided: total number, education, tutor/work experience, or ethnicity.

The next three quality indicators related to *treatment*: (a) specification of instructional conditions for experimental and control conditions; (b) description of tutor training; and (c) treatment fidelity established. Studies could receive a 0, .5, or 1 for each

Table 1. Definition of the 10 Experimental and Quasi-Experimental Quality Indicators

Quality Indicator	Definition
Equivalent Mortality Rates	Studies received a 0 or 1. Studies received a 0 if they did not report similar attribution rates from pre and post-test AND did not state that a similar number of students left both the control and experimental groups.
Pretest Equivalence of Students	Studies received a 0 or a 1. Studies received a 0 if they did not analyze pre-treatment performance data to show pretest equivalence.
Pretest Equivalence of Tutors	Studies received a 0, .5, or 1. Studies received a 0 if they did not provide any of the following: number of tutors, level of tutor education, tutor work/tutoring experience, tutor ethnicity. Studies received a .5 if they provided information for at least one of the areas. Studies received a 1 if they provided information in more than 1 area.
Instructional Conditions Described	Studies received a 0, .5, or 1. Studies received a 0 if insufficient data (you could not replicate based on information provided) were provided for both the control and experimental condition. Studies received a .5 if replication could occur for at least one of the conditions. Studies received a 1 if replication could occur for both conditions.
Tutor Training Described	Studies received a 0, .5, or 1. Studies received a 0 if tutor training was not described. Studies received a .5 if a general plan and total time in training were described, but specific components of training were omitted. Studies received a 1 if procedures could be replicated. Such studies reported total time and specific training components.
Treatment Fidelity	Studies received a 0, .5, or 1. Studies received a 0 if treatment fidelity was not described. Studies received a .5 if treatment fidelity was described, but it was unclear the number of sessions observed or whether the treatment was in fact delivered as described. Studies received a 1 if at least 25% of all treatments were observed/recorded and data offered to show treatment was delivered as stated.
Control for Hawthorne Effects	Studies received a 0 or a 1. Studies received a 0 if no information was provided about whether the tutees behaviors were affected by knowledge of participating in the study. Studies received a 1 if they provided some type of treatment in the control condition.
Dependent Variables Reliable	Studies received a 0 or a 1. Studies received a 0 if they did not report this information or if the reliability reported was below .60. Studies received a 1 if the dependent variables' reliability were above .60.
Control for Tutor Effects	Studies received a 0 or a 1. Studies received a 0 if they did not randomly assign tutors to tutees or tutors did not work with students in both control and experimental conditions. Studies received a 1 if they did randomly assign tutors to tutees or tutors did work with students in each condition.
Random Assignment	Studies received a 0 or a 1. Studies received a 0 if they did not randomly assign tutees to experimental and control conditions. Studies received a 1 if they did randomly assign tutees to both conditions.

of these criteria. Studies received more than a .5 for specification of instructional conditions if such procedures could be replicated. Studies received more than .5 for tutor training if more than total time and a brief overview were provided. Finally, studies received more than a .5 for treatment fidelity if at least 25% or more of all sessions were observed, across all phases of the study.

The final four quality indicators related to *outcome measures* and *data analysis*. The three quality indicators related to outcome measures were: (a) control for Hawthorne effects, (b) dependent variable reliability, and (c) control for instructor effects (i.e., randomly selecting tutor/tutee pairs). The final quality indicator related to *data analysis*: random vs. non-random assignment. Each of these criteria was scored as either 0 or 1.

Results for Experimental and Quasi-experimental Studies

Table 2 includes data related to the 19 experimental and quasi-experimental studies that examined the effectiveness of one-to-one tutoring interventions led by nonprofessional tutors. This includes information about the (a) study's purpose, research questions and focus of instruction; (b) number and ethnicity of participating students, and geographic location; (c) dependent variables; (d) tutors' experience, compensation, training, and treatment fidelity; (d) results as well as mean effect sizes reported across each outcome measure; and (e) overall quality of the study (based on the 10 quality indicators). Table 3 presents data related to the overall quality of each of the 19 studies.

Table 2. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Structured Programs with Multiple Foci (Comprehension, Fluency, and Word Recognition *(n* =13)						
Al Otaiba et al. (2005)	Evaluated effects of 8-month Tutor Assisted Intensive Learning Strategies (TAILS) instruction on multiple reading measures. Treatment: 4 days/week Control: Regular Practices (Comprehension, Fluency, Word Recognition)	N[1] = 24 Kinder. AA (n = 47) O (n = 1) URBAN (During School)	Comprehension (n =1) - Woodcock Reading Mastery Test Revised (WRMT-R) – passage comprehension Word Recognition (n =2) – WRMT-R – word identification and word attack subtests (All measures reliable)	N2 = 12 (7 without previous tutoring experience, aged 30 to 65) Received a small stipend Training: 3 sessions totaling 13 hours Treatment Fidelity was collected and provided.	The 4-day treatment was better than control for short-term gains. Comprehension: .73 Word Recognition: .55	80%
Burns et al., 2004	Investigated effectiveness of a web-based learning tutorial, HOSTS (Helping One Student to Succeed), on reading achievement measures. Treatment: 6 HOSTS schools (4 days/week) Control: 4 Schools -Regular Practices (Comprehension, Fluency, Word Recognition)	N[1] =129 Kinder. to Grade 4 C (n = 94) AA (n = 32) H (n = 3) URBAN, RURAL, and SUB (During School)	Comprehension (n =1) - Gray Oral Reading Test (GORT) Fluency (n =2) - GORT and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) test of oral reading fluency (All measures reliable)	Recruited from community No pay Training: 2 hours but supervised 100% Treatment Fidelity: Not provided.	Statistically significant differences for HOSTS tutored group on fluency and comprehension measures. Comprehension: .50 Fluency: .37	25%

Table 2 continued.. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Mc-Grady (1983)	Examined effects of 10-month programmed tutoring (Houghton Mifflin Tutorials) intervention on reading comprehension. Treatment: 5 days/week Control: Regular Practices (Comprehension, Fluency, Word Recognition)	N^1 =35 Grade 4 RURAL (During School)	Comprehension *(n* =1) Reading Comprehension Test, Iowa Tests of Basic Skills – mean normal curve equivalent scores. (All measures reliable)	N^2 =36 Nonprofessionals, 14 tutors had NOT tutored before Paid minimal hourly wage Training: 15-20 hours Treatment Fidelity: Not provided	Experimental group made gains but post- test differences were not statistically significant. Comprehension: - .35[3]	15%
Morris et al. (1990)	Evaluated effectiveness of 2 years of an 8-month/year Howard Street tutoring program on multiple reading measures. Treatment: 4 days/week Control: Regular practices (Comprehension, Fluency, Word Recognition)	N^1 = 30 Grades 2 to 3 URBAN (After School)	Fluency *(n* =1) sentence reading – basal (unique scores based on grade- level accuracy plus an additional 10 achievement points) Word Recognition *(n* =3) identified isolated words on three measure (Reliability not provided)	Nonprofessionals, 14 tutors had NOT tutored before Paid minimal hourly wage. Training: Minimal Treatment Fidelity: Not provided.	Similar results across both years. Tutored exceeded control on word recognition and fluency measures. Fluency: .69 Word Recognition: .46	30%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Nielson (1991)	Compared effectiveness of 9-month one-to-one adult tutoring (used Ginn Reading Series, World of Reading) on reading comprehension. Treatment: Students received points for each reading activity done completely the first time. Control: Regular Practices (Comprehension, Fluency, Word Recognition)	N[1] = 29 Grade 3 25% with behavioral problems RURAL (During School)	Comprehension *(n* =1), reading comprehension subtest of the Stanford Achievement Test (SAT) (All measures reliable)	Paid: Parents tutors given free lunches on tutoring days Training: Time not provided Treatment Fidelity: Not provided.	No statistically significant differences. Comprehension Post: 1.31[2] Maintenance: - 1.91[2] *(n* = 38, 26 in treatment group and 12 in control)	30%
Ramey (1990)	Examined effectiveness of 1985-1987 HOSTS reading intervention in Seattle on California Test of Achievement test results. Treatment: HOSTS Control: No intervention (Comprehension, Fluency, Word Recognition)	N[3] = 33 Grades 2 to 3 *(n* = 8) & Grades 4 to 6 *(n* = 25) URBAN (After School)	Comprehension *(n* =2) – California Test of Achievement (CAT) total reading score combining both the primary and intermediate scores. (Reliability not provided)	Paid: No Description Training: No Description Treatment Fidelity: Not provided.	There were no statistically significant differences between tutored group and control. Comprehension (Grades 2&3) Post: -1.27[2] Maintenance: -1.01[2] (Grades 4 to 6) Post: -.94[2] Maintenance: -.94[2]	0%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Ramey (1991)	Examined effectiveness of HOSTS 1988-1990 reading intervention in Seattle on California Test of Achievement test results. Treatment: HOSTS Control: No intervention (Comprehension, Fluency, & Word Recognition)	N^3 = 77 Grades 2 to 3 (*n* = 18) & Grades 4 to 6 (*n* = 59) URBAN (After School)	Comprehension (*n* =2) – California Test of Achievement (CAT) total reading score combining both the primary and intermediate scores. (Reliability not provided)	Paid: No Description Training: No Description Treatment Fidelity: Not provided.	Differences in gain scores were not statistically significant. Comprehension (Grades 2&3) Post: -.99[2] Maintenance: -1.04[2] (Grades 4 to 6) Post: -1.19[2] Maintenance: -1.27[2]	0%
Rimm-Kaufman et al., (1999)	Examined efficacy of 8-month Intergenerational Reading Program on fluency measures. Treatment: Reading for meaning tutoring model Control: Not specified (Comprehension & Word Recognition)	N^1 = 21 Grade 1 AA (*n* =5) C (*n* =7) HC (*n* =7) Other (*n* =2) SUB (During School)	Fluency (*n* =1) - Reading level determined by 1/6 subtests of Observational Survey Word Recognition (*n* =1) – Students read aloud 20 words (Reliability not provided)	N^2 = 21 Most retired and well-trained, most retired, >50% had not been teachers No pay Training: Total time not provided. Treatment Fidelity: Not Provided	No statistically significant differences at end of intervention on fluency measures. Fluency: .43 Word Recognition: - .04	25%
Sank-aranar-ayanan (1998)	Examined effects of 10-month Intergenerational Literacy Project on reading measures; results related to cognitive profiles? Treatment: 3 days/week Control: No tutoring (Fluency & Word Recognition)	N^1 = 63 Grade 1 AA (*n* = 95) URBAN (During School)	Fluency (*n* =2) - Reading two different texts (Reliability not provided)	N^2 = 32 Mostly elderly nonprofessional adults (M = 53;range 23 to 79) No pay Training: 12 hours Treatment Fidelity: Not reported	Statistically significant gains on one fluency measures. No differences on post-test scores. Weakest readers did not benefit as much. Fluency: .22	35%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Vadasy et al. (1997a)	Examined efficacy of 6-month Sound Partners (SP) reading program on multiple reading measures, both immediately after treatment and a year after treatment ended. Treatment: 4 days/week Control: regular practices (Fluency and Word Recognition)	N^1 = 17 Grade 1 Minority (*n* = 17) URBAN (After School)	Fluency (*n* =2) - Analytical Reading Inventory (ARI) and Oral Reading Fluency on two 1-min passages Word Recognition (*n* =3) Wide Range Achievement Test-Revised (WRAT-R) reading subtest; Woodcock Johnson Revised (WJ-R) word Attack subtest, and Dolch raw scores Maintenance included 1 fluency measure and 2 word recognition measures (2 WJ-R subtests) (Reliability not provided)	N^2 = 10 5 parents, AA (*n* = 4), C (*n* = 1) Received $5/hour Training: 5 hours Treatment Fidelity: 5 observations but data not provided.	Moderate effects on word recognition at the end of first and second grade. Moderate effects on fluency in 1st grade and negative effects in 2nd grade. Post Fluency: .43 Word Recognition: .31 .49[1] Maintenance Fluency: - .10 Word Recognition: .53[1]	40%
Vadasy et al. (1997b)	Examined efficacy of 6-month SP reading program on multiple reading measures. Did not examine if effects varied based on how much of a treatment students receive because only 6 students in group. Treatment 1: 4 days/week Control: Regular Practices (Fluency and Word Recognition)	N^1 = 20 Grade 1 Minority (95%) URBAN (After School)	Fluency (*n* =1) – ARI (1st grade level) Word Recognition (*n* =5) WRAT-R (raw) reading subtest; Dolch Word recognition; WJ-R word attack subtest (raw); Bryant Pseudoword; Pseudoword List (Reliability not provided)	N^2 = 10 4 parents Received "nominal hourly wage" Training: 9 hours total, 6 hours before treatment and 3 after treatment Treatment Fidelity: 25% of sessions observed (1-6 scale)	Minimal effects observed. Fluency: .16 Word Recognition: .30	50%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Vadasy et al. (2000)	Evaluated effects of revised 6-month SP reading program on multiple reading measures. Asked if there be improvements if tutors explicitly taught how to sound out words and match stories read to skills learned- Asked if these effects maintained after 1 year. Treatment: 4 days/week Control: Regular Practices (Fluency and Word Recognition)	*N*[1] = 23 Grade 1 Minority (*n* =20) URBAN (After School)	Fluency (*n* =2) - ARI (primary and 1st grade level) Word Recognition (*n* =4) WRAT-R reading subtest; Dolch Word recognition; WJ Word Attack; Bryant Pseudoword (WJ word attack and word identification used at maintenance) Maintenance included 1 fluency measure and 2 word recognition measures (2 WJ-R subtests) (Reliability not provided)	*N*[2] = 8 1 non-parent Paid $5/hour for training and tutoring Training: 14 hours total, 8 hours before treatment and 6 hours after treatment Treatment Fidelity: 25% of sessions observed (1-6 scale)	Tutored group outperformed control group on almost all measures at post-test but these effects did not maintain. Post: Fluency: .49 Word Recognition: .94 1.06[1] Maintenance Fluency: - .09 Word Recognition: .52[1]	50%
Wallach & Wallach (1976)	Examined effectiveness of 8-month tutoring program for children with low reading readiness by using methodologically sound procedures. Treatment: 5 days/week Control: Regular Practices (Fluency and Word Recognition)	*N*[1] = 36 Grade 1 URBAN (During School)	Fluency (*n* =1) - sentence reading – basal Word Recognition (*n* =2) word identification (Reliability not provided)	*N*[2] = 4 Non-education; mid-30's to mid-50's; AA (*n* = 4) Paid by an outside source – commensurate with other school positions. Training: "intensive" Treatment Fidelity: Not provided	Treatment group outperformed both control groups. Following are effects when compared with matched control Fluency: .73 Word Recognition: .68	50%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Structured and Unstructured Programs with single focus on fluency or comprehension *(n* = 6)						
Baker et al. (2000)	Evaluated effectiveness of a 2-year reading program (6 months each year) called Start Making a Reader Today (SMART) on students' reading comprehension scores. Treatment: 2 days/week Control: Regular Practices Unstructured Program (Comprehension)	*N*[1] = 43 Grades 1 to 2 URBAN and RURAL (After School)	Comprehension *(n* =3) (WRMT-R) passage and word comprehension (W score) Fluency *(n* =3) - two first grade passages and one second grade passage Word Recognition *(n* =2) two time-points from WRMT-R word identification subtest (W score) (All measures reliable)	1/3 between 30-45, 1/3 between 45-65, and 1/3 over 65. "literate adults" Businesses paid Training: 50% trained in 1 to 2 hours Treatment Fidelity: Not provided	Statistically significant differences only after 2nd year on all but passage comprehension measures. Comprehension: .33[1] Fluency: .46 Word Recognition: .39[1]	55%
McPeak (2000)	Examined effects of 3 month Reciprocal Teaching intervention on reading comprehension. Treatment: 1 day/week Control: 10 sessions of one-to-one unguided adult-child interactions Structured Program (Comprehension)	*N*[1] =23 Grade 3 C *(n* =36) A *(n* =9) H *(n* =1) SUB (During School)	Comprehension *(n* =1) group administered Gates-MacGinitie Reading Test (GMRT) – standardized score (All measures reliable)	*N*[2] = 27 7 retired community members and 5 from community organization, all HS graduates No pay Training: 4 hours Treatment Fidelity: Provided	Although means for both groups improved over time, two groups' differences were not statistically significant. Comprehension: -.19[1]	70%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Meyer et al. (2002)	Assessed effectiveness of 6-month structured strategy instruction program, called Intergenerational Tutoring (ITP), on reading comprehension scores. Treatment: Senior Citizens led students through structure strategy intervention via a computer. 3 days/week Control: Regular Practices Structured Program (Comprehension)	N^1 = 20 Grade 5 -- RURAL (During School)	Comprehension *(n* =2) – Recall broken into two scores of both total recall and recall of the most important information from long (immediate posttest-generalization) and short (maintenance) passages. (All measures reliable)	N^2 = 11 Highly educated senior citizens (between ages 62 and 80), 5 with PhDs. Paid (personal communication) Training: 12 hours Treatment Fidelity: Monitored but did not provide data.	The differences were statistically significant at maintenance, but not on immediate post test; however, post-test passages were much longer than those used during treatment. Comprehension Post Test: .52 Maintenance: .80	80%
Vadasy et al. (2002)	Examined the effectiveness of an 8-month adapted Sound Partners program (called Thinking Partners, TP) on several reading measures. Treatment: 4 days/week Control: Regular practices but tutoring often occurred during regular class reading time. Structured Program (Comprehension)	N^1 = 10 Grade 2 C *(n* =3) A *(n* =6) O *(n* = 1) URBAN (During School)	Fluency *(n* =1) – 1 min score for 2nd grade reading Word Recognition *(n* =4): WRMT-R word identification and word attack subtests; Test of Word Reading Efficiency (TOWRE); and WRAT-R reading subtest (Reliability not provided)	40% did not have previous tutoring experience Paid small stipend Training: 4 hours Treatment Fidelity: Yes	There were no statistically significant differences between the treatment and control groups. Fluency: -.05 Word Recognition: .32[1]	45%

Table 2 continued. Experimental and Quasi-Experimental One-to-One Tutoring Interventions Led by Nonprofessional Tutors

Study	Purpose/ Research Questions/ (Focus of Instruction)	Students/ Description/ Setting	Dependent Variables	Tutors (Experience)/ Compensation/Training/ Treatment Fidelity	Results (Effect Size)	Quality Score
Weiss et al. (1989)	Evaluated the effectiveness of a 3-month adult tutoring program (offering more opportunities to read) on fluency measures. Treatment: 4 days/week Control: Regular practices Unstructured Program (Fluency)	N^1 = 17 Grades 3 to 5 SUB (During School)	Fluency *(n* =2) – Basic Achievement Skills Individual Screener reading subtest raw scores; CBM reading measure based on Holt Reading Series. (All measures reliable)	N^2 = 12 Mostly senior citizens *(n* = 8) without professional teaching experience *(n* = 8) No pay Training: 5 hours Treatment Fidelity: Not provided	No statistically significant differences between the groups. Fluency: -.20	40%
Wilfong (2006)	Determined if a 3-month reading intervention that used poetry (and repeated readings) would improve reading skills of struggling 3rd grade readers. Treatment: 1 day/week Control: Regular Practices Unstructured Program (Fluency)	N^1 = 36 Grade 3 C *(n* = 85) RURAL (During School)	Fluency *(n* =1) a CBM that provided words correct per minute (Reliability not provided)	N^2 = 6 3 non-education majors Most were not paid although one was paid by OhioReads Training: 2 hours Treatment Fidelity: Not provided	The gains made by the treatment group were statistically significant when compared to the treatment group but when post-test means were used to calculate effect sizes, results indicated an overall negative effect. Fluency: - 2.1	25%

Notes. Dashes indicate that data were not reported. N^1 = total number of students in the treatment group in which effect sizes were calculated for, treatment and control; Kinder. = Kindergarten; AA = African American; AI = American Indian; A = Asian; C = Caucasian; HC = Haitian Creole; H = Hispanic; O = Other. Location: SUB = suburban; N^2 = total number of adult tutors working with the treatment group; participants in the study, [1] indicates ES calculated based on standard scores.

Table 3. Overall Percent of Quality Indicators Met (10-point scale) and Individual Quality Indicator Results

Study	Total Quality Score	Total Percent	Equiv. Mort. Rates	Pretest Equiv. of Ss	Equiv. of Ts	Instructional Conditions Described	Tutor Training Described	TX Fidelity	Control for Hawthorne Effects	DV(s) Reliable	Control for Tutor Effects	Random
Al Otaiba et al. (2005)	8	80%	1	1	0	1	.5	.5	1	1	1	1
Burns et al. (2004)	2.5	25%	0	0	0	0	.5	0	0	1	0	1
McGrady (1983)	1.5	15%	0	0	0	0	.5	0	0	1	0	0
Morris et al. (1990)	3	30%	0	1	0	0	1	0	0	0	0	1
Nielson (1991)	3	30%	0	0	0	0	0	0	1	1	0	1
Ramey (1990)	0	0%	0	0	0	0	0	0	0	0	0	0
Ramey (1991)	0	0%	0	0	0	0	0	0	0	0	0	0
Rimm-Kaufman et al. (1999)	2.5	25%	0	1	0	0	.5	0	0	0	0	1
Sankaranara-yanan (1998)	3.5	35%	0	1	0	.5	1	0	0	0	0	1
Vadasy et al. (1997a)	4	40%	1	0	0	1	1	0	0	0	0	1
Vadasy et al. (1997b)	5	50%	0	0	1	1	1	1	0	0	0	1
Vadasy et al. (2000)	5	50%	0	1	0	1	1	1	0	0	0	1

Table 3 continued. Overall Percent of Quality Indicators Met (10-point scale) and Individual Quality Indicator Results

Study	Total Quality Score	Total Percent	Equiv. Mort. Rates	Pretest Equiv. of Ss	Equiv. of Ts	Instructional Conditions Described	Tutor Training Described	TX Fidelity	Control for Hawthorne Effects	DV(s) Reliable	Control for Tutor Effects	Random
Wallach & Wallach (1976)	5	50%	0	1	1	1	1	0	0	0	0	1
Baker et al. (2000)	5.5	55%	1	1	0	1	.5	0	0	1	0	1
McPeak (2000)	7	70%	0	0	0	1	1	1	1	1	1	1
Meyer et al. (2002)	8	80%	0	1	1	1	1	0	1	1	1	1
Vadasy et al. (2002)	4.5	45%	0	1	0	1	.5	1	1	0	0	0
Weiss et al. (1989)	4	40%	0	0	0	1	1	0	0	1	0	1
Wilfong (2006)	2.5	25%	0	0	1	.5	1	0	0	0	0	0

Overall quality of studies. Before examining the outcomes, I first address the overall quality of the studies included in this review. As can be seen in Table 3, approximately half of the studies met criteria for: pretest equivalencies, describing the instructional conditions and tutor training, reliability for dependent variables, and random assignment of students to treatment and control conditions. However, the majority of studies (more than 70%) did not meet the following criteria: equivalent mortality rates of experimental and control participants, pretest equivalence of the tutors (reporting more than just the total number of tutors or a brief description of their background), treatment fidelity, control for Hawthorne effects, and control for tutor effects.

The findings for the nonprofessional adult-led literacy interventions are presented next. Results are provided for each of the three treatment categories: structured programs with multiple foci, structured programs with two foci, and structured and unstructured programs with a single focus. For each specific treatment (e.g., TAILS; Al Otaiba et al., 2004) within each category (e.g., structured programs with multiple foci) results are described using a similar format. First, student participants and treatment (e.g., TAILS) details are provided. Second, tutor descriptions and outcome measures (e.g., word recognition at post-test and maintenance) are presented. Finally, the effect sizes and study quality are summarized. The effect sizes are provided first for raw and standard scores. The final section, in contrast, summarizes effects across all measures (reading comprehension, oral reading fluency, and word recognition) regardless of type of score used.

Summaries are provided at the end of each treatment category (e.g., structured programs with multiple foci) and at the end of the results section. The format for

summaries is similar to the format used throughout the results section (e.g., description of students and treatment details). In the final summary section, I summarize effect sizes across variables (e.g., type of outcome measure) that other researchers have found influenced the overall effectiveness of tutoring programs: subject matter, program structure, outcome measures, treatment duration, response to treatment, and tutor type/training (Cohen et al., 1982; Erlbaum et al., 2000; Wasik, 1998; Wasik & Slavin, 1993; Wasik et al., 2002). The first category of treatments descried next is structured programs with multiple foci.

Structured programs: Focus on comprehension, fluency, and word recognition. Fourteen studies examined the effectiveness of nonprofessional adult tutors delivering instruction with multiple foci in a structured format (see Table 2). As noted earlier, all of the studies evaluated effects of reading instruction on student achievement. The results for the structured programs are separated for those programs in which comprehension, fluency, and word recognition were the focus of instruction and those in which only two out of the three were the focus of instruction.

Seven out of the 14 studies used reading treatments that focused on comprehension, fluency, and word recognition (Al Otaiba et al., 2005; Burns et al., 2004; McGrady, 1983; Morris et al., 1990; Nielson, 1991; Ramey, 1990, 1991). The five treatments used in the six studies (Tutor Assisted Intensive Learning Strategies, TAILS; Helping One Student to Succeed, HOSTS; Houghton Mifflin Tutorials; Howard Street tutoring program; Ginn Reading Series) all used multiple activities during the, 4 to 5 times weekly, and from 15 to 60 mins per tutoring sessions (median was 30 mins). The majority of the programs were in effect for 8 to 10 months (TAILS; Houghton Mifflin

Tutorials; Ginn Reading Series) with one program lasting 2 years (Howard Street tutoring program) and another program's length unknown (HOSTS).

Although all of the programs had multiple components, the students' specific activities, dependent variables, and tutors varied from study to study. In the following section the four treatments are described using a similar format: (a) treatment group and program materials; (b) tutor descriptions and outcome measures; and (c) effect sizes and study quality. A summary of the four programs is provided at the end of this section. As stated earlier, only results related to reading comprehension, oral reading fluency, and word recognition were presented.

***TAILS*.** The tutors in Al Otaiba et al.'s (2005) study taught 24 kindergarten students in an urban school during the school day. During each session, the tutors first focused on phonological awareness activities through a variety of games. Next, the tutors helped the students decode words through direct instruction and brief speed games. Both of these techniques were taken from the Peer Assisted Learning Strategies program (PALS; see Al Otaiba et al.). Finally, the tutors used books to help students improve their vocabulary and reading comprehension. The tutors not only read but also employed a wide variety of questioning techniques.

Seven out of the 12 TAILS tutors had no previous tutoring experience and were between 30 and 65 years old (Al Otaiba et al., 2005). The tutors received a small stipend for participating and received 13 hours of training. Al Otaiba et al. evaluated effects of the treatment on student reading comprehension and word recognition using three subtests of the Woodcock Reading Mastery Test Revised (WRMT-R). Standardized scores used to evaluate effects of program. All measures were found reliable.

Providing the TAILS treatment had a moderate to large overall effect on improving passage comprehension and word recognition (see Table 2). The effect size for the reading comprehension measure was .73 and the averaged effect size for the word recognition measures was .55. Confidence can be placed in these findings as the researchers met the majority of the quality indicators (see Table 3).

***HOSTS*.** The tutors in the HOSTS treatments (Burns et al., 2004; Ramey, 1990, 1991) used specific web-based activities to teach 239 students in grades two through six. The schools in the studies were located in a variety of areas (rural, urban, and suburban) and the actual teaching occurred both during and after the school day. Burns et al. provided a description of the HOSTS tutoring program. Students taught using the HOST language arts program received personalized lessons generated on computers. Each program was individualized based on the student's needs and were aligned with state and local reading standards (see Burns et al.). The activities included tutors providing students direct instruction in the following areas: "decoding skills, vocabulary, literature, writing, and critical thinking" (Burns et al., p. 89).

Very little information is known about the HOSTS tutors. Burns et al. (2004) stated that the tutors were community nonprofessional adults that received no pay, two hours of training, and constant supervision. Ramey (1990, 1991) did not provide any tutor description. Burns et al. measured effect of the HOSTS program on one informal reading comprehension measure (Gray Oral Reading Test; GORT) and two fluency measures (GORT and the Dynamic Indicator of Basic Early Literacy Skills; DIBELS; see Burns, 2004). All measures were reliable. Ramey evaluated effects of HOSTS program on two

standardized reading comprehension measures (California Test of Achievement, CAT; see Ramey, 1990, 1991). Reliability of assessment measures was not provided.

The HOSTS treatment had varied effects across the three studies. The effects were small to moderate in Burns et al.'s (2004) study with an effect of .50 for comprehension and .37 for fluency. The effects were negative in the Ramey studies (1990, 1991) although the treatment groups were considerably smaller not equivalent at pre-test). The average effect of the HOSTS program on improving second and third graders reading comprehension was -1.13 *(n* = 2; range -1.27 to -.99) at post-test and -1.03 *(n* = 2; range -1.04 to -1.01) during maintenance (Ramey, 1990, 1991). The average effect on improving fourth through sixth graders' reading comprehension was -1.07 *(n* = 2; range -1.19 to -.94) at post-test and -1.14 *(n* = 2; range -1.27 to -1.01) at maintenance. Caution must be used when interpreting results as all three studies failed to meet several of the quality indicators (see Table 3).

***Houghton Mifflin tutorial*.** The Houghton Mifflin tutorial (McGrady, 1983) was not described in much detail, however available data supported its placement within the multi-focus category. The tutors used materials from the Houghton Mifflin Comprehension and Word Attack (CAWA) Skills Books (McGrady, 1983) to tutor 35 fourth grade students. Tutoring occurred during the school day in a rural school setting. The CAWA skills books contained activities designed to improve students' oral reading fluency, reading comprehension, and word recognition (McGrady).

Thirty-six tutors participated in the study, many with no tutoring experience *(n* = 14; McGrady, 1983). The tutors were paid a minimal hourly wage and received 15 to 20 hours of training. McGrady evaluated the effects of the program on one standardized

reading comprehension measure (Iowa Test of Basic Skills). The reliability for this measure was provided.

The Houghton Mifflin tutorial yielded a negative effect on the fourth grade students reading comprehension. The effect size was -.35 at post test (see Table 2). The results need to be tempered because the overall quality of the study was low. The only quality indicators met were describing tutor training, reporting no ceiling or floor effects, and providing reliability data for dependent variables (see Table 3).

Howard Street Tutoring program. The tutors in Morris et al.'s (1990) study taught 30 second and third grade students using Howard Street Tutoring materials. The study took place over two years in an urban after-school setting. Over the course of the 2-year study, the nonprofessional adults worked individually with the students for 1 hour per day, 4 days per week. The tutors first read with the students, with a focus on comprehension and word recognition (Morris et al.). Next, the tutors engaged tutees in word recognition activities by using games and sorting activities. The students then wrote short stories with little emphasis being placed on correct spelling until later edits. The tutors then listened to students read books that the child could easily read. Finally, the tutors read to the child.

Although the total number of tutors used in the study was not provided, 14 of the tutors had no experience tutoring (Morris et al., 1990). The tutors were paid a minimal wage for their participation and also received minimal training. Morris et al. evaluated the effects of the Howard Street Tutoring program on one oral reading fluency measure and three word recognition outcome measures (see Table 2). Reliability was not provided on any measure.

The Howard Street Tutoring program had a moderate to large effect on improving second and third grade students oral reading fluency and word recognition (see Table 2). The effect size for oral reading fluency measure was .69 and the averaged effect size for the word recognition measures was .46. The results must be tempered as only 4 of the 11 quality indicators were met.

Ginn Reading Series tutorial. In Nielson's (1991) study, 29 third grade students (25% with behavioral problems) were taught reading strategies using the Ginn Reading Series tutorial. The 29 students were divided into two groups tutored by parents *(n* = 14) or nonprofessional adults *(n* = 15). Results are presented for all 29 students as a comparison was not made between the nonprofessional adults only and a control group. The tutors read stories (with a focus on vocabulary and oral reading fluency and improving reading comprehension) and reviewed vocabulary through precision teaching techniques (Nielson, 1991). Students received points for reading words correctly on the first try and were then able to exchange these points every month for small articles (e.g., pencils, books, and so forth).

Very little information is provided on the nonprofessional adults who participated in Nielson's (1991) study. Nonprofessional adults received a small compensation (free school lunch) for participating. Nielson evaluated the effect of the program on one standardized comprehension measure (Stanford Achievement Test; SAT). Reliability was provided.

The Ginn Reading Series tutorial had a large effect immediately after treatment and a negative effect at maintenance. The effect size on the SAT comprehension measure was 1.31 at post test and -1.91 during maintenance. The overall quality of this study

suggests that interpretation of the results must be tempered. In all, Nielson's (1991) study met 36% of the quality indicators, omitting information related to several important areas such as describing differences in mortality rates and pre-treatment equivalence.

Summary of structured programs: Focus on comprehension, fluency, and word recognition. In summary, nonprofessional tutors using a structured program with multiple foci had varied effects on improving 357 Kindergarten through sixth grade students reading comprehension, oral reading fluency, and word recognition. Five treatments were described in seven studies (TAILS, Al Otaiba et al., 2005; HOSTS, Burns et al., 2004, Ramey, 1990, 1991; Houghton Mifflin Tutorials, McGrady, 1983; Howard Street Tutoring program, Morris et al., 1990; Ginn Reading Series, Nielson, 1991). Studies occurred in a variety of locations (rural, urban, and suburban) both during and after school. Treatment duration ranged from 8 (Al Otaiba et al.) to 9 (Nielson, 1991) to 10 (McGrady, 1983), to 12 months (two consecutive 6-month studies, Morris et al., 1990).

Less information can be surmised about nonprofessional adults used in the studies. Of those that reported volunteer data (Al Otaiba et al., 2005; McGrady, 1983; Morris et al., 1990), at least 62 nonprofessional adults were used to deliver instruction. In four of the treatments, nonprofessional adults received a small stipend (money, free lunch) for participating (Al Otaiba et al.; McGrady, 1983; Morris et al., 1990; Nielson, 1991). Four of the studies described tutor training with total time in training ranging from minimal (Morris et al.) to 2 hours (Burns et al., 2004) to 13 hours (Al Otaiba et al.) to 15 to 20 hours (McGrady, 1983).

The treatments had a wide range of effects across the seven studies (see Table 2). Most studies reported effects of the program on standardized outcome measures (Al Otaiba et al.; McGrady, 1983; Nielson, 1991; Ramey, 1990, 1991). The standardized outcome measures were collected during posttest (Al Otaiba et al.; McGrady) as well as posttest and maintenance (Nielson; Ramey, 1990, 1991). On the standardized comprehension measures, taken at posttest, the effect sizes ranged from 1.31 (Nielson) to .73 (Al Otaiba et al.) to -.35 (McGrady) to -1.27 (Ramey, 1990). These effects diminished during maintenance. On the standardized comprehension measures, taken during maintenance, the effect sizes ranged from -1.91 (Nielson) to -1.27 (Ramey, 1991). Al Otaiba et al. also evaluated the effects of the program on word recognition. The effect size of the treatment on this outcome variable was .55.

Others reported effects of treatments using raw scores (Burns et al., 2004; Morris et al., 1990). Burns et al. evaluated the effects of the HOSTS program on three informal outcome measures assessing comprehension and fluency. Results were slightly higher for comprehension (ES = .50) when compared to fluency (ES = 37). Morris et al. assessed effects of Howard Street Tutoring program on four informal outcome measures assessing fluency and word recognition. Effect sizes were larger for the fluency measure (ES = .69) when compared to the three word recognition measures (ES = .46).

Combining raw and standard score data, the average effect size on comprehension was -.28 at post-test (range -1.27 to 1.31) and -1.23 at maintenance (range -1.91 to -.94). The average effect size on word recognition was .51 at post-test (range .46 to .55). The average effect size on fluency was .53 at post-test (range .37 to .69).

The overall results must be tempered due to the overall low quality of studies. The majority of the studies were randomized control trials (4/7; 57%), described tutor training (4/7, 57%), established that there were no ceiling or floor effects (5/7; 71%), and established dependent variable reliability (4/7; 57%). However, Al Otaiba et al.'s (2005) study was the only study that met more than half of the quality indicators (only received a 0 for pretest equivalence of tutors and control for tutor effects). The quality indicators not met by the remaining 6 studies are the following: equivalent mortality rates, pretest equivalence of tutors, instructional conditions described, treatment fidelity, and control for tutor effects (Burns et al., 2004; McGrady, 1983; Morris et al., 1990; Nielson, 1991; Ramey, 1990, 1991). This raises questions about whether the treatment actually occurred as described. It also raises questions about whether the procedures could be replicated.

The remaining studies found in the first section of Table 2 are described next. These studies evaluated the effects of structured programs with a focus on word recognition and comprehension (Rimm-Kaufman et al., 1999) or word recognition and fluency (Intergenerational Literacy Project; Sankaranarayanan, 1998; Sound Partners, Vadasy et al., 1997a, 1997b, 2000; Wallachs' tutoring program; Wallach & Wallach, 1976). The programs are described in the next section using the same format used for the first group of studies. First data related to student participants and treatment description is provided. Next, tutor participants and outcome measures are summarized. Finally, effect sizes and study quality is highlighted. After the four treatments are given separate descriptions, an overall summary is provided.

Structured programs: Focus on word recognition and comprehension or word recognition and fluency. Four treatments are described in the following section.

First, Rimm-Kaufman et al. (1999) evaluated the effectiveness of an 8-month Intergenerational Reading Program that focused on reading for meaning. Second, Sankaranarayanan (1998) examined effects of 10-month Intergenerational Literacy Project. Third, Vadasy et al. (1997a, 1997b, 2000) assessed effects of the 6-month reading program called Sound Partners (SP). Fourth, Wallach and Wallach (1976) examined the effects of an 8-month reading program. Four of the of the reading programs had activities that primarily focused on fluency and word recognition (exception is Rimm-Kauffman et al.) and evaluated the effects on students' fluency and word recognition.

Intergenerational Reading Program. The first study examined the effect of the Intergenerational Reading Program that focused on fluency and comprehension (Rimm-Kaufman et al., 1999). Twenty-one first-graders were provided one-on-one tutoring by adult volunteer tutors (Rimm-Kaufman et al.). The study took place during the school day in a suburban elementary school. The treatment occurred three times a week for 45 mins each session and lasted 8 months (Rimm-Kaufman et al.). The treatment included activities designed to increase "reading for meaning versus simply decoding" (Rimm-Kaufman et al., p. 144). The tutors first focused on helping students make connections between pictures and words. Tutors then taught students phonetics through stories, games, and other activities.

There were 21 tutors that provided tutoring to each of the 21 students in the treatment condition (Rimm-Kaufman et al., 1999). Most tutors were retired and more than half did not have teaching experience. The tutors did not receive any payment for their participation and were identified as "well trained" (p. 144). Rimm-Kaufman et al.

evaluated the effect of the program on an informal reading fluency measure (Observational Study) and word recognition measure. Raw scores were reported. Reliability was not provided for these measures.

The treatment provided by the tutors in the Rimm-Kaufman et al. (1999) study had moderate effects on improving reading fluency and no effects on improving word recognition (see Table 2). The effect size for the fluency measure was .43, whereas the effect size for the word recognition measure was -.04. Caution must be taken when interpreting results as only 23% of the 11 quality indicators were met.

The next three treatments are the remaining three structured treatments with multiple foci. These treatments are all similar in that they primarily focused on activities related to fluency and word recognition. The programs are the Intergenerational Literacy Project (Sankaranrayanan, 1998), Sound Partners (Vadasy et al., 1997a, 1997b, 2000), and the Wallach and Wallach tutoring program (Wallach & Wallach, 1976).

Intergenerational Literacy Project. The tutors in Sankaranrayanan's (1998) study taught 63 first grade students in an urban school during the school day. The 10-month Intergenerational Literacy Project (ILP) occurred 3 to 4 times per week and lasted approximately 35 mins. Tutors read familiar books, taught phonics through various games and sorting activities, elicited writing samples that reinforced lessons, and introduced new books.

There were 32 volunteer tutors in the Intergenerational Literacy Project (Sankaranrayanan, 1998). The majority of the tutors were senior citizens with a mean age of 53 (range 23 to 79). The tutors did not receive pay for participating and were required to attend 12 hours of training. Sankaranrayanan (1998) evaluated the effects of the

treatment on two fluency measures (informal assessments). Raw scores were reported. Reliability was not provided.

The effect of the Intergenerational Literacy Project on improving reading fluency and word recognition was small. The effect size for the two oral fluency measures was .22. The overall quality of the study was also low, with only 32% of the quality indicators met. The next three studies in Table 2 comprise a similar treatment, Sound Partners.

Sound Partners. The effects of the Sound Partners tutoring program were evaluated in three separate studies (Vadasy et al., 1997a, 1997b, 2000). In all, 60 1st-grade students were taught using the Sound Partners structured tutoring program. The three 6-month studies occurred in urban elementary schools and were conducted after the school day. The Sound Partners treatment is comprised of specific activities related to letter names and sounds, rhyming, segmentation, word families, spelling, writing using invented spelling (tutors corrected by end of session), and reading books. Each session is 30 mins long and occurs 4 days per week.

Twenty-eight tutors were used in the three different studies (Vadasy et al., 1997a, 1997b, 2000). Many of the tutors were parents of other elementary students (16/28; 57%). The tutors received some sort of reimbursement for participating in the study and were required to attend from 5 to 15 hours of training (Vadasy et al., 1997a, 1997b, 2000). The total time in training increased as the studies were conducted because the researchers saw a need to better equip their tutors in order to obtain better effects (Vadasy et al., 2000). Across all three studies, Vadasy et al. evaluated the effects of the Sound Partners on fluency (Analytical Reading Inventory, ARI) and word recognition outcome measures (one subtest of Wide Range Achievement Test-Revised, WRAT-R; two word attack

subtests of Woodcock Johnson Revised, WJ-R; and Dolch word recognition). In the last study, Vadasy et al. also evaluated word recognition using the Bryan Pseudoword list. Effects were drawn from standardized scores for the following measures: WRAT-R (Vadasy et al., 1997a, 2000) and WJ-R (Vadasy et al., 1997a, 2000). Reliability was not provided for any measures.

Providing the Sound Partners reading program had small to moderate effects on improving fluency and word recognition, with marked diminishing effects on fluency measures at maintenance (see Table 2). The effect sizes, using raw scores, are described first. The average effect size for fluency was .36 (range .16 to .49) at post-test and -.10 (range -.09 to -.10) at maintenance. The average effect size for word recognition was .52 (range .30 to .94) at post-test. Only standardized scores were used to calculate effects of program on word recognition at maintenance.

The effect sizes for word recognition improved slightly when calculated with measures that used standardized scores, (Vadasy et al., 1997a, 2000). When standard scores were used to calculate effect sizes, the effect on improving word recognition was .78 at post-test (compared to .52 when raw scores were used to calculate effect sizes) and .53 at maintenance. There should be moderate confidence in these results as each study met at least 6 of the 11 (55%) quality indicators (see Table 3). The Wallach and Wallach tutoring program (Wallach & Wallach, 1976) is the final structured treatment with multiple foci presented in this section.

Wallach and Wallach tutoring program. In the Wallach and Wallach (1976) study, tutors instructed 36 first-grade students in an attempt to increase their fluency and word recognition. The 8-month study occurred during the school day in an urban

elementary school. During each 30 min session (delivered 5 days per week) the tutors progressed students through phoneme-related activities. The instruction occurred at three different levels (Wallach & Wallach). First basic letter sound relations were taught. Next, students were taught how to connect letters and distinguish between words. Finally, they were taught to apply those skills in their classroom reading materials.

Four African American women delivered the Wallach and Wallach tutoring program (Wallach & Wallach, 1976). All of the tutors were non-education majors between 30 and 50 years old. The tutors were paid for their involvement and received intensive training (Wallach & Wallach). It appears these tutors were paid more than the other tutors in the reviewed studies, receiving money that was commensurate with other paid school positions. Wallach and Wallach evaluated the effectiveness of the program on fluency (one informal measure) and word recognition (one informal and one standardized measure). Reliability of measures was not provided.

The Wallach and Wallach (1976) treatment had large effects on first-grade students' fluency and word recognition. The effect size for the measures ranged from .68 (word recognition) to .73 (fluency). Moderate confidence can be placed in these results as just over half of the quality indicators were met (55%; see Table 3).

Summary of structured programs: Focus on word recognition and comprehension or word recognition and fluency. In summary, nonprofessional tutors using structured programs that focused on fluency and comprehension or fluency and word recognition had varied effects on improving 180 first-grade students oral reading fluency and word recognition. There were four treatments described in six studies (Rimm-Kaufman, 1999; Intergenerational Literacy Project, Sankaranrayanan, 1998;

Sound Partners, Vadasy et al., 1997a, 1997b, 2000; Wallach & Wallach, 1976). The studies occurred in urban and suburban locations both during and after school. The studies lasted 6 (Vadasy et al., 1997a, 1997b, 2000), 8 (Rimm-Kaufman et al., 1999; Wallach & Wallach, 1976), or 10 months (Sankaranrayanan, 1998) with treatment occurring between three to four times per week.

There were 85 tutors in the six studies. Two of the treatments provided tutors with a small stipend (SP; Vadasy et al., 1997a, 1997b, 2000), whereas the other provided pay commensurate with paraprofessionals employed by the school (Wallach & Wallach, 1976). Two of the programs, evaluated in four studies, provided total volunteer training time (Sankaranarayanan, 1998; Vadasy et al., 1997a, 1997b, 2000). The average amount of time in training across these studies was 10 hours (range 9 to 15). Vadasy et al. (1997b, 2000) gradually increased the amount of tutor training provided in the three studies evaluating Sound Partners' effectiveness.

All of the studies examined the effectiveness of treatment on fluency and word recognition outcome measures. Treatments had small to moderate effects at post-treatment with effects diminishing during maintenance. Most studies reported effects using raw scores (Sankaranrayanan, 1998; Vadasy et al., 1997b, Wallach & Wallach, 1976). Studies assessed effectiveness at post-test (Rimm-Kaufman et al., 1999; Sankaranrayanan; Vadasy et al., 1997; Wallach & Wallach, 1976) and maintenance (Vadasy et al., 1997a, 2000). The average effect size (using measures with raw scores) for fluency was .44 (range .22 to .73) at post-test and -.10 (no range) at maintenance. The average effect size for word recognition was .31 (range -.04 to .68). Two studies also included standardized scores when they calculated effects on word recognition (Vadasy

et al., 1997a, 2000). The average effect size on word recognition (using standardizes scores) was .78 at post-test and .53 at maintenance.

Combining both standard and raw scores, the average effect size on fluency was .41 at post-test *(n* = 6; range .16 to .73) and -.10 at maintenance *(n* = 2; no range). The average effect size on word recognition was .47 at post-test *(n* = 8; range -.04 to 1.06) and .53 at maintenance *(n* = 2; no range).

Moderate confidence can be placed in these findings based on the number of quality indicators met. First, all of the studies were randomized control trials (see Table 3). Second more than half of the studies established pretest equivalence of students, described instructional conditions, described tutor training, and had no ceiling or floor effects. Some caution should be taken in interpretation as the majority of these studies did not establish equivalent mortality rates, pretest equivalence of the tutors, treatment fidelity, control for Hawthorne effects, dependent variable reliability, and control for tutor effects (see Table 2).

The final six treatments described in the next section are those structured and unstructured programs delivered by nonprofessional adult tutors with focus on one (Baker et al., 2000; McPeak, 2000; Meyer et al., 2002; Vadasy et al., 2002; Weiss et al., 1989) or two activities (Wilfong, 2006). These treatments tended to be less formal and less intense (fewer sessions per week or briefer tutoring sessions). Results are described in a similar format as was used in this section, with a summary provided at the end.

Structured and unstructured programs: Single focus on fluency or comprehension. Six treatments are described in six studies. First, Baker et al. (2000) evaluated the effectiveness of a 2-year Start Making a Reader Today program (SMART).

Next, McPeak (2000) reported the effects of a 3-month reciprocal teaching program. Third, Meyer et al. (2002) assessed effects of a 6-month structured strategy instruction program called Intergenerational Tutoring. Fourth, Vadasy et al. (2002) examined the effects of an 8-month Thinking Partners (TP) program (variation of the SP program). Fifth, Weiss et al. (1989) evaluated effectiveness of 3-month Paired Reading program that provided more opportunities for students to read. Finally, Wilfong (2006) examined effects of a 3-month Poetry Academy. All of the reading programs evaluated the effects of the treatment on reading comprehension (Baker et al., McPeak, Meyer et al., Vadasy et al.), oral reading fluency (Weiss et al., 1989), or comprehension and fluency (Wilfong, 2006).

Start Making a Reader Today. The tutors in the SMART reading program taught 43 first- and second-grade students in urban and rural schools (Baker et al., 2000). The tutoring occurred immediately after the school day ended and lasted for 6-months during the students first grade year and 6-months during their second grade year. The SMART tutoring program was unstructured with tutors conducting four main activities during each 30-min (two times per week) session. First, the volunteer read to each tutee. Second, the volunteer and tutee read in unison (choral reading). Third, the volunteer read a short section and the tutee repeated. Fourth, the volunteer asked students questions to increase comprehension skills (Baker et al.)

The majority of the tutors used in the SMART program were nonprofessional adults from the business community. No information was provided on total numbers or tutoring experience but information related to tutor pay and tutor demographics was provided. First, Baker et al. (2000) reported that tutors were reimbursed by businesses.

Second, Baker et al. identified tutors as adults between the ages of 30 to 45 (33%), 45 to 65 (33%), and over 65 (33%). Third, some tutor training occurred with 50 percent of tutors receiving between one to two hours of training.

Effects of program were assessed on reading comprehension, fluency, and word recognition (see Table 2). Comprehension was assessed using three assessments (standardized scores on 3 WRMT-R subtests). Fluency was assessed using three assessments (raw scores on first- and second-grade reading passages). Word recognition was assessed using two assessments (standardized scores on 2 WRMT-R subtests). All measures were reliable.

Providing the SMART program had small to moderate effects on comprehension, fluency, and word recognition (see Table 2). All measures were taken at post-test. The effect sizes ranged from .33 (reading comprehension) to .39 (word recognition) to .46 (fluency). Moderate confidence can be placed in these findings as the study met half of the quality indicators (see Table 3). The quality indicators not met were pretest equivalence of tutors, treatment fidelity, no ceiling or floor effects, and controlling for Hawthorne and tutor effects. This raises questions about whether the treatment occurred as stated and whether effects were truly a result of the treatment. The next treatment presented in Table 2 is the 3-month reading program evaluated by McPeak (2000).

***McPeak*.** The tutors in McPeak's (2000) study taught 23 third-grade students in a suburban elementary school during the school day (total program lasted approximately 3 months). The sessions involved tutors using a collaborative reciprocal teaching approach during each 45 min tutoring session (occurred once per week). The activities were structured and primarily focused on reading comprehension (McPeak). The tutors taught

students how to make predictions, generate questions, summarize, and clarify unknown vocabulary and ideas.

There were 27 tutors in all. Seven were retired community members and 4 were from community organizations. All of the tutors were high school graduates. The tutors did not receive payment for participating and were required to attend 4 hours of training. McPeak (2000) evaluated the effects of the tutoring on reading comprehension. All measures were reliable.

Although the two groups both improved over time, the tutoring program in the McPeak's (2000) study did not have an effect on improving reading comprehension (see Table 2). The effect size for comprehension was -.19. This outcome may have occurred because the two groups were not equivalent at pre-test. Confidence can be placed in these findings as the study met 8 out of the 11 quality indicators (73%). The quality indicators not met were establishing equivalent mortality rates and pre-test equivalence of students and tutors (see Table 3). The next study evaluated the effects of a tutoring program that used retired senior citizens to deliver instruction via the computer (Meyer et al., 2002).

Intergenerational Tutoring Program (ITP). Tutors in Meyer et al.'s (2002) study taught 20 fifth-grade students in a rural elementary school via computers. The Intergenerational Tutoring Program (ITP) was structured, occurred 60 mins each week (3 times per week), and lasted six months (Meyer et al.). The activities involved teaching students *structure strategies* aimed at improving their reading comprehension (Meyer et al.). The lessons involved teaching the following plans to improve total recall: using comparisons, identifying problems and solutions, adding cause and effects, using the plan

strategy, and utilizing the sequence plan. Students received lessons from their tutors via email and were prompted to show their understanding through writing.

Eleven tutors delivered the ITP. All of the tutors were senior citizens (between 62 and 80 years old) and were *highly educated* (Meyer et al., 2002). The tutors received some compensation for their participation and were required to participate in 12 hours of training. Meyer et al. (2002) evaluated the effects of the program on two comprehension outcome measures. Both outcome measures were taken from students' writing and evaluated for total recall and recall of most important information (raw scores; Meyer et al.). The reading samples at post-test and maintenance were different. The post-test texts were longer than those used during instruction, so during maintenance the researchers shortened the text. Reliability was provided for both measures.

The ITP program had moderate to large effects on students' reading comprehension. The magnitude of the effect varied depending on when students were tested (or possibly passage length). On the longer text, given immediately after treatment, the effect size was .52. On the shorter text, given at maintenance (over 2 months after the treatment ended), the effect size was .80. Confidence can be placed in these findings as Meyer et al. (2000) met 9 out of the 11 quality indicators (82%). The indicators not met were establishing equivalent mortality rates and recording treatment fidelity (although they monitored fidelity, they did not provide fidelity data; see Table 2 and 3). The next treatment described in Table 2 was a variation of the SP program, delivered to 2nd grade students (Vadasy et al., 2002).

Thinking Partners (TP). In their most recent study, Vadasy et al. (2002) evaluated the effects of extending the Sound Partners (SP) programs into the second

grade with a program called Thinking Partners (TP). The activities in the SP program focused on fluency and word recognition skills whereas the TP program focused on reading comprehension. Vadasy et al. (2002) included three treatments in the study (SP only; TP only; and SP and TP combined), but the control group only met criteria for comparing to the TP-only group. Therefore the effects of the other treatments are not included in this review.

The TP tutoring program was a highly structured program modeled after the SP program (Vadasy et al., 2002). In the current study, tutors taught ten 2nd-grade students in an urban elementary school during the school day. As with the SP tutoring, TP tutoring occurred 30 mins per day, 4 days per week. The first two weeks of the 6-month program were spent reviewing letter-sound relations and decoding skills. The rest of the time was spent with tutors teaching reading comprehension using five strategies: keep track, understand words, make connections, think ahead, and make questions. Tutors used scripted lessons and spent at least four days on each strategy before proceeding on to the next.

Vadasy et al. (2002) provided little information about their tutors. At least 40 percent of the tutors did not have previous tutoring experience and were paid a small stipend for their participation. They were also required to attend four hours of training. Vadasy et al. (2002) evaluated the effects of TP on student reading comprehension but only percentages were provided. Vadasy et al. (2002) also evaluated effects on fluency using one 1-min test and word recognition using two WRMT-R subtests, the Test of Word Reading Efficiency (TOWRE) and one WRAT-R subtest (all four word recognition measures used standard scores). Reliability for these measures was not provided.

Effects of program on comprehension measures could not be evaluated as percentages were used to assess overall change (see Vadasy et al., 2002). Providing the TP treatment had no effect on reading fluency and small effects on word recognition. The effect size for fluency was -.05. The effect size for word recognition was .32. Some caution must be taken when interpreting results as only 50% of the 11 quality indicators were met. Those indicators not met were: equivalent mortality rates for treatment and control groups, pretest equivalence for tutors, reliability for dependent variables, control for tutor effects, and randomization. The next treatment described is the tutoring program evaluated by Weiss et al. (1989).

Paired Reading. The tutors in Weiss et al.'s (1989) study provided nine 3rd through 5th-grade students more opportunities to read. The tutoring program was unstructured, took place in a suburban elementary school, and was conducted four times per week during the school day. The study lasted approximately 3 months. Each 20-30 min session involved tutors using a paired reading approach (see Weiss et al.). The tutor first asked student about the previous lesson. Next, the tutor and student read together. Finally, the student read independently. The primary focus of the intervention was on improving reading fluency. When students were unsure of a word, the tutor read it and then asked the students to repeat. If they continued to miss the word, tutors used flash cards to drill correct reading.

Twelve tutors provided the Paired Reading instruction (Weiss et al., 1989). The tutors were mostly senior citizens without professional teaching experience (8/12; 67%). The tutors were not paid and were required to attend 5 hours of training. Weiss et al.

(1989) evaluated effects of the paired reading program on reading fluency. All measures were reliable.

The Paired Reading instruction did not improve the students' reading fluency (see Table 2). The effect size for fluency was -.20. Caution must be used when interpreting the results because the study did not meet the majority of the 11 quality indicators (5/11; 45% met). The following six quality indicators were not met: equivalent mortality rates, pretest equivalence of students and tutors, treatment fidelity, and control for Hawthorne and tutor effects. The final study (Wilfong, 2006) evaluated the effects of another unstructured program, the Poetry Academy.

***Poetry Academy*.** The final treatment in Table 2 is the 3-month Poetry Academy (Wilfong, 2006). Thirty-six students in the third grade were presented with a reading intervention aimed at improving their fluency and reading comprehension. The study occurred four classrooms at one rural elementary school. The program was based on Fast Start (repeatedly reading poetry with parents to increase fluency) and Reading Recovery (rereading of familiar text to increase fluency). The procedures were unstructured, with tutors arranging 10 mins each week to spend with each tutee. During that time they repeatedly read engaging poems with the 3rd grade tutee.

Three out of the six community tutors used in Wilfong's (2006) study were non-education majors. Only one of the tutors was paid from an outside school source, the others were not paid. All of the tutors were required to attend 2 hours of training. Wilfong evaluated the effects of the Poetry Academy on reading fluency using one curriculum-based measure. Reliability for measures was not provided.

Although Wilfong (2006) reported larger gain scores for the students receiving the Poetry Academy treatment, calculation of effect sizes (using pre and post-test scores and standard deviations) indicated there was no effect on reading fluency. The effect size for fluency was –2.1. This negative effect was most likely because the treatment and control group were not equivalent at pre-test (see Wilfong). Caution must be used when interpreting results as 7 quality indicators (32%) were not met: equivalent mortality rates, pre-test equivalence, treatment fidelity, control for Hawthorne effects, dependent variable reliability, control for tutor effects, and randomly assigning students to groups. The following is a summary of the results for the six structured and unstructured programs that do not have multiple foci.

Summary of unstructured and structured programs with a single focus on fluency or comprehension. The six unstructured (SMART, Baker et al., 2000; Paired Reading, Weiss et al., 1989; Poetry Academy, Wilfong, 2006) and structured programs (McPeak, 2000; ITP, Meyer et al., 2002; TP, Vadasy et al., 2002) with singular foci had varied effects on improving comprehension, fluency, and word recognition of 141 students in grades 1 to 5. Studies occurred in a variety of locations (rural, urban, and suburban) both during and after school. The studies lasted three months (McPeak, 2000; Weiss et al., 1989; Wilfong, 2006) to six months (Baker et al., 2000; Meyer et al., 2002; Vadasy et al., 2002) and instructional sessions occurred between one (McPeak; Wilfong), two (Baker et al.), three (Meyer et al.), and four times per week (Vadasy et al., 2002; Weiss et al., 1989).

There were at least 56 tutors across the six studies. Two of the studies did not provide total tutor numbers (Baker et al., 2000; Vadasy et al., 2002). Half of the

treatments provided tutors with a small stipend (TP; Vadasy et al., 2002) or some other compensation (SMART, Baker et al., 2000; ITP; Meyer et al., 2002). All six studies provided total volunteer training time. The average amount of time in training across these studies was close to 5 hours (range 1.5 to 12). Structured programs provided more tutor training (Mean = 6.67) than unstructured programs (Mean = 2.83).

Overall the structured and unstructured programs with a singular focus had small effects on improving reading comprehension and word recognition and negative effects on improving oral reading fluency. Studies measured effectiveness on comprehension measures (Baker et al., 2000; McPeak, 2000; Meyer et al., 2002), fluency measures (Baker et al., 2000; Vadasy et al., 2002; Weiss et al., 1989; Wilfong, 2006), and word recognition measures (Baker et al., 2000; Vadasy et al., 2002). The average effect on improving reading comprehension was .22 *(n* = 3; range .22 to .52). The average effect on improving word recognition was .36 *(n* = 2; range .32 to .39), whereas the average effect on improving oral reading fluency was only -.47 *(n* = 4; range -2.1 to .46). Because the effects varied considerably across the studies, the effects are now presented based on treatment duration, as this variable should be associated with positive versus negative effects.

The six structured and unstructured treatments lasted between three months (McPeak, 2000; Weiss et al., 1989; Wilfong, 2006) and at least six months (Baker et al., 2000; Meyer et al., 2002; Vadasy et al., 2002). The longer programs reported moderate effects on improving comprehension, fluency, and/or word recognition at post-test (Baker et al.; Meyer et al.; Vadasy et al.), whereas the shorter programs reported all negative effects (McPeak; Weiss et al.; Wilfong). Averaging across all of the outcome measures

(e.g., word recognition), the average effect of the longer programs on outcome measures was .33 *(n* = 6; range -.05 to .52). The average effect of the shorter programs on outcome measures was -.83 *(n* = 3; range -2.1 to -.19). Another possible reason that the three shorter programs (McPeak; Weiss et al; Wilfong) reported negative effects was that they all failed to establish pretest equivalence of treatment and control groups (see Table 3).

Moderate confidence can be placed in these findings based on the number of quality indicators met (see Table 3). Although only three of the studies were randomized control trials (Baker et al., 2000; Meyer et al., 2002; Weiss et al., 1989), the majority met at least half of the 11 quality indicators (exception Wilfong, 2006). The average number of quality indicators met was 5.9 (54%). At least five out of the 6 studies met the following quality indicators: instructional conditions described, tutor training described, and no ceiling or floor effects. Some caution should be taken as the majority of studies failed to meet the following quality indicators: equivalent mortality rates, pretest equivalence of the tutors, treatment fidelity, and control for tutor effects (see Table 3).

Summary of Treatments, Study Quality, Outcome Measures, Student Characteristics, and Tutor Characteristics

A summary of results from all 19 experimental and quasi-experimental studies is presented next (see Table 4). The results are presented in a format similar to summaries for each category of studies (e.g., unstructured and structured treatments with a single focus). Namely, student participants and treatments (e.g., structured programs with multiple foci) are summarized. Then, tutor descriptions and outcome measures (e.g., word recognition at post-test and maintenance) are presented. Next, the effect sizes are

Table 4. Summary of Data Across all 19 Experimental and Quasi-Experimental Studies

Type of Reading Program	N	Avg. Quality Score (out of 10)	Avg. effects on ORF	Avg. effects on RC	Avg. effects on WR	Tutee Character. (Grade)	Treatment Duration / Intensity	Tutor Characteristics
Across All 19 studies	19	3.9 (range 0.0 to 8.0)	.14 (n = 12 range -2.1 to .73)	-.14 (n = 9 range -1.1 to 1.3)	.50 (n = 11 range -.04 to 1.1)	678 students (K-6)	Avg. Duration: 7.6 months (range 3 to 16 months; *n* =16/19) Avg. Intensity: 3.5 days /week (range 1 to 5 days; *n* =14/19)	At least 203 tutors 11/19 studies (9 treatments) provided small stipends or compensation Average tutor training: 8.0 hours (range 1.5 to 20 hours; *n* =13/19)
Structured Programs with Focus on WR and RC or WR and ORF	6	4.2 (range 2.5 to 5.0)	.41 (*n*=6 range .16 to .73)	--	.47 (*n*=8 range -.04 to 1.06)	180 students (1)	Avg. Duration: 7.3 months (range 6 to 10 months; *n* =6/6) Avg. Intensity: 4.0 days /week (range 3 to 5 days; *n* =5/6)	At least 85 tutors 4/6 studies (2 treatments) provided small stipends Average tutor training: 10 hours (range 9 to 15 hours; *n* =4/6)
Structured and Unstructured programs with a singular focus	6	5.3 (range 2.5 to 8.0)	-.47 (*n*=4 range -2.1 to .46)	.22 (*n*=3 range -.19 to .52)	.36 (*n*=2 range .32 to .39)	141 students (1-5)	Avg. Duration: 5.8 months (range 3 to 12 months; *n* = 6/6) Avg. Intensity: 2.5 days /week (range 1 to 4 days; *n* = 6/6)	At least 56 tutors 3/6 studies (3 treatments) provided small stipends or compensation Average tutor training: 5 hours (range 1.5 to 12 hours; *n* = 6/6)
Structured Programs with Multiple Foci	7	2.6 (range 0.0 to 8.0)	.53 (*n*=2 range .37 to .69)	-.28 (*n*=6 range -1.27 to 1.31)	.51 (*n*=2 range .46 to .55)	357 students (K-6)	Avg. Duration: 10.8 months (range 8 to 16 months; *n* =4/7) Avg. Intensity: 4.3 days /week (range 4 to 5 days; *n* =3/7;)	At least 62 tutors 4/7 studies (4 treatments) provided small stipends or compensation Average tutor training: 11.7 hours (range 2 to 20 hours; *n* =3/7)

Notes. N = Number of studies; Avg. = Average. Effects = Effect Sizes Calculated using *Cohen's d*; ORF = Oral Reading Fluency; RC = Reading Comprehension; WR = Word Recognition; Character. = Characteristics; K = Kindergarten; -- = no data.

summarized. The effect sizes are presented using categories found influential in other reviews of 1:1 tutoring (e.g., program structure; Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). First, the data related to student participants and treatments are presented.

Student participants and treatments. Nineteen experimental and quasi-experimental studies examined the effectiveness of delivering academic tutoring via nonprofessional adults (all non-parents) on academic student outcomes (e.g., word recognition). There were three types of treatments/programs evaluated: structured programs with a focus on *two* of the aforementioned outcomes *(n* = 6); unstructured and structured programs with a focus on *one* of the aforementioned outcomes *(n* = 6); and structured programs with a focus on comprehension, fluency, and word recognition *(n* = 7; see Table 4). Only two of the 15 programs (i.e., HOSTS, SP) were evaluated in more than one study (Burns et al., 2004; Ramey, 1990, 1991; Vadasy et al., 1997a, 1997b, 2000; see Table 2). Across the three types of treatments (e.g., unstructured and structured treatments with a single focus), there were 678 students in grades Kindergarten through sixth tutored by at least 203 nonprofessional adult tutor (all non-parents; see Table 4). The studies occurred in a variety of locations (rural, suburban, and urban), both during the school day *(n* = 12; 63%), and immediately after the school day had ended *(n* = 7; 37%; see Table 2). The average length of treatment across the 16 studies was 7.6 months (studies evaluating HOSTS studies did not provide this data). Treatment length ranged from 3 months (McPeak, 2000; Weiss et al., 1989; Wilfong, 2006) to 12 months (Morris et al., 1990). Fifteen studies provided information on the intensity of treatment (i.e., number of days per week). On average, tutors worked individually with students 3.5 days per week (range 1 to 5 days).

Tutor descriptions and outcome measures. There were at least 203 nonprofessional adults trained to tutor school-aged children. Seven studies did not report total tutor numbers (Baker et al., 2000; Burns et al., 2004; Morris et al., 1990; Nielson, 1991; Ramey, 1990, 1991; Vadasy et al., 2002). All of the adults worked one-on-one with a student. Sixty percent (9/15) of the treatments provided some sort of incentive for tutor participation (e.g., small stipend). Tutor training was described for 14/15 treatments (Rimm-Kaufman, 1999 did not report). The average time in training was 8.0 hours (range 1.5 to 15 hours).

Across the 19 studies, the effectiveness of the programs was evaluated on reading comprehension (9/19; 47%), oral reading fluency (12/19; 63%), and/or word recognition (9/19; 47%) outcome measures (see Table 4). Most studies examined the effectiveness immediately after treatment (13/19; 68%), although some collected maintenance data as well (Meyer et al., 2002; Nielson, 1991; Ramey, 1990, 1991; Vadasy et al., 1997a, 2000). At post-test, the majority of effect sizes for reading comprehension outcome measures were calculated using standard scores (9/11; 82%). The word recognition outcome measures were split fairly evenly between standard scores (5/11; 45%) and raw scores (6/11; 55%). Two studies reported effects on word recognition using both raw and standard scores (Vadasy et al., 1997a, 1997b). All of the effects on oral reading fluency were reported using raw scores (12/12).

The majority of the 19 studies (13/19; 68%) presented in Table 2 reported positive effects for nonprofessional adult-led tutoring programs on reading comprehension, oral reading fluency, and word recognition, but the effects varied. The following section highlights the overall effects.

The first section provides a summary of effect size differences related to outcome measures (e.g., word recognition). The differences related to the type of score (raw versus standard) used to calculate effect sizes are presented first followed by an overall comparison of effects on outcome measures (e.g., word recognition) regardless of type of score used (raw versus standard). In addition to differences related to outcome measures, previous reviews of 1:1 tutoring programs (non-parent delivered) have shown that subject matter, program structure, treatment duration, response to treatment, and tutor type and training were sources of effect size variability (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). Thus, in this next section, overall effects are also organized using similar variables (with the exception of subject matter, as all of the studies in this review focused on reading programs). This included organizing effects for: (a) program structure, (b) treatment duration, (c) tutee grade level, (d) tutor type, and (e) tutor training.

Effect sizes for reading comprehension, oral reading fluency, and word recognition measures overall and by study characteristics.

Differences between outcome measures by type of score and overall averages. Effect sizes were calculated across three outcome measures at both post-test and maintenance and reported for both raw (when available) and standard scores. The effectiveness of nonprofessional adult-led tutoring on reading comprehension and word recognition was evaluated using both standard and raw scores (see Table 2). These differences are presented first.

The average effect size for reading comprehension measures varied when effect sizes were calculated using raw or standard scores, however these differences did not exist for word recognition or oral reading fluency measures. On average, the seven

nonprofessional adult-led tutoring programs (TAILS, Al Otaiba et al., 2005; HOSTS, Burns et al., 2004; Ramey, 1990, 1991; Houghton Mifflin, McGrady, 1983; Ginn Reading Series, Nielson, 1991; SMART, Baker et al., 2000; Reciprocal teaching, McPeak, 2000; Intergenerational Tutoring Program, Meyer et al., 2002) had moderate effects on improving reading comprehension for students in grades Kindergarten through six when calculated using raw scores and negative effects when calculated with standard scores. The overall effects diminished at maintenance. Using raw scores, the average effect on reading comprehension was .51 *(n* = 2; range .50 to .52) at post-test and .80 at maintenance, whereas it was -.28 *(n* = 9; range -1.27 to .73) at post-test and -1.23 *(n* = 5; range -1.91 to -.94) at maintenance for standard scores.

The differences between effect sizes calculated using standard and raw scores were not observed on word recognition measures (effect sizes for fluency measures were only calculated using raw scores). On average, there were moderate effects across the eight nonprofessional adult-led tutoring programs (TAILS, Al Otaiba et al., 2005; Howard Street, Morris et al., 1990; Intergenerational Reading Program, Rimm-Kaufman et al., 1999; Intergenerational Literacy Project, Sankaranarayanan, 1998; Sound Partners, Vadasy et al., 1997a, 1997b, 2000; Wallach & Wallach, 1976; SMART, Baker et al., 2000; Thinking Partners, Vadasy et al., 2002) that evaluated effects on improving word recognition for students in Kindergarten through third grade. The effects were similar when calculated using raw and standard scores and maintained after programs ended. Using standard scores, the average effect size on word recognition was .56 *(n* = 5; range .32 to 1.06) at post-test and .53 *(n* = 2; no range) at maintenance. Using raw scores, the average effect size on word recognition was .44 *(n* = 6; range -.04 to .94) at post-test (no

maintenance data available). Effects for fluency measures are not disaggregated as they were only reported using raw scores.

Combining raw and standard scores, the overall effect of providing nonprofessional adult-led tutoring varied depending on the outcome measure (e.g., word recognition) and time of assessment (e.g., maintenance). The smallest effects were for reading comprehension, with slightly stronger effects for oral reading fluency. The largest effects of treatments were on word recognition measures. With the exception of word recognition (although only two effects were used to calculate the average effect size at maintenance for word recognition measures), effects diminished at maintenance.

The average effect size on reading comprehension was -.14 *(n* = 11; range -1.27 to 1.31) at post-test and -.90 *(n* = 6; range -1.91 to .80) at maintenance. The average effect size on oral reading fluency was .14 *(n* = 12; range -2.1 to .73) at post-test and -.10 *(n* = 2; no range) at maintenance. The average effect size on word recognition was .50 *(n* = 11; range -.04 to 1.06) at post-test and .53 *(n* = 2; no range) at maintenance.

Effect sizes for five study characteristics. Effect sizes varied across certain outcome measures (e.g., reading comprehension) when separated and compared by program structure, treatment duration, tutee grade level, tutor compensation, and tutor training. Effect size differences associated with program structure are presented first.

The 19 studies were separated into two main categories: structured programs with multiple foci and unstructured and structured programs with a singular focus. The structured programs with multiple foci were further separated into two categories: those that focused on three reading areas (e.g., word recognition) and those that focused on only two. Comparisons were only made if there were at least three effect sizes in each

category (e.g., structured programs with multiple foci). Two comparisons were made with the unstructured and structured programs with a singular focus. Comparisons were made for reading comprehension and oral reading fluency outcome measures.

On reading comprehension measures, the average effect size was smaller for studies identified as structured with multiple foci when compared to unstructured and structured programs with a singular focus (see Table 2). The average effect size for reading comprehension for the unstructured and structured singular focused programs was .22 *(n* = 3; range -.19 to .52), whereas it was -.28 *(n* = 7; range -1.27 to 1.31) for structured programs with multiple foci. The reverse relation was found for oral reading fluency. The average effect size on oral reading fluency was .41 *(n* = 6; range .16 to .73) for structured programs with two foci (e.g., SP; Vadasy et al., 1997a, 1997b, 2000), but -.47 *(n* = 4; range -2.1 to .46) for unstructured and structured programs with a singular focus. Due to the low number of effect sizes, no other comparisons were made.

The overall effects of nonprofessional adult-led tutoring on certain academic outcome measures (e.g., reading comprehension) also varied depending on the duration of the program (3 months or less versus 6 months or longer; there were no studies that lasted more than 3 months but less than 6 months). Effect sizes varied on reading comprehension and oral reading fluency measures for studies that lasted 6 months or longer versus those that lasted 3 months or less. The average effect size for reading comprehension was .51 *(n* = 5; range -.35 to .73) at post-test for studies that lasted 6 months or longer and -.47 *(n* = 4; range -1.11 to .50) for studies that lasted three months or less. The average effect size for oral reading fluency was .40 *(n* = 9; range -.05 to .73) at post-test for studies that lasted 6 months or longer and -.64 *(n* = 3; range -2.1 to .37)

for studies that lasted 3 months or less. Word recognition measures were only evaluated in studies that lasted six months or longer.

The average effects of nonprofessional adult-led tutoring programs on academic outcome measures were further separated based on tutee grade-level. Effect sizes were separated first by outcome measure (e.g., word recognition) and then by primary (Kindergarten through third grade) versus intermediate (fourth through sixth grade) grade levels. The Burns et al. (2004) study was grouped with the primary-level group although some fourth graders were involved in the study.

Effect size differences related to tutee grade level were only calculated for reading comprehension, as the other measures were primarily assessed with young students (students in grades 1 to 3). For example, all of the nonprofessional adult-led reading intervention studies that measured the effects on word recognition measures (9/9 studies) were conducted with young students. Similarly, nearly all of the studies that measured the effects on oral reading fluency measures were conducted with young students (10/12 studies) The average effect size for reading comprehension was small for students in grades kindergarten to three and negative for students in grades four to six. The average effect size was .06 *(n* = 7; range -1.27 to 1.31) for younger students and -.49 *(n* = 4; range -1.19 to .52) for older students. The remaining effect sizes for oral reading fluency and word recognition outcome measures were all conducted with Kindergarten through fourth grade students. The final effect size comparison was made for tutor-related characteristics.

Next, the average effect on improving reading comprehension changed from moderate to negative when studies were separated into those that provided tutor

compensation and those that did not. There was an overall moderate effect (ES = .51; n = 5; range -.35 to 1.31) on reading comprehension when tutor compensation was provided, whereas there was a negative effect (ES = -.47; n = 4; range -1.11 to .50) when such compensation was not provided (or described). There were similar outcomes, however, for both groups on fluency measures (ES = .10 for studies that provided compensation versus .21 for studies that did not). Almost all of the effect sizes for word recognition measures (10/11; 91%) were in studies that provided some sort of tutor compensation, so no comparisons could be made.

Finally, the average effect on all three outcome-measures improved when more tutor training was provided. Studies were separated into two groups: studies that provided two hours or more of training and studies that did not provide total tutor training time (e.g., Wallach & Wallach, 1976) or provided two hours or less of tutor training. The average effect size on reading comprehension was -.32 *(n* = 7; range -1.27 to 1.31) for studies that provided less tutor training, compared to .18 *(n* = 4; range -.35 to .73) for studies that provided more tutor training. The differences were not as great on oral reading fluency and word recognition measures, however, average effect sizes were larger for studies that provided more training.

The average effect size on oral reading fluency improved from .10 *(n* = 6; range -2.1 to .73) for studies that provided less tutor training to .18 *(n* = 6; range -.20 to .49) for studies that provided more tutor training. The average effect size on word recognition changed from .37 *(n* = 4, range -.04 to .68) for studies that provided less training to .57 *(n* = 7; range .30 to 1.06) for studies that provided more tutor training.

In the final results section, the results related to qualitative research are reviewed.

The same inclusion criteria were used to gather qualitative research as were used to gather experimental and quasi-experimental research, however only one qualitative study was obtained.

Results for Qualitative Study

One qualitative study was found that examined the effects of nonprofessional adult-led tutoring/teaching on academic student outcome measures (Neuman & Roskos, 1993). The findings support those cited for the 19 experimental and quasi-experimental studies (i.e., nonprofessional adult-led tutoring has a positive impact on students' academic outcomes.

Neuman and Roskos (1993) evaluated the effectiveness of a nonstructured program that used nonprofessionals to increase 65 pre-school students' literacy-related activities during free-time activities at three Head Start locations. Neuman and Roskos changed play settings into "enriched office settings" (p. 101): demarcating the different activities within the office area (e.g., Office Exit) and grouping several literacy-related activities around familiar objects (e.g., telephone, telephone book, and message pad were located on a desk). The nonprofessional adults (all parents but not assigned to their own child's class) worked in the designated literacy areas, within each preschool class, three days per week for 5 months. Each session lasted 1 hour. The adults were instructed to actively engage the students in literacy related activities (e.g., taking a phone message) and modeling literate behavior (e.g., writing a grocery list). All interactions were dependent on children's interest and varied across the three classrooms.

Neuman and Roskos (1993) solicited three parents to participate in the treatment

condition. The Head Start parents were required to volunteer in classrooms in lieu of not paying for their children's pre-school program. All participants were women with high school diplomas. All participants could write at least one paragraph. Neuman and Roskos evaluated the effect of the program on several outcome measures. The outcome relevant to this review was word recognition (ability to read 7 posted signs in the office).

The Head Start literacy-enriched office intervention was effective at improving word recognition when compared to a control condition (Neuman & Roskos, 1993). The students in both groups were equivalent at pre-test on the Test of Early Reading Ability and low in engaging in literacy-related activities. After treatment, the students in the treatment group outperformed the control group in correctly reading the specific words associated with the office center (posted on the signs).

Discussion

There is evidence that nonprofessional adult-led remediation is a viable mechanism for improving Pre-kindergarten through sixth grade reading outcomes. Although research related to all types of academic outcomes was sought, only those related to reading outcomes were found. The current review of nonprofessional adult-led academic remediation is important because it extends previous research and provides data-derived recommendations for future practices and research.

More specifically, I included the following in the current study as a result of the review: (a) adequate tutor training, (b) tutor compensation, (c) maintenance data, (d) detailed information related to tutor-related characteristics such as previous tutoring experience. The results of the literature review also highlighted the need to conduct more

rigorous research when evaluating the effectiveness of nonprofessional adult volunteer-led literacy interventions. Thus, I conducted high quality research, using a single subject design that could establish experimental control (multiple-probe design). Finally, the review highlighted the need to move beyond reading and evaluate the effects of a nonprofessional adult volunteer-led *writing* intervention. My current study does this. More details regarding these recommendations is provided later in this chapter.

How findings support, extend, or contradict previous findings. There have been three major reviews of the effectiveness of adult-led remediation or tutoring programs on student achievement (Erlbaum et al., 2000; Ritter et al., 2009; Wasik, 1998). The current review examined the effects of previously identified program characteristics on academic outcomes. Results supported, contradicted, and extended results published in previous reviews (see Table 5). These similarities, differences, and extensions are explored next within the context of the five program characteristics: program structure, outcome measures, treatment duration, response to treatment, and tutor type/training.

Program structure. The current review extends previous research by evaluating effects in light of the overall quality of the study and contradicts previous findings that suggested all structured programs outperform unstructured programs. Previous reviews of the effectiveness of adult-led remediation on student achievement reported that structured programs were more effective than unstructured programs (Erlbaum et al., 2000). The average effect size of structured programs to programs that focused on visual perception skills was .50 versus .03 (Erlbaum et al., 2000). Similarly, Wasik et al. (1998) reported that more complex programs were superior to other programs. Findings from this review did not support her conclusion.

Table 5. Summary of Effect Sizes Across 3 Different Reviews of Adult-Led Literacy Interventions: Separated Across Oral Reading Fluency (ORF), Reading Comprehension (RC), and Word Recognition (WR) Measures By Study Characteristics (when at least three effect sizes were available)

Focus (N)	Program Structure	Tutee Characteristics	Tutor Compensation/ Tutor Training	Treatment Duration
Nonprofessional adult tutors (19)	Differences across RC Small effect sizes for unstructured and structured programs with singular focus (.22, n = 3, range -.19 to .52), but negative effect sizes for structured programs with multiple foci (-.28, n = 7, range -1.27 to 1.31) Differences across ORF Small effect sizes for structured programs with 2 foci (.41, n = 6, range .16 to .73), but negative effect sizes for unstructured and structured programs with a singular focus (-.47, n = 4, range -2.1 to .46)	Differences across RC Small effect sizes for students in grades kindergarten to 3 (.06, n = 7, range -1.27 to 1.31), but negative for students in grades 4 to 6 (-.49, n = 4, range -1.19 to .52) No other comparisons could be made because most treatments delivered to young students.	Differences across RC Moderate effect sizes for treatments that provided tutor compensation (.51, n = 5, range -.35 to 1.31), but negative effects for treatments that did not provide or did not describe (-.47, n = 4, range -1.11 to .50). Small effect sizes when 2 or more hours of tutor training was provided (.18, n = 4, range -.35 to .73), but negative effects when less tutor training was provided or not described (-.32, n = 7, range -1.27 to 1.31) Differences across WR Moderate effect sizes when more training provided (.57, n = 7, range .30 to 1.06), but small with less training (.37, n = 4, range -.04 to .68)	Differences across RC Moderate effect sizes for treatments that lasted 6 months or longer (.51, n = 5, range -.35 to .73), but negative effects for treatments that lasted 3 months or less (-.47, n = 4, range -1.11 to .50) Differences across ORF Small effect sizes for treatments that lasted 6 months or longer (.40, n = 9, range -.05 to .73), but negative effects for treatments that lasted 3 months or less (-.64, n = 3, range -2.1 to .37)
Adult Tutors (29) (Erlbaum et al., 2000)	Across all measures Small effect sizes for programs focused on mixed skills (.50, n = 30), but negative when inadequate information provided (-.07, n =4)	Across all measures Small effect sizes for students in grades 1 through 3 (.46, n = 28) and 2 to 3 (.37, n = 8), but small for students in grades 4 to 6 (.06, n = 5)	Across all measures Moderate effect sizes when community nonprofessional adults were trained (.59, n = 5), but negative when not trained or described (-.17, n = 3).	Across all measures Large effect sizes for programs that lasted 20 weeks or less (.65, n = 14), but small effect sizes for shorter programs (.37, n = 16). However, no statistically differences when evaluated according to total time in treatment.
Volunteer Tutors (21) (Ritter et al. 2009)	Differences across global reading outcomes Moderate effects for highly structured programs (ES = .59), but small effect sizes for less structured programs (ES = .14)	No statistically significant differences due to grade level	Did not find that effect sizes varied across the following tutor types: community volunteer, parent, or college student. Did not examine differences across tutor compensation/tutor training.	Did not examine across this variable.

In the current review, treatments were separated based on structure and complexity (number of elements that the program focused on). The first finding was that studies that examined structured programs with multiple foci were of poor quality (only 1 out of 6 programs met more than 50% of the 11 quality indicators). Additionally, the only measure where an effect size could be averaged (more than 3) was reading comprehension. The average effect size for reading comprehension was -.28 *(n* = 8; range -1.27 to 1.31), whereas the effect size for unstructured and structured programs with singular foci on reading comprehension was .22 *(n* = 3; range -.19 to .52).

In the current review, the studies that reported the largest and most consistent effects were structured programs with a focus on word recognition and oral reading fluency (e.g., Sound Partners, Vadasy et al., 1997a, 1997b, 2000). The average effect size across these programs on word recognition measures was .53 *(n* = 7; range -.04 to 1.06) and on oral reading fluency measures was .41 *(n* = 6; range .16 to .73; post-treatment data only). Additionally, effect size differences across program types (e.g., structured programs with multiple foci) differed as a result of outcome measure (reading comprehension and oral reading fluency measures).

***Outcome measures*.** The current review extended as well as supported previous findings of the effects of adult-led instruction on specific outcome measures (e.g., word recognition). First, the findings were extended by the inclusion of qualitative results (Neuman & Roskos, 1993). The findings from the qualitative study confirm what Erlbaum reported in 2000: nonprofessionals can have a positive effect on improving word recognition. The results from the experimental and quasi-experimental studies also supported and extended Erlbaum et al.'s (2000) finding related to increased effect sizes

for word recognition measures, however the effect sizes were not as large. In the current review, the average effect size across word recognition measures was .49 *(n* = 11; range -.04 to 1.06). In Erlbaum et al.'s review the average effect size across word recognition measures was .94.

The current review also extended findings from previous reviews by separating outcomes not only into specific outcome measures (e.g., word recognition), but also by time after treatment (e.g., post-treatment versus maintenance). Conclusions could not be drawn for maintenance data, as there were too few effect sizes to calculate an average effect. The current review also supported Erlbaum et al.'s (2000) conclusion that there are no apparent differences between raw and standard scores. Cohen et al. (1982) reported that non-standardized scores were associated with larger effects, whereas Erlbaum et al. reported no difference. In the current review, such differences (raw versus standard scores) among effects sizes could only be calculated across word recognition outcome measures (only outcome measures with at least 3 effect sizes in each category). As with Erlbaum et al., these differences did not appear dissimilar (ES = .56 for standard scores and ES = .44 for raw scores). The next section highlights differences in findings related to treatment duration.

***Treatment duration*.** The current findings both extended and contradicted previous findings related to treatment duration. Previous reviews of adult-led academic remediation found that shorter programs were associated with stronger effects, however the current review found the opposite to be true. Effect sizes were larger across two outcome measures (see below) for programs of six months or longer when compared to programs three months or less (there were no programs that lasted between three and six

months).

In the current review, the average effect size across reading comprehension measures was .51 *(n* = 5; range -.35 to 1.31) for programs that lasted 6 months or longer and -.47 *(n* = 4; range -1.11 to .50) for programs that did not provide duration data or lasted 3 months or less. The effects across oral reading fluency measures were also larger for longer treatments with an average effect size of .40 *(n* = 9; range -.05 to .73) for programs that lasted 6 months or longer and -.64 *(n* = 3; range -2.1 to .37) for programs that lasted three months or less. Comparisons could not be made across word recognition measures as all were evaluated in programs lasting at least 6 months. Next, I offer support for findings related to response to treatment.

***Age of students*.** Programs for younger students were again associated with larger effects, but only across certain outcome measures. Effect sizes were first separated across similar outcome measures and then separated across two grade-levels: (a) Kindergarten through grade three, and (b) grades four through six. As in previous reviews, effect sizes were larger for younger students (Erlbaum et al., 2000), but comparisons could only be made across reading comprehension measures. The average effect size on reading comprehension for younger students (included Burns et al., 2004 study although a percentage of those students were in grade 4) was .06 *(n* = 7; range -1.27 to 1.31), whereas for older students it was -.49 *(n* = 4; range -1.19 to .52). Erlbaum reported a much larger average effect size for younger students (ES = .42 for younger students). The final program characteristic evaluated related to tutor-level characteristics.

***Tutor type/training*.** The current review extended previous research of the effectiveness of adult-led academic remediation/tutoring by focusing only on the effects

of programs delivered by nonprofessional adults. Previous reviews have included some applicable studies (Erlbaum et al., 2000), but have not assessed effectiveness of programs in a systematic way, tempered results based on study quality, or provided clear connections between data and recommendations (Wasik 1998; Wasik & Slavin, 1993; Wasik et al., 2002). Others have suggested that tutor compensation (e.g., small stipend for participation) is related to overall program success, but this had not been examined with data until the current review.

Results indicated that in fact tutor compensation was associated with stronger effects but only across certain outcome measures (e.g., reading comprehension). The average effect size for reading comprehension was .51 when tutors were paid versus -.47 when tutors were not paid. Similar findings were not found across oral reading fluency measures (similar effect sizes of .10 and .21). Such a comparison could not be made for word recognition measures as most studies provided tutor compensation.

Having examined how the current review extended, supported, and contradicted previous reviews, I now provide recommendations. Recommendations are made both for practice as well as for future research.

Recommendations for practice and future research. In this section I highlight three recommendations for implementing school-based nonprofessional adult-led academic remediation programs and five recommendations for future research. Recommendations are data-driven and tempered by overall quality of studies from which they were drawn. The five recommendations had a direct influence on the design of the current study.

1. Schools should increase the use of nonprofessional adults in academic remediation programs for students in grades pre-Kindergarten through six. The findings from this review indicate that both structured programs with a focus on word recognition and fluency and unstructured and structured programs with a singular focus are effective at improving oral reading fluency and word recognition. Confidence can be placed in these findings as at least 67% of the studies within each category (e.g., unstructured and structured programs with a singular foci) met at least 50% of the 11 quality indicators. Effect sizes were larger in programs that provided more tutor training as well as tutor compensation (see recommendations numbers two and three).

2. School-based nonprofessional adult-led academic remediation programs should include adequate training for the nonprofessional adults. Effect sizes were consistently larger across the three outcome measures (e.g., word recognition) when more extensive tutor training was provided. *Extensive* is defined *total time* in training (rather than the extent to which tutors were trained to criterion) as more complete tutor training descriptions were often not provided. When more training was provided, effect sizes increased across reading comprehension (from -.32 to .18), oral reading fluency (from .10 to .18), and word recognition measures (from .37 to .57). Confidence can be placed in these recommendations as 7 out of the 10 studies that provided more than two hours of tutor training also met at least 50% of the 11 quality indicators. Future researchers should adequately describe the tutor training that occurred. The total number of hours in training may be less critical than the extent to which tutors were taught to criterion,

3. Nonprofessional adults should be provided some sort of compensation (from free lunches to small hourly wages) for their participation. Wasik et al. (2002) suggested

that tutor compensation was important to a program's overall success. Tutor compensation did make a difference in overall effect sizes across reading comprehension measures in the studies evaluated in this study. The overall effect at post-test on reading comprehension measures increased from -.47 *(n* = 4; range -1.11 to .50) when no tutor compensation was provided to .51 *(n* = 5; range -.35 to 1.31) when compensation was provided. Although this relation was not tested experimentally, the large change in effect sizes seem to warrant a closer look at such practices. Confidence can be placed in these findings as three of the five studies (60%) that provided tutor compensation and measured effects on reading comprehension met at least 50% of the 11 quality indicators.

4. Future research of nonprofessional adult-led academic remediation programs should include maintenance data and examinations of diminished effects. There were six studies that evaluated the enduring effects of treatments (Nielson, 1991; Meyer et al., 2002; Ramey, 1990, 1991; Vadasy et al., 1997a, 2000). In the majority of the studies (4/6; 67%), effects diminished at maintenance. Differences between post-test to maintenance were calculated across reading comprehension measures (only outcome variable with three or more effect sizes at maintenance). Effects on reading comprehension decreased from -.14 *(n* = 11; range -1.27 to 1.31) at post-test to -.90 *(n* = 6; range -1.91 to .80) at maintenance. Caution must be used when interpreting maintenance results, as only three of the six studies met at least 50% of the 11 quality indicators (Meyer et al., 2002; Vadasy et al., 1997a, 2000). Nevertheless, as lasting effects are important to consider in evaluating overall success, future research must continue to collect and analyze maintenance data and alter programs accordingly.

5. Future research of nonprofessional adult-led remediation should more carefully examine program effectiveness as it relates to tutor-related characteristics. Although 12 studies (63%) provided total tutor numbers, the majority of studies did not provide adequate information about the tutors. For example, only 9 studies (47%) indicated whether tutors had previous teaching or tutoring experience (4/9 studies met at least 50% of the 11 quality indicators). Just as tutor compensation appeared to have an impact on overall program effectiveness, so to could tutors experience (e.g., Were they retired teachers?) be associated not only with effectiveness, but also with program sustainability. Future research should explore these associations. Such information would greatly improve remediation services.

6. The quality of future research related to nonprofessional adult-led academic remediation must improve. The majority of the research evaluated in this review failed to meet the following quality indicators (number and percentage of studies that failed to meet quality indicator in parentheses): equivalent mortality rates (16/19; 84%); pre-test equivalence of tutors (14/19; 74%); treatment fidelity (14/19; 74%); control for Hawthorne effects (14/19; 74%); reliability of dependent variables (11/19; 58%); control for tutor effects (16/19; 84%). Finally, the majority of the studies failed to establish pre-test equivalence of tutees (10/19; 53%). Several evaluations of program effectiveness seemed to be influenced by these pre-test differences (e.g., Ramey, 1990, 1991; Wilfong, 2006) and led to my seventh recommendation.

7. Future research should use experimental designs that can establish control, even when large groups are not available for testing, when pre-test differences make drawing accurate post-test conclusions difficult, and/or when best practices mandate that

all students receive a treatment. Several programs could not be evaluated because the researchers were unable to establish adequate control (i.e., no control group). Single subject designs provide an excellent way to not only establish control, but also demonstrate possible progress, regardless of pre-test non-equivalence (Horner et al., 2005).

8. Finally, future research on the effectiveness of nonprofessional adult-led literacy remediation should go beyond reading interventions to include writing interventions as well. There is abundant research to support both the need to improve students' writing and the effectiveness of writing programs that have currently only been led by teachers and researchers (see Rogers & Graham, 2008). This review suggests that incorporating nonprofessional adults as tutors would be effective.

This review of the effectiveness of nonprofessional adult-led remediation/tutoring improved the overall understanding of such programs (program structure, outcome measures, treatment duration, age of students, and tutor type/tutor training). However, there are limitations related to comprehensiveness of the findings and methods used to compare overall effectiveness.

Limitations. The first limitation is related to exploring the varied effects of program structure that others have suggested are important contributors to a program's overall success. For example, Wasik et al. (1993, 1998) suggested that programs (led by knowledgeable coordinators) that used individually paced lessons, active student participation, explicit modeling, and scaffolded instruction are superior to those that do not. Studies within this review were not grouped according to these attributes (such thorough descriptions of treatments were often not provided within text) and therefore

possible relations were not explored. Other relations cited as important (e.g., relation between tutor incentives and overall effectiveness of program) were explored. Future research should attempt to examine more of these possible relations.

The second limitation is related to the methods used to compare overall effectiveness. The differences were not assessed using meta-analytic procedures. Differences among the five program features (e.g., shorter versus longer treatments) must therefore not be identified as statistically significant. There would have been much noise or error in the findings to do this, as the total number of effect sizes in each category never surpassed 12 and were typically less than 10. Much has been accomplished since previous reviews (more research that used random assignment to groups and more description of tutor training; see Erlbaum et al. 2000). Hopefully this improvement will continue so that the variability that exists among these findings can be more clearly evaluated.

Concluding comments. Is it a viable option to use nonprofessional adults in the community in academic remediation positions within school settings? Will there be an overall positive effect on improving one or more academic skills? Will these effects maintain? Do the results vary depending program-level characteristics (e.g., treatment duration), tutor-level characteristics (e.g., tutor compensation), or student-level characteristics (e.g., grade level)? The results from this review of literature suggest that nonprofessional adult-led reading instruction can be effective, but effects vary across outcome measures (e.g., word recognition) and across several study characteristics (e.g., program structure).

Future research must examine whether these positive effects can be carried over

to other academic subjects such as writing. There are currently no evaluations of the effect of nonprofessional adult-delivered academic instruction outside of reading. This is problematic because deficits exist among areas such as writing, with frequent calls to put effective practices into more use (see Graham and Perin, 1997a, 1997b). Using nonprofessional adults could provide a viable alternative to implementing effective practices that have only delivered, to date, by researchers and teachers (see Rogers & Graham, 2008).

In the next section, I describe a model that has been used repeatedly to improve students' writing performance: Self-Regulated Strategy Development model (SRSD; Harris, Graham, & Mason, 2003). SRSD has been well validated in the literature (Graham, 2006; Graham & Perin, 2007; Rogers & Graham, 2008). In the next section, I first provide a description of the SRSD model and then provide a summary of the effectiveness of SRSD instruction using several recent meta-analyses.

Review of SRSD Literature

In contrast to the previous section, I did not conduct a meta-analysis on SRSD, as there are several recent meta-analyses available (Graham, 2006a; Graham & Perin, 2007a; Rogers & Graham, 2008). In the proposed study, nonprofessional adults recruited from a school community will teach a story-writing strategy embedded within a Self-regulated strategy development framework (SRSD; Harris, Graham, and Mason, 2003). The SRSD framework provides teachers (peers or adults) with a structure to teach the strategies (academic and self-regulation) needed to improve students' writing (SRSD has been found effective for teaching math and reading as well, see Harris et al., 2003).

SRSD has been well validated in the literature (see Rogers & Graham, 2008). Before reviewing the literature related to SRSD, I provide a description of this model.

SRSD Defined

Characteristics of. SRSD has several key characteristics. First, SRSD is collaborative. The teacher and student set goals for writing (include all seven story parts, make sense) and then plan and write together, with students gradually directing the process and instructors providing support as needed (see Harris et al., 2006). Instruction requires that teachers understand a student's current approach to writing as well as how the student self-regulates the writing process. Second, SRSD is individualized. Instruction is criterion-based with students receiving instruction until they are able to demonstrate successful independent performance. The ultimate goal is for students to consistently apply the strategy when needed (Harris et al., 2003). Third, SRSD also stresses good teaching. Teachers are encouraged to (a) deliver the instruction with enthusiasm, (b) be responsive, (c) seek out a support network to problem solve and generate appropriate changes needed for future lessons, and (d) provide ongoing assessment.

When using the SRSD approach to teach writing, there are three main goals (see Harris et al., 2003). First, students are explicitly taught the strategies needed for planning and drafting, writing, revising, and/or editing their writing. Second, students are taught (and given support with) how to regulate their own behavior through self-talk (e.g., "This is hard, but I know there's a strategy I can use.") and other self-regulation strategies (goal setting, self-assessment, self-instruction, self-reinforcement). Finally, students are taught

how to increase their motivation for completing a task through activities, such as obtaining visible evidence of their progress.

Six instructional stages. There are six main stages of instruction within the SRSD model: *Develop Background Knowledge*, *Discuss* the strategy, *Model* the strategy, *Memorize* the strategy, *Support* the students' use of the strategy, and *Independent Performance* (Harris et al., 2003). These stages are used to teach self-regulation (goal setting, self-assessment, self-instruction, and self-reinforcement) and target strategies and procedures, are non-linear, and should be thought of as recursive with re-teaching occurring when necessary (Harris et al., 2003). There are specific goals for each of the six SRSD instructional stages (for a more complete description of each SRSD instructional stage, see Harris et al., 2003).

The first part of SRSD instruction involves the stage, *Develop Background Knowledge*. During this stage, instruction is centered on (a) teaching related concepts and vocabulary needed to use the target strategies and self-regulation procedures and (b) generating excitement about learning these strategies and procedures (Harris et al., 2003). During this stage, it is essential that students are equipped with the necessary information needed to make learning successful. For example, if students are being taught how to write a persuasive essay, they must first understand what it means to persuade. Further, students must understand how different self-regulatory processes affect the learning process. For example, students may be introduced to self-statements at this stage and shown how such statements can either inhibit or promote progress. Also during this stage, the instructor is also obtaining information about the student so that more individualized instruction can be provided (see *Model It* stage below). The *Develop*

Background Knowledge stage is often combined with the *Discuss It* stage.

The next SRSD instructional stage, *Discuss It,* involves helping students understand the *what*, *when*, and *how* of the strategy being taught. In other words, students are introduced to the strategies, learning when and where to use them. This includes describing the strategies and procedures to be learned and the rationale for each. During this stage, students are also asked to commit to putting forth their best effort to both learn and use the strategies. As stated earlier, the SRSD instructional stages are recursive, and the *Develop Background Knowledge* and *Discuss It* stages are often combined within a lesson.

The third SRSD instructional stage is *Model It*. During this stage, the instructor overtly models how to use the target strategies and self-regulation procedures, involving students as collaborators in the use of those techniques, but taking the lead in showing how to use them (see Harris et al., 2003). In other words, students learn how to identify the writing problem and apply a special set of writing and self-regulation tricks (i.e., strategies) to successfully tackle the problem. Just as important, the instructor models how to use self-regulation procedures to manage students use of the strategies, writing process, or their behaviors. For example, as the instructor is already well aware of the type of self-statements the students use (see *Develop Background Stage*), he/she is now able to individually model appropriate self-statements while applying the strategy using language that is familiar to the child (see Harris et al., 2003). For example, the instructor may model how to use positive self-statements while modeling use of the strategy. After students demonstrate a readiness to become more of a leader in the learning process, the instruction moves to the next SRSD stage, *Support It*. However, before describing this

stage, I first describe the Memorize It stage.

The fourth SRSD instructional stage is *Memorize It*. During this stage, students are asked to memorize steps in the strategy and the meaning of any mnemonic that is used to represent the strategy steps or part of the steps (see Harris et al., 2003). The Memorize It stage often begins during the first or second lesson. Harris et al. (2003) stated that memorizing is most appropriate for students who have memory difficulties. By being able to memorize important elements of a strategy, students will improve the ease with which they are able to recall and use the strategies/processes and require less assistance and prompting from the teacher. This is necessary for them to move to the next SRSD instructional stage, *Support It.*

The fifth SRSD instructional stage is *Support It*. Harris et al. (2003) have compared the instruction during this stage to scaffolding used when making a building. The support provided is said to be scaffolded because teachers provide more support early on and then gradually remove this support as students demonstrate a more solid understanding of and ability to independently use strategies and self-regulation procedures (e.g., goal setting, self-monitoring, or self-reinforcement). By appropriately supporting the student, the instructor improves the likelihood that the student will maintain motivation and were able to maintain improvements and generalize both cognitive and affective changes (see Harris et al., 2003). Criterion levels are gradually increased (student becomes more and more responsible for performing said task with minimal prompts, interaction, or guidance) until students reach their final goal: independent performance.

The final SRSD instructional stage is *Independent Performance*. During this

stage, students are able to use strategies and self-regulation processes without prompts, interactions, or guidance (see Harris et al., 2003). Additionally, students are encouraged to work out steps quietly ("in your head"). This is done so that students can use strategies and processes in multiple environments (e.g., testing situation). During this final SRSD stage, it is important to continually monitor students' performance. If students demonstrate decreased performance, they should be provided booster sessions until they are once again ready to proceed to independent use of skills.

As indicated earlier, there is extensive research on the effectiveness of using the SRSD model to improve student outcomes, including as a means to improve story-writing skills for young struggling writers (see Rogers & Graham, 2008). Such instruction has resulted in a number of positive effects, involving enhancing students' writing performance, knowledge of writing, and approach to writing (Graham, 2006a; Rogers & Graham, 2008). In the next section, I present recent meta-analyses that have examined the effectiveness of SRSD in writing (Graham, 2006a; Graham & Perin, 2007a; Rogers & Graham, 2008). This is followed by a review of one recently conducted SRSD writing study (not included in any of the previously mentioned meta-analyses; Reid, Luschen, & Lienemann, 2009). The study conducted by Reid et al. supports the need to further investigate the effectiveness of delivering non-teacher/non-research-led SRSD instruction.

Meta-analyses related to SRSD

Within the past three years, Graham and his colleagues have conducted several meta-analyses of writing treatments (Graham, 2006a; Graham & Perin, 2007a; Rogers &

Graham, 2008). These meta-analyses are detailed in the following section so as to provide the evidence that SRSD has been well validated in the literature and to provide the rationale for the current proposed study. Although other meta-analyses exist (e.g., Harris & Graham, 2003), only the most recent meta-analyses of SRSD are presented, as they include studies examined in the earlier reviews.

Graham (2006a). The first meta-analysis described was conducted by Graham (2006a). In his analysis, Graham conducted a meta-analysis of all strategy instruction studies and compared the effectiveness of studies (group comparisons and single subject designs) that used the six-stage SRSD instruction to those studies that did not. The second meta-analysis described was conducted by Graham and Perin (2007a). In this meta-analysis, the authors evaluated the effects of a broader range of writing studies (i.e., included strategy instruction approaches as well as other writing interventions) delivered to adolescent students (grades 4 to 12). Finally, Rogers and Graham (2008) completed a third meta-analysis of single subject writing research. This review differed from the 2006 and 2007 meta-analysis because it focused only on single-subject design studies and, in contrast to Graham and Perin (2007a), included a broader range of students (grades Kindergarten to 12). The outline used to highlight each meta-analysis is as follows: (a) the studies included in the review and SRSD's representation among them, (b) methods employed in meta-analysis, (c) overall results, (d) results for SRSD studies, and (e) summary.

One of the earliest meta-analyses to compare the effectiveness of SRSD to other strategy instruction writing approaches was conducted by Graham (2006a). Graham included studies in the meta-analysis that (a) evaluated the effects of a writing treatment

in which students in grades 1 to 12 were taught one or more strategies for planning and drafting, revising, and/or editing compositions, and (b) were experimental, quasi experimental, or single subject design investigations. Studies that used SRSD instruction were well represented in this analysis making up almost half of the group studies (45%; 9/20) and well over half of the single subject studies (68%; 13/19).

Graham (2006a) evaluated group comparison studies by calculating effect sizes. Effect sizes were calculated by subtracting the posttest mean of the control group from the posttest mean of the treatment group and dividing by the standard deviation for the control group. Effect sizes were considered small if .20 or smaller, medium if .50, and large if .80 (see Graham, 2006a). Single subject effect sizes were calculated using a nonparametric approach: the percentage of nonoverlapping data (PND; Scruggs & Mastropieri, 2001). The PND is the percentage of data points in treatment that represent an improvement over the most positive data point obtained during baseline. PND were interpreted using criteria proposed by Scruggs and Mastropieri (2001): PND greater than 90% was a large effect; PND between 70.1% and 90% was a moderate effect; PND between 50.1% and 70% was a small effect; and PND 50% or below was classified as not effective. It should be noted that PND does not indicate the magnitude of an effect, rather the percentage of treatment data points that represent an improvement over the strongest baseline data point.

Graham (2006a) found that the overall impact of strategy instruction on improving students' writing outcomes was large. When the effect sizes from group comparison studies were averaged, the overall effect on writing outcome measures (quality, length, revisions, and mechanics) was 1.15 at post-test (n = 110; SD = 1.44).

These effects remained large several weeks after treatments had ended. The average effect size at maintenance (data only available in SRSD studies) was 1.32 (n = 24; SD = .93).

The overall impact of strategy instruction in the single subject designs was similar to that found for the group comparison studies. The average PND at post-test (across all writing outcome measures) was 89% (n = 58; SD = 19%). There were more maintenance data available in the single subject studies (74% of studies). The average PND at maintenance was 93% (n = 35; SD = 16%). When generalized to different persons or settings, strategy instruction effects remained large with a 90% PND (n = 18; SD = 20%). The effects were also assessed when generalized to different writing genres. The average PND for this type of outcome was 84% (n = 4; SD = 6%).

Graham (2006a) evaluated whether there were statistically significant differences among studies that used all SRSD instructional stages as opposed to those that did not. Graham reported that there were statistically significant post-test differences for group comparison studies that used SRSD versus those treatments that did not use SRSD ($p < .02$). In fact, the average effect size at posttest for studies that used all SRSD instructional stages was almost double that of the other studies (see Graham, 2006a). Comparisons of SRSD and non-SRSD studies could not be made at maintenance, as there were not enough data.

In summary, Graham (2006a) was among the first to compare the effectiveness of SRSD instruction with other strategy instruction using meta-analytic procedures. Graham conducted a meta-analysis of relevant group designs and single subject designs using effect size (group) and PND (single subject) calculations. The overall effect of using

strategy instruction on improving students' writing outcomes was large. When the group effect sizes were further analyzed, the overall effect of using the six-stage SRSD instructional model on improving writing almost doubled the effect observed in the other strategy instruction studies.

Graham and Perin (2007a). The second meta-analyses described went beyond examining the effect of using strategy instruction to include a wide range of writing instructional approaches (Graham & Perin, 2007a). This meta-analysis is described using the same five-part format: (a) the studies included in the review and SRSD's representation among them, (b) methods employed in meta-analysis, (c) overall results, (d) results for SRSD studies, and (e) summary.

The second meta-analysis was published in 2007 (Graham & Perin, 2007a), and evaluated the effects of writing instruction (including those that used SRSD instruction) delivered to adolescent (grades 4 to 12) students attending regular public or private schools. Writing studies were included if they (a) were an experimental or quasi-experimental group design, (b) included a reliable measure of writing quality, (c) provided data needed to calculate an effect size, and (d) were published in peer-refereed or non-refereed sources (e.g., journal articles as well as unpublished dissertations). The authors modeled much of their meta-analysis' procedures after Hillocks' (1986) seminal review of writing practices (see Graham & Perin, 2007a). For example, they identified separate categories of writing interventions similar to this previous review (i.e., strategy instruction, word processing, grammar instruction, sentence combining) and examined the effects of specific writing treatments on the quality of students writing by calculating average weighted effect sizes (see description below). Again, SRSD studies were well

represented in their findings. Out of the 107 total studies reviewed by Graham and Perin (2007a), approximately 19% (20/107) were strategy instruction studies. Among the strategy instruction studies, approximately 40% (8/20) used the SRSD instructional model.

The authors evaluated the overall effects by calculating mean weighted effect sizes and comparing effects between treatments that yielded four or more effect sizes (see Graham & Perin, 2007a). The authors also examined a study's overall writing quality (based on 9 quality indicators). Again, effect sizes were only calculated for reliable writing quality measures. The authors used *Cohen's d* or the standardized mean difference to calculate effect sizes (see Graham & Perin, 2007a). The effect sizes were calculated by subtracting the posttest mean of the control group from the posttest mean of the treatment group and dividing by the pooled standard deviation of the control and treatment group. The authors computed only one effect size per study. Effect sizes were also weighted by multiplying each effect size by its inverse variance (see Graham & Perin for a complete description). As in the other meta-analysis, effect sizes were considered small if .20 or smaller, medium if .50, and large if .80 (see Graham, 2006a). Study quality was also assessed using 9 quality indicators. These indicators were based on recommendations from leaders in the field (see Gersten et al., 2005). For example, studies were evaluated for having included treatment fidelity measures or ensuring that there was pre-treatment equivalence among the treatment and control groups.

The authors provided detailed results for the 15 different types of writing treatments (e.g., strategy instruction, word processing, grammar instruction, sentence combining). Those summarized here are provided to further the understanding of SRSD's

effectiveness. First, the authors reported that the overall quality of studies had improved over time. When the quality scores for all of the studies were analyzed, it was found that the difference between the more recent studies' quality scores and the earlier studies' quality scores was statistically significant ($r = 32$, $p < .01$). Of particular interest to this paper was the overall high quality of SRSD studies. All of the SRSD studies met over 55% (5/9) of the quality indicators. The average quality score among the 8 SRSD studies was 7.0 (range 5.0 to 9.0).

As in Graham's earlier meta-analysis (Graham, 2006a), this review showed that SRSD instruction had a large overall effect on improving adolescent students' writing. Effects were positive across all strategy instruction studies. Additionally, as in Graham's (2006a) earlier review, strategy instruction that used the SRSD model were superior to strategy instruction that did not use the SRSD model. The difference between SRSD and non-SRSD studies was statistically significant with SRSD studies surpassing other strategy instruction studies. The average weighted effect size for SRSD was 1.14. This was almost twice as large as the average weighted effect size for non-SRSD interventions (.62).

Additionally, the authors also compared differences in weighted mean effects among those writing treatments that used some form of explicit teaching (i.e., grammar, sentence combining, strategy instruction [SRSD and non-SRSD], and summarization). The authors did not calculate whether these differences were statistically significant, but the differences are worth noting. For example, the mean weighted effect size for SRSD (1.14) was higher than grammar (-.32), sentence combining (.50), and summarization (.82). Of further interest to the study proposed here, all writing treatments included in this

meta-analysis were delivered by teachers or researchers (see Graham & Perin, 2007a).

In summary, a meta-analysis of writing instruction delivered to adolescent students attending regular and public school revealed overall positive results, in particular the results were especially positive for those students who received SRSD strategy instruction (ES = 1.14; Graham & Perin, 2007a). The mean weighted effect size was large compared to other types of strategy instruction as well as compared to other explicit instruction (e.g., grammar instruction). The mean quality score for SRSD studies was also high with all studies meeting at least 55% of the 9 quality indicators (mean quality score among SRSD studies was 7.0).

Rogers and Graham (2008). The third and final meta-analysis was conducted to analyze single subject writing research (e.g., those conducted using reversal or multiple baseline designs; Rogers & Graham, 2008). The authors located 88 single subject design writing intervention studies. A sizable portion of these studies evaluated the effect of SRSD instruction on improving writing performance (31%, 27/88). Rogers and Graham used several indices to measure an intervention's overall effect. One nonparametric approach used to summarize and evaluate single subject design data was the percent of nonoverlapping data (see earlier description for Graham, 2006a).

Rogers and Graham also assessed whether experimental control was established within each study. In other words, did the intervention and the intervention alone cause the observed changes to occur? In single subject designs, experimental control is established by actively manipulating (instituting, modifying, or removing) the independent variable (or treatment) to determine its effects on the dependent (outcome) measures (Horner et al., 2005). For example, in multiple baseline designs (used in 61/88

[69%] of the studies evaluated in Rogers & Graham, 2008), experimental control is achieved by first establishing baseline stability across participants, behaviors, or settings and then systematically introducing a treatment (in a staggered fashion) across one or more legs or tiers. When the treatment demonstrates a predicted pattern of change across participants, behavior, or settings when the intervention is introduced, but remains stable in untreated legs/tiers, experimental control has been achieved. It is generally agreed that experimental control is not established until there is at least one demonstration and two replications showing that the manipulation had the predicted impact (see Kennedy, 2005). Therefore, Rogers and Graham indicated that experimental control had only been achieved when the predicted covariation between the introduction of the treatment and changes in the dependent measure(s) were demonstrated through at least three demonstrations in an experiment.

Rogers and Graham calculated mean and median PNDs across similar outcome measures (e.g., productivity) within each treatment type (e.g., the SRSD instruction that involved teaching students how to plan and draft narrative, expository, and persuasive text). PND were interpreted using criteria proposed by Scruggs and Mastropieri (2001; see earlier description for Graham, 2006a).

The studies that evaluated the effects of teaching students strategies for planning and drafting narrative and expository text (all of the studies used the SRSD stages of instruction to teach the target strategies) had large effects (both at post-treatment and maintenance), generalized to untaught genres, frequently established experimental control, and were of high quality. The following findings were for the SRSD writing instruction delivered to struggling writers in grades 2 to 8 and typical writers in grades 4

to 8. Teachers or researchers delivered all of the interventions. The median/mean PND for treatment/post-treatment were as follows: elements (100%/96%), productivity (95%/91%), and quality (99%/99%). These effects also maintained 3 weeks or longer after treatment had ended. The median/mean PND for maintenance were as follows: elements (100%/90%), and productivity (100%/86%).

There was further evidence that SRSD was effective. First, the results indicated that SRSD generalized to untaught genres. When generalized to an untaught genre, the median/mean PND for elements was 86%/85%. Second, the majority of the SRSD treatments also established experimental control (78%; 21/27). Finally, confidence can be placed in these findings as more than 74% (20/27) of the SRSD studies met at least 8/11 (72%) quality indicators (quality indicators included such criteria as establishing reliability for dependent variables, providing adequate participant information, and collecting treatment fidelity and social validity data; see Rogers & Graham, 2008).

In summary, there were 27 single subject studies that evaluated the effect of using SRSD instruction to improve students writing outcomes (i.e., number of story elements). As in other reviews, SRSD instruction resulted in large overall effects on a variety of writing outcome measures both at post-test (range 91% to 100% PNDs) and at maintenance (range 86% to 100%). The findings from this review also indicate that effects generalize to other individuals and writing genres. As in other reviews, this review also revealed that SRSD studies evaluated using single subject designs, published to date, have only been conducted by researchers or teachers.

SRSD study related to the current study. Very recently, researchers have investigated the effects of using individuals other than researchers or teachers to deliver

SRSD writing instruction (Reid, Luschen, & Lienemann, in press). Reid et al. conducted a pilot study in which one paraeducator was trained to deliver an SRSD writing intervention (POW and WWW) to three struggling second-grade writers. The researchers cited several reasons to use individuals other than students' own teachers to deliver an evidence-based writing treatment. First, the researchers stated that teachers are often unable to provide the individualized instruction required to help struggling writers learn the necessary skills. Second, the researchers suggested that there is an untapped/underutilized resource available within our schools: paraeducators. The authors hypothesized that this group (often nonprofessionals) could be trained to successfully deliver SRSD writing instruction to struggling young writers and that the outcome would positively impact students' writing performance (e.g., writing productivity, holistic quality, and inclusion of story elements).

The Reid et al. study (2009) evaluated effectiveness using a multiple baseline design across participants with multiple probes collected during baseline. One paraeducator delivered the instruction to one male and two female struggling second-grade writers (all Caucasian) in a rural elementary school in the Midwest. The paraeducator had a college degree and two years experience delivering individualized reading and math instruction in the school (no experience delivering SRSD instruction). Although tutor training was provided, total time in training was not reported. The authors collected baseline data, whereas the paraeducator collected post-treatment (or independent performance) data. Students were shown black and white picture prompts and asked to plan a story, include all the parts of a good story, and to write as much as they could. Maintenance data were also collected four and eight weeks following the

completion of the independent performance phase, however, only two data points for two students and one data point for one student were collected. The students' writing was evaluated for number of story parts included (range of 0 to 7), number of words written (range of 10 to 100), and holistic quality (range of 0 to 7).

The paraeducator-delivered SRSD writing intervention resulted in improved writing outcomes (at least three demonstrations) at post-test for number of story elements. As there were less than three data points during the maintenance phase (across all three writing outcome measures), PND and experimental control for maintenance data were not assessed. Experimental control was not established for writing productivity or holistic quality, although some improvements were made. These results are described next.

First, the students included more essential story parts in their writing during Independent Performance when compared to baseline. All students' baseline data were stable before the intervention was introduced. Following intervention (or independent performance), all students included more essential story elements. The mean PND also indicated a moderate effect (mean PND = 89%). Two of the three students had 100% nonoverlapping data points and one student had 66% nonoverlapping data. The student with a 66% PND showed a decreasing trend during Independent Performance, but received booster sessions. The booster sessions resulted in improved scores during the maintenance phase.

Second, although some progress was observed, experimental control was not established for writing productivity and holistic quality. The mean PND for writing productivity was 33% (0%, 100%, 0%). Additionally, only two of the three students' data

were stable during the baseline condition. The student in the third tier/leg had four ascending data points during baseline before the intervention was introduced. The mean PND for holistic quality was 67%. Two of the students had stable baseline data with 100% nonoverlapping data during Independent Performance, however the final student's data were unstable during baseline with a PND of 0%.

The researchers also collected treatment fidelity information. While teaching, the paraeducator used a checklist with step-by-step directions for each lesson. Additionally, one of the authors observed thirty-two percent of all sessions during instruction, using the same checklist to assess treatment fidelity. Treatment fidelity was reported as 97% (range not provided).

The Reid et al. study (2009) is extremely relevant to the proposed writing study. First, it shows that someone other than a teacher or trained researcher can effectively deliver structured SRSD writing instruction (97% treatment fidelity across 32% of sessions). Second, it also indicates that that a non-teacher/research-delivered SRSD writing treatment can have a positive effect on young struggling writers' writing performance (as measured by including more essential elements in stories). Finally, the results from this study (as well as the results from other meta-analyses described earlier) support the need to continue to research the effectiveness of similar writing interventions (e.g., interventions led by individuals who are able to assist, but to date have not typically been called on to deliver structured writing interventions). For example, it is still unclear whether adults without formal teacher training or tutoring experience and not employed by a school district were able to deliver a structured writing treatment as it was intended to be delivered, or whether students were to benefit from such instruction. The results

from recent studies (i.e., Reid et al., 2009) indicate that adults other than researchers or teachers can teach struggling young writers with positive effects. In an effort to promote the widespread use of effective strategies, more research should be conducted to examine the effectiveness of utilizing other adult tutors such as nonprofessional adult volunteers.

Conclusion

Future research on the effectiveness of nonprofessional adult-led literacy interventions should go beyond reading interventions to include writing interventions as well. First, there is abundant research to support the need to improve students' writing and the effectiveness of writing programs, such as SRSD instruction, that have only been led by teachers, researchers (Rogers & Graham, 2008), or paraprofessionals with tutoring experience (Reid et al., in 2009). Second, there is evidence that other literacy programs (none of the programs evaluated effects on writing other than on improving spelling, see Morris et al., 1990) delivered by nonprofessional adults are effective at improving reading outcomes for students in grades pre-Kindergarten through six. These findings can be used to shape and improve future research in which nonprofessional adults are used to deliver literacy-related instruction.

First, the findings from the 19 nonprofessional adult-led literacy studies suggest that adequate *tutor training be provided.* As indicated in an earlier section, the studies in the review provided little information on tutor training other than total time. Thus, I calculated differences based on time. Studies were coded as having two or fewer hours of tutor training or having more than two hours of tutor training in order to create two groups with equal numbers. The overall effect at post-test on reading comprehension

measures increased from -.32 (n = 7; range –1.27 to 1.31) when two or fewer hours of tutor training was provided to .18 (n = 4; range -.35 to .73) when more than two hours of tutor training was provided. Future researchers should describe tutor training in more detail, as total time in training may be less important than the extent to which tutors were trained to criterion.

Second, nonprofessional adult *tutors should be provided some sort of compensation* for their participation (e.g., free lunches or small hourly wage). The findings from the review suggest that providing tutor compensation may improve overall effects. For example, in the current review the overall effect at post-test on reading comprehension measures increased from -.47 (n = 4; range -1.11 to .50) when no tutor compensation was provided to .51 (n = 5; range -.35 to 1.31) when tutor compensation was provided.

Third, future research examining the effectiveness of nonprofessional adult-led literacy instruction *should include maintenance data*. Only 32 percent (6/19) of the studies reviewed examined effects several weeks after the treatment had ended. In the majority of the studies that collected maintenance data (4/6; 67%), effects diminished. For example, effects on reading comprehension decreased from -.14 (n = 11; range -1.27 to 1.31) at post-test to -.90 (n = 6; range -1.91 to .80) at maintenance.

Fourth, future research should more carefully examine program effectiveness as it relates to tutor-related characteristics. One way to further this research is to *provide better descriptive statistics for tutors*. In the current review, approximately 63 percent of the studies (12/19) provided total tutor numbers, but most did not provide any other information about the tutors. Less than half indicated whether the tutors had previous

tutoring experience (9/19; 47%). This information is important when discussing the external validity of a treatment.

Finally, the findings from the current review stress the need to conduct research that is of higher quality, using experimental designs better suited to research questions. In the current review, only 37 percent (7/19) of the studies met at least 5 of the 10 quality indicators. For example, the majority of the studies (more than 70%) did not report treatment fidelity. *Such information (e.g., treatment fidelity) should be included in future studies*. Additionally, the experimental design should be chosen to best answer research questions. For example, *future research should use an appropriate design* to demonstrate experimental control, even when large numbers (30 or more) are not available for testing. All of the 19 studies evaluated for this review used group designs although the majority (10/19; 53%) included fewer than 30 students in the treatment group. All recommendations for future research were addressed in the current study.

CHAPTER III

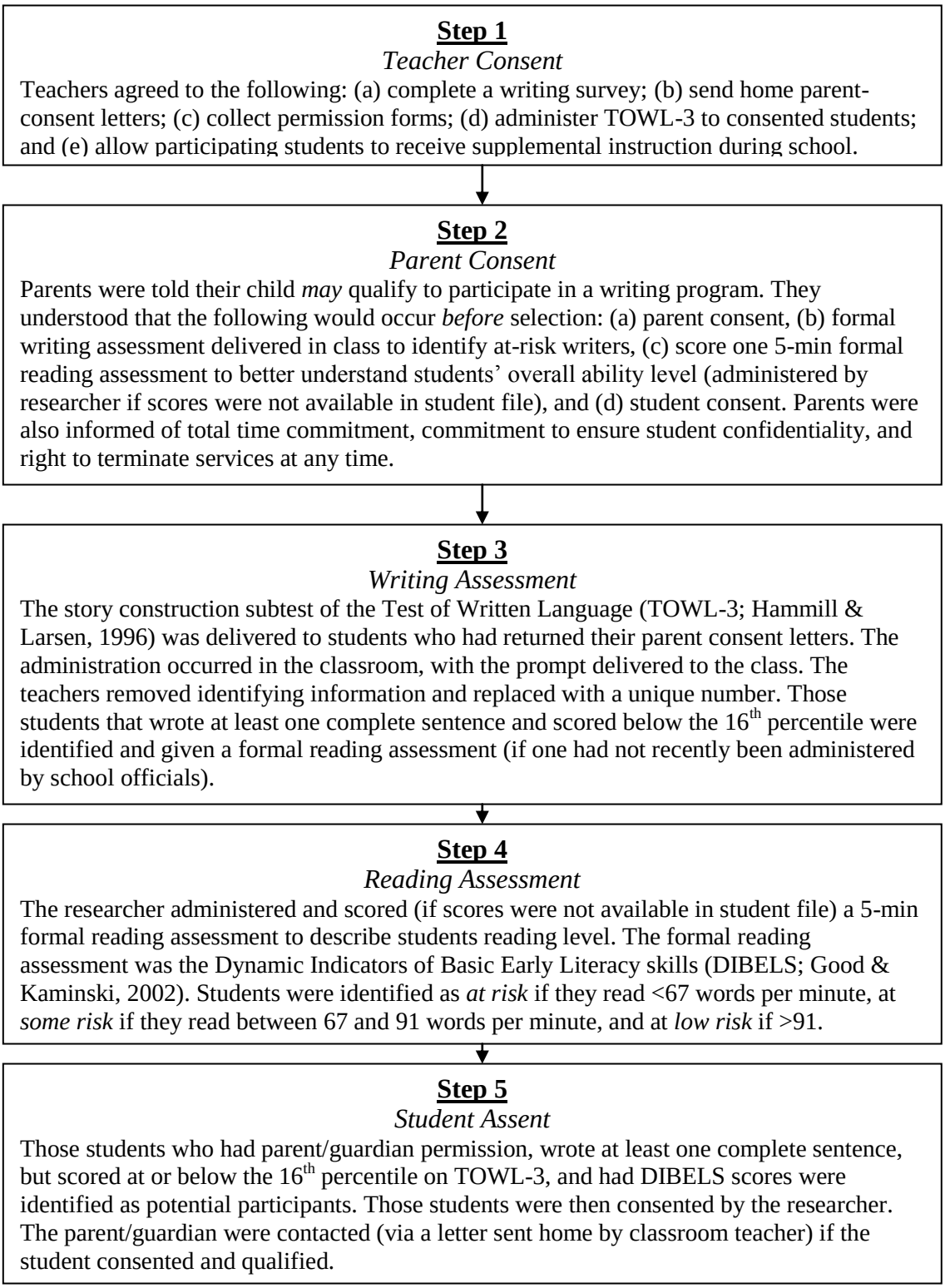
METHOD

Participants

Setting. The study was conducted with six struggling third-grade writers at a rural elementary school in Middle Tennessee. Each tutoring session was conducted during the school day. The elementary school was identified after receiving district approval and confirming principal interest. According to the 2009 Tennessee Department of Education (TDOE), the school houses approximately 300 students in kindergarten through fourth grade. The majority of the students at the elementary school were White (210/301; 69.8%) followed by African American (74/301; 24.6%), Hispanic (9/301; 3.0%), American Indian (0.3%), and Asian (1/301; 0.7%). Approximately 61 percent of students in the school participated in the free/reduced lunch program. A small percentage of students in the school were identified as needing special education services (6.7%).

Student identification. At the time of the study, the participating school had three third-grade classrooms. All three third-grade teachers were interested in participating in the study. Several steps were taken to ensure equitable selection of students and to preserve individual student's rights to confidentiality. The third-grade students in this study were selected based on several step-wise criteria (see Figure 1). The researcher was not aware of the students' identifying information until the end of step three.

Figure 1. Student Selection Procedures



First, their classroom teacher indicated a willingness to participate in the study (see Figure 1). Second, the parents/guardians provided consent for students to participate in the formal non-identifying assessments that could possibly allow their child to qualify for individualized writing assistance. Steps three through five are described in greater detail next.

During step three, the teachers administered the Spontaneous Writing Prompt from the *Test of Written Language-3* (TOWL-3; Hammill & Larsen, 1996) in the classroom to consented students. The results from this assessment were used to assess a student's ability to write a complete and interesting story. The test involves showing students a picture and giving them 15 minutes to write a story (assessed by examining if specific thematic elements were included in the story, see Graham et al., 2005). After identifying information was removed and replaced with a students' unique number, the papers written by students were given to the researcher to score. Identification criteria included writing at least one complete sentence and scoring at or below the 16th percentile on the TOWL-3. The TOWL-3 has good reliability at the third-grade level (.89), with moderate correlations with other writing measures. Two individuals scored the writing assessments.

Next, the researcher informed the teachers as to which writing prompts (still marked with non-identifying numbers) included at least one complete sentence and were scored at or below the 16th percentile. At this point, the teachers first confirmed that identified students were writing below grade-level in class and then identified the at-risk writers by name. One third-grade teacher was absent during this step due to an

unexpected medical leave. The remaining two teachers stated they were comfortable with providing this information as the third-grade team frequently collaborated together regarding their students (identified students from this classroom were Nick and Alexis, see Table 6).

The fourth step involved assessing the student's oral reading fluency. The oral reading fluency of students was assessed using the scores obtained through the school-administered Dynamic Indicators of Basic Early Literacy Skills assessment (DIBELS; Good & Kaminski, 2007; see description in Outcome Measures section in this chapter). I accessed the test scores from consented students' files that were administered by school officials the month before the study began (December, 2009). These scores were used for descriptive purposes.

The final step involved obtaining student assent and contacting parents/guardians. Those students who qualified were approached, with their teachers present, to determine their willingness to participate in the writing program (called the *Writing Academy*). The students were asked if they would like to participate in a writing program at their school. I told the students the approximate time commitment for participating in the Writing Academy, described the writing treatment, and identified program goals. The students signed an IRB-approved assent form stating that they wished to participate, but were also told that they were free to stop participating at any time.

These selection and exclusion procedures resulted in the identification of six struggling third-grade writers from three different third grade classes (see Table 6). Cecilia was a 9-year, 1-month old Caucasian female who had been referred for special

Table 6. Student Demographic Data

Student (TX Group)	Gender	Ethnicity	Age	DIBELS n^1 (Status)	TOWL-3 n^2 (%)
*Cecilia (1)	Female	Caucasian	9.1	49 (At-risk)	2 (9)
*Justin (1)	Male	Caucasian	8.5	94 (No risk)	0 (5)
*Michelle (2)	Female	Caucasian	10.3	79 (Some risk)	3 (9)
*Nick (2)	Male	Caucasian	10.1	59 (At-risk)	0 (1)
*Alexis (3)	Female	African American	9.9	86 (Some risk)	2 (9)
*Lucas (3)	Male	Hispanic	8.6	98 (No risk)	2 (16)

Notes. TX = treatment group; treatment was delivered one-on-one, groups refer to the order in which treatment was introduced. * = names provided are pseudonyms. n^1 = The median oral reading fluency score taken from the December 2009 school-administered Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment. Status = likelihood that a third-grade student has reading difficulties; < 67 = at-risk; 67-91 = some risk; > 91 = low risk. n^2 = The story construction subtest raw score taken from the December 2009 teacher-administered Test of Written Language (TOWL-3) assessment.

education services, but had not been identified. Justin was an 8-year, 5-month old Caucasian male. Michelle was a 10-year, 3-month old Caucasian female. Nick was a 10-year, 1-month old Caucasian male. Alexis was a 9-year, 9-month old African American female previously diagnosed with a language disorder and currently receiving speech services. Lucas was an 8-year, 6-month old Hispanic male. (All names provided are pseudonyms. See Table 6 for participant characteristics.) Two students were identified as

at-risk for reading difficulties (Cecilia and Nick). Two students were identified as having some risk for reading difficulties (Michelle and Alexis). Two students were identified as having no risk for reading difficulties (Justin and Lucas).

The students' writing scores on the Test of Written Language-3 fell at the 16th percentile or below. The students wrote between 12 (Alexis) and 69 (Michelle) words on their 15-minute classroom writing assessment (mean = 36.1). The majority of the students' writing scores placed them at the 9th percentile (Cecilia, Michelle, and Alexis) followed by the 5th percentile (Justin) and 1st percentile (Nick).

I assigned students to tutoring groups based on classroom teacher preference. This procedure was complicated by an unforeseen medical absence of Nick and Alexis's classroom teacher when students were assigned. It was decided that one of the teachers who was present would have her students, Cecilia and Justin, begin the first leg of the writing treatment with the male tutor (James, described later in Table 7). This decision was made based on this teacher's eagerness to have her students start the writing treatment and her desire to have her male student, Justin, paired up with a male tutor. As the return date of the remaining third-grade teacher was unknown, we decided to divide the second and third treatment groups between the two classrooms. This procedure allowed both teachers to have students receiving treatment during the second leg, as opposed to waiting for the third leg of instruction to begin.

In summary, several steps were taken to identify third-grade struggling writers (see Figure 1) and to place them in the appropriate treatment group. To identify students, first their teachers were consented. Second, the students' parents/guardians were consented. Third, the students wrote at least one complete sentence and scored at or

below the 16th percentile on the TOWL-3 writing assessment. Fourth, each student's oral reading fluency was measured using the DIBELS assessment. Finally, students agreed to participate in the writing program. To place students, a combination of random assignment and teacher preference were used. In the next section, I describe the methods used for recruitment and identification of the nonprofessional adult writing tutors and provide demographic information on selected individuals.

Nonprofessional adult identification. Access to potential nonprofessional adults occurred through the school's administrator. The school's principal identified individuals who frequently volunteered at the school in clerical roles (e.g., photocopying, cutting, assisting on field trips) and introduced the researcher to these individuals. Said individuals were given a copy of the IRB-approved consent form that provided detailed information on the (a) writing intervention, (b) the estimated time commitment, and (c) the specific teaching qualities of individuals sought. The writing intervention was described as a tutoring treatment aimed at improving story-writing skills of third-grade students. I informed individuals that the writing treatment had repeatedly been shown to have positive effects, but had only been tested using researchers and classroom teachers as instructors. I informed the potential candidates that the evidence would suggest that nonprofessionals (those without formal teaching experience or teacher training) would also be successful at improving students' writing. I did not provide further details about the SRSD model at this point.

Second, I outlined the total time commitment for participation. Nonprofessional adults would work with individual third-grade students during the school day. The writing instruction would occur in the school, 3 times per week, with each session lasting

approximately 30 mins. The estimated total time in instruction would be no more than 6 weeks or 18 sessions. The volunteers were also informed that the writing instruction would be very detailed with scripts and training provided. In order to feel comfortable in delivering the writing instruction, the researcher would provide an estimated 6 to 8 hours of training prior to the start of the study. It is important to note that this amount of training was based on the nonprofessional adults' lack of previous training and procedures used in previous SRSD studies in which adults were trained to a specified criterion (see Lane et al., 2008). The total time commitment for participation was outlined as approximately 17 hours. The nonprofessional adults were also told they would be given a small stipend for participating in the study (i.e., $50 for completing the 6- to 8-hour training, and $100 for tutoring each child).

Finally, the letter outlined tutor qualifications. Those individuals encouraged to apply were post high school graduates without formal teaching experience or training (e.g., had not been trained to tutor). The individuals could have third-grade children, but they would not be assigned to work with their own child. Additionally, individuals were told they needed to maintain student confidentiality. I indicated that up to eight nonprofessional adults would be able to receive the training, but no more than six tutors would be selected as writing tutors.

Identifying potential volunteers via the school's administration resulted in identification of one tutor. The school administrator gave me permission to recruit potential adults from a nearby community. This resulted in identification of two additional adult volunteers. I then interviewed each potential tutor in person. This interview is detailed next.

I met with each volunteer separately to both ensure they understood the intervention and time commitment and to assess whether they qualified to participate. At this point I read through the IRB-approved consent form with each individual and asked whether they agreed to maintain student confidentiality; provide demographic information (race, age, gender, educational background, tutoring experience); had a high school diploma or above, did not have a college education major, had no previous writing tutor experience; could attend a two ½-day training session; could tutor three times per week (30-min sessions) during a 3- to 6-week program; would audio tape each session; be observed weekly; be available weekly for discussion/feedback; maintain a daily journal; and attend a focus group at the end of the study. These selection and exclusion procedures resulted in the identification and placement of three nonprofessional adults (see Table 7 for a description of the nonprofessional adult tutors).

All individuals (n=3) who volunteered to participate in the current study met the inclusion criteria. The first volunteer, James, was a 60-year-old Caucasian male. James lived in the same community as the researcher and had heard about the Writing Academy directly from the researcher. James had a B.S. in Business Administration and a Masters of Divinity, no tutoring experience or educational training, and was a retired military chaplain. He and his wife did not have any children.

The second volunteer, Jessica was a 41-year-old Caucasian female. Jessica also lived in the same community as the researcher and had similarly heard about the Writing Academy directly from the researcher. Jessica had a B.S. in nursing and was currently not employed outside of her home. Jessica did not have any tutoring experience outside of

her experience with her own children who were in the second and third grade. These children did not attend the same school in which the study occurred.

The third volunteer, Phyllis, was a 41-year-old Caucasian female. Phyllis was identified as a possible tutor by the school's principal. Phyllis had a B.S. in marketing and was currently not employed outside of her home. Phyllis spent several hours per

Table 7. Nonprofessional Adults Demographic Data

Tutor (TX Group)	Gender	Ethnicity	Age	Background Education Tutoring Experience (exper.) Current Occupation
*James (1)	Male	Caucasian	60	B.S. – Business Admin. M.Div No previous tutoring exper. (retired)
*Jessica (2)	Female	Caucasian	41	B.S. – Nursing No previous tutoring exper. (not currently employed)
*Phyllis (3)	Female	Caucasian	41	B.S. – Marketing No previous tutoring exper. (not currently employed)

Notes. TX = treatment group; treatment was delivered one:on:one, groups refer to the order in which treatment was introduced. * = names provided are pseudonyms. B.S. = Bachelor of Science degree. M. Div = Masters in Divinity.

week volunteering at her children's school primarily assisting with clerical help (e.g., photocopying and cutting material). Phyllis also assisted as a "coach" or "co-manager"

(with one third-grade teacher) in an after-school program called *Destination Imagination (DI)*. The DI program occurred weekly and was a forum in which a small group of third-grade students worked in teams to solve challenging problems and compete against other teams at tournaments. At the tournaments teams are required to think quickly, work together, and come up with original solutions to satisfy challenge requirements (see http://www.idodi.org/ for a more complete review of DI program). The one-hour weekly after-school sessions were meant to help prepare students for the tournament. Phyllis did not receive any training to "coach" this group, stating that her role was primarily to provide feedback, purchase necessary materials (to build props which would be used at tournaments), and manage behavior (e.g., reducing off-task behaviors). Phyllis's two children (kindergarten and third-grade) attended the school in which the study occurred. Phyllis's third-grade student did not participate in the current study.

The three tutors were assigned to treatment groups based on their volunteer date. Although Phyllis was approached first, she did not officially commit to tutor until after James and Jessica had volunteered. The first individual to volunteer (James) was assigned to the first tutoring group (Cecilia and Jacob). The next individual to volunteer (Jessica) was assigned to the second tutoring group (Michelle and Nick). The third individual to volunteer (Phyllis) was assigned to the final tutoring group (Alexis and Lucas).

In summary, there were several steps taken to identify suitable nonprofessional adults for this study and to place them with their tutees. First, the school's principal identified possible candidates. Those individuals were approached and identified using set inclusion and exclusion criteria. Next, the school's principal gave permission to recruit suitable tutors from a nearby community. These individuals were also approached

and identified using criteria described above. Third, prior to starting the tutoring sessions, the nonprofessional adults received training and demonstrated an ability to teach each lesson. Finally, tutors began working with students who had been assigned to groups by classroom teacher preference. In the next section, I describe the treatment. I detail the general instructional procedures related to (a) SRSD story-writing instruction, (b) preparation of nonprofessional adults, and (c) procedural fidelity.

Treatment

General instructional procedures. The purpose of this study was to extend what is known about the effectiveness of SRSD instruction. Did the intervention improve struggling students' writing, when nonprofessional adult volunteers delivered the instruction? The story writing intervention was consistent with that described in previous studies (see Lane et al., 2008). However, the procedures in the current study did not involve specific behavioral reinforcement (such as providing positive behavioral support tickets), as was done in some prior studies (Lane et al., 2008) as I did not anticipate that students in this study would need this type of reinforcement to be successful. The procedures involved SRSD story-writing instruction, nonprofessional adult training, and procedural fidelity. These are described next.

***SRSD story-writing instruction*.** The nonprofessional adults taught each tutee using eight scripted SRSD story-writing lessons (see Lessons 1a to 7 in Appendix B). The following description of SRSD is organized according to the six SRSD instructional stages (e.g., Develop and Activate Background Knowledge, described next). It is important to note that although SRSD stages and lessons are related (progression in both

related to quest for independent performance), but more than a single SRSD stage is typically covered in each lesson. For example, in lesson 1a the nonprofessional adults Develop and Activate Background Knowledge and Discuss It. The SRSD stages covered in each lesson are highlighted at the top of each lesson plan (see Appendix B). Next, I provide a brief overview of SRSD and then provide detailed information related to each SRSD instructional stage. More detailed information about SRSD can be found in Harris et al. (2003).

General overview of SRSD. With SRSD, students were explicitly taught strategies and self-regulation procedures needed to successfully manage the writing task (Harris et al., 2003). Before students were taught target strategies, they were also provided the tools needed to be successful (e.g., knowledge of related vocabulary; understanding of benefits of using the strategies, and commitment to becoming a better writer). More specifically, students were explicitly taught the mental operations needed to successfully manage and complete specific writing tasks (i.e., the writing strategies) and self-regulation procedures (such as goal-setting, self-monitoring, self-reinforcement, and self-instructions) needed to use the strategy, manage the writing task or their writing behavior. Throughout these activities, students collected concrete evidence of their progress and set goals for what they would do.

SRSD instruction was scaffolded with teachers gradually shifting responsibility for activating strategies and applying skills from teacher to student (see Harris et al., 2003). SRSD instruction was structured through the use of detailed lesson plans (which served to guide the instructors actions). Students progressed through these lessons at their own pace and were provided with individualized levels of support and feedback (as

determined through active collaboration).

Instruction involved the following six stages: *Develop and Activate Background Knowledge, Discuss it, Model it, Memorize it, Support it, and Independent Performance*. Teaching persisted until the student was able to demonstrate ownership and independent use of targeted writing and self-regulation strategies. In this study, students were taught how to write stories using SRSD (see Appendix B for lesson plans).

The students were first taught a general planning and drafting strategy for writing. There were three steps in this strategy represented by the mnemonic POW. The first step in POW was ***Pick*** an idea for writing a story. The next step was ***Organize*** notes (students were taught how to do this and provided a graphic organizer to aide them [see WWW… below]). The final step in POW was ***Write*** and say more (which reminded them to continue planning and drafting while writing). To help them carry out the *O* in POW (*Organize notes*), the students were taught a separate genre-specific strategy for story writing. This strategy was represented by the mnemonic, W-W-W, What = 2, How=2. The nonprofessionals prompted the student to generate ideas and make brief notes for the following questions (related to basic story elements): ***W***ho is the main character? ***W***hen does the story take place? ***W***here does the story take place? ***What*** does the main character do or want to do; what do other characters do? ***What*** happens when the main character tries to do it; what happens with other characters? ***How*** does the story end? and ***How*** does the main character feel; how do other characters feel? The SRSD instruction covered seven specific stages (e.g., *Develop and Activate Background Knowledge*). These stages (and related activities) are more fully described below.

SRSD instructional stages. The first part of SRSD instruction involved the stage,

Develop and Activate Background Knowledge. The nonprofessional adults introduced related vocabulary and concepts to lay the groundwork for learning and using both the general writing strategy and story-writing planning and drafting strategy. The nonprofessional adults (a) built rapport; (b) introduced and discussed strategies (or tricks) for better writing (general and story writing); and (c) used graphic organizers and stories to help make these concepts more concrete.

First, as the students were unfamiliar with the nonprofessional adults, time was spent building rapport. During this time, the nonprofessional adult asked the student questions about their current writing methods (or approach) and generated excitement about learning tricks that would make writing more fun and interesting to write and read. The nonprofessional adults also asked the students to commit to putting forth their best effort to learn and use the strategies. Next, the trick for writing better was introduced.

The trick for better writing (in general) was taught using the mnemonics POW and WWW. As stated earlier, POW stood for *Pick* my idea(s), *Organize* my notes, and *Write* and say more. Next, the trick for better story writing was introduced. This was taught using the mnemonic WWW, What = 2, How = 2 (also described above). The nonprofessional adults further introduced the concept of using "million dollar words" (MDWs) in writing to make writing more fun to read. The students were told that MDW were vocabulary words that are interesting and/or colorful and make a story more exciting.

When discussing each of these tricks, the nonprofessional adults encouraged the student to become an active participant in the learning process (see Harris et al., 2003). In addition, the nonprofessional adults and students collaborated; working together to

identify the essential parts of a story (WWW) in reading material, including writing brief notes for each part in a story-parts graphic organizer (identical to the planning and drafting organizer for WWW used when writing stories). This helped to ensure students were familiar with each part of a story. There was both discussion and repetition examining stories as well as generating possible story parts until it was certain that students had mastered these concepts. Students also examined and read stories to identify other important characteristics of good writing (e.g., MDW). For example, in the first lesson, the nonprofessional adults first reviewed story parts (WWW) and use of interesting words (MDW) and then read the story, Albert, out loud. On the second read through, the nonprofessional adults asked the students to say when he/she heard one of the story parts or MDWs (see Lesson 1a in Appendix B). As SRSD instruction is recursive, the nonprofessional adults continued to review POW, WWW, What = 2, How = 2, and MDW over the next several lessons (and four stages of instruction), spending time to discuss what each part stands for and their overall importance (see Appendix B).

During stage two, *Discuss it*, the nonprofessional adults and students further discussed the importance of using POW and WWW when writing stories. The nonprofessional adults further introduced the term, *transfer*, and discussed how some of the to-be-learned material (e.g., POW) could be used in situations outside of the Writing Academy. The nonprofessional adults told the student that it was the student's responsibility to come to every Writing Academy lesson once they learn how to use the strategy with at least two examples of how they were able to transfer the writing strategies to other writing tasks (e.g., using POW when prompted by their classroom teacher to write about their weekend). Again, as SRSD instruction is recursive, the

mnemonics POW and WWW were rehearsed, their parts discussed, and concrete examples provided (e.g., looking for parts in stories and recording brief notes on graphic organizer). In the next two sections, the SRSD stages *Model It* and *Memorize It* are described in more detail, although several of the activities described to this point have already demonstrated examples of each (e.g., rehearsing POW and WWW).

The third SRSD instructional stage was *Model It*. During this stage, the nonprofessional adults provided repeated opportunities for the student not only to observe how to plan and write a better story, but also to participate in the process, mainly through generating ideas for the story. In other words, the nonprofessional adults led the process, demonstrating how to use the POW and WWW strategies for story writing, but frequently solicited students' input. The modeling was natural and enthusiastic, using language familiar to the child (see Harris et al., 2003). The modeling activities were drawn from similar studies (c.f., Danoff, Harris, & Graham, 1993; Lane et al., 2008). The nonprofessional adults modeled how to (a) define the problem (e.g., "What is it I have to do? I have to write a good story. A good story has to have all 7 parts and make sense."), (b) plan a story (e.g., "Remember that the first letter in POW, P, is for pick my idea."), (c) use coping statements (e.g., "If I just take my time, a good idea will come to me."), (d) self-evaluate (e.g., "Have I transferred all my notes into my story?"), and (e) self-reinforce (e.g., "I have 7 good parts. Now I can blast off my rocket!"). The nonprofessional adults told the student that saying things out loud *and* silently in one's head helps with the writing process.

To model, the nonprofessional adults used black and white line drawing (similar to those used in baseline), students' self-statement sheets, and the POW and WWW

graphic organizer (see Lesson 2 in Appendix B). The nonprofessional adults first modeled (out loud) how to identify the problem and set goals for writing. For example, the nonprofessional adults looked at the black and white drawing picture with the student and said something such as, "What is it I have to do? I have to write a good story. A good story should be interesting, have all seven story parts, and be fun to read and write."

The nonprofessional adults then asked the students to help him/her plan and use coping statements. To begin with, the nonprofessional adults referred the student back to the now known POW strategy (e.g., "Remember that the first letter in POW is to come up with a good idea for a story. Can you look at this picture with me and help me come up with a good idea for a story?"). If the student had problems generating ideas, the nonprofessional adults modeled using coping statements (e.g., "I can't think of anything to write! Ok, if I just take my time, a good idea will come to me."). The nonprofessional adults helped students generate one to two coping statements that were most helpful for students and wrote these down on their self-statement sheets (see Lesson 2, Appendix B).

Next, the nonprofessional adults modeled how to plan when writing a story by showing students how to organize their ideas (the O in POW) using the specific trick for story writing (WWW, What = 2, How = 2). Throughout this process, they continued to model how coping statements can aid in the writing process. Together, the students and nonprofessional adults used the WWW graphic organizer and, in no particular order, generated ideas for all seven story parts (e.g., "What ideas do I see in this picture? Can you help me? What do we think we should name this character?" [who], "When do you think it happened?" [when], "Let's see, where to you think the story took place?" [where]). During this modeling the nonprofessional adults used more coping statements

(e.g., “I have so many ideas. It’s so hard to pick just one. I know, I’ll put one idea down for each of the seven parts and add more later if I still have time.”).

After the students and nonprofessional adults wrote brief notes for each of the seven parts, the nonprofessional adults modeled self-evaluation procedures. He/she modeled how to look back at the notes and decide whether more could be added to any of the story parts to make the story more interesting to read. The nonprofessional adults also modeled how to assess whether any million-dollar words (MDWs) were used. The nonprofessional adults then modeled how planning aided the writing process by showing students how they could now use their brief notes to write and say more (the W in POW).

While writing and saying more, the nonprofessional adults frequently referred to the completed WWW graphic organizer and the students’ self-statement sheets. He/she talked himself/herself through writing a story (again, the student was constantly encouraged to assist in the process although during this stage the nonprofessional adults were still taking the lead). More specifically, the nonprofessional adults pulled out a clean piece of paper and said something such as “How shall I start? I need to tell who, when, and where.” During the process of writing, the nonprofessional adults also modeled how ideas can be added as one writes (e.g., “Wait, I just thought of a MDW that would make that sentence so much more fun to read!”). He/she also modeled how to use encouraging and positive self-statements such as “Good work, I’m done. It’ll be fun to share my story with others.”

After the story was written, the nonprofessional adults again modeled self-evaluation by asking the students to determine whether they had used all their planning tricks when writing their story. They asked the students to identify the trick they used to

get started (e.g., " What is it we said to ourselves when we started? We asked ourselves what it was we had to do. We said we had to write a good story that had all seven parts, was fun to read and make sense?). They also asked the students what trick they had started with that could be used for all writing (e.g., "How did we start writing our story? Yes, we had to pick an idea, organize our notes, and then write and say more."). Finally, they asked the students what trick they had used to write a good story that has all seven parts (e.g., "How did we organize our notes so that we could write a great story? Yes, we wrote brief notes for WWW, What = 2, How = 2. Then we did the W in POW. We wrote and said more!"). In the later lessons (see Lesson 6, Appendix B), the nonprofessional adults also modeled how organization can occur even when pre-made WWW graphic organizers are not available. Throughout all the lessons, the nonprofessional adults continued to model how coping statements and self-evaluation can make story writing easier and result in better stories.

The nonprofessional adults also modeled self-monitoring and recording procedures by showing students how to graph their stories on a rocket. By doing so, the nonprofessional adults showed students a fun way to assess whether each story had all seven parts and million dollar words. The rockets were hand drawn and divided into seven equal parts. There were six stars drawn around each rocket. The nonprofessional adults modeled locating each part in their writing and coloring in the corresponding part on the rocket. They also modeled how to fill in one star per million-dollar word included in the story. Finally, the nonprofessional adults modeled how they were able to blast off the rocket if each of the seven parts had been filled in.

Although the nonprofessional adults decreased the amount of modeling they

provided during lessons three through six, they continually monitored the students and provided whatever modeling was necessary so that students could write good stories with seven story parts. In other words, although it was important to allow the student to lead as much as possible as the lessons progressed, the nonprofessional adults modeled as much as was necessary until the student could write a story independently (see Lesson 6, Appendix B). The nonprofessional adults received guidance (at least weekly) from the researcher as to the amount of modeling that was appropriate.

The next SRSD instructional stage was *Memorize It*. As stated earlier, the activities described thus far provided numerous ways in which the nonprofessional adults helped the student memorize the tricks for writing better stories (POW and WWW strategy). Throughout each lesson, students had multiple opportunities to practice reciting the strategies (e.g., WWW strategy) and discussing their importance. For example, in an early lesson, the nonprofessional adults told the students that POW was an important trick to learn because good writers often use it for many things they write (see Lesson 1a, Appendix B).

Students were also tested at the beginning of every lesson (beginning with lesson 1b, see Appendix B) on the mnemonics and their meanings. At first this was completed orally, but as the students progressed, the nonprofessional adults asked them to write their WWW trick down the side of a blank piece of paper. This practice allowed students to have numerous opportunities to practice making their own planning sheet, even before it was necessary for them to stop using the pre-made graphic organizer while writing stories The nonprofessional adults also helped the student practice POW and the WWW tricks using pre-made flashcards. The activities to improve recall were adapted so that they fit

each student's unique learning styles (e.g., some students enjoyed timed games whereas others did not). For example, one student had a particularly difficult time remembering the What = 2. The nonprofessional adult came up with a trick of putting a post-it with WHAT written boldly across the top on her head and retrieving the student from class with the post-it note adhered. From that day on, the student was able to recite the What = 2 part of the trick without any verbal (or visual) assistance. Students were verbally praised for remembering some or all of the parts and were encouraged to continue to practice rehearsing outside of the Writing Academy.

Along with helping the student memorize the writing tricks (POW and the WWW tricks) and modeling what the tricks looked like in action, the nonprofessional adults encouraged students to write with minimal assistance. During stage six, *Support It*, the nonprofessional adults provided the students with the support needed to learn to apply POW and the WWW strategy as well as the self-regulation procedures independently and correctly. As SRSD stages are recursive, there was much interchange between the *Model It* and *Support It* stages. For example, the nonprofessional adults first modeled how to plan a story together using POW, WWW, MDW, the graphic organizer, and the students' self-statements sheet, but during this planning and drafting, the students were encouraged to take the lead whenever possible (e.g., students generated ideas for brief notes while nonprofessional adults wrote on graphic organizer). If the student was unable to do so without assistance, or did so incorrectly, the nonprofessional modeled the correct procedures again.

There were several lessons dedicated to supporting the students as they progressed toward independent use of the taught strategies and self-regulation procedures

(see Lessons 3 to 6 in Appendix B). Throughout these lessons, the students were encouraged, as much as possible, to independently apply these processes and procedures. These lessons were repeated as often as needed until the student could apply the strategies and self-regulation procedures correctly and without assistance or the use of any of the graphic organizers. During each lesson during the *Support It* stage, students were encouraged to write a good story with all seven parts, to make the writing more interesting by including more MDWs, to use their coping statements, to review their work, and to praise themselves for goals achieved.

An example of how students were supported in their efforts at self-monitoring is found in Lesson four (see Appendix B). By Lesson four, the students had had multiple opportunities to identify the essential story elements and MDWs in read stories (e.g., Smokey, see Appendix B). Now the student was given the opportunity to read his/her own work that was written during the baseline phase. He/she found the parts and MDWs in his/her own story. The students were given an opportunity to rewrite the baseline paper and were verbally praised for their work. This activity provided another chance for the student to self-evaluate, seeing the progress he/she had made (e.g., comparing baseline performance to performance during the *Model It* and *Support It* stages), and practice taught strategies and self-regulation procedures. The repeated opportunities of supported practice at using the taught strategies and self-regulation procedures allowed the students to eventually move toward independent performance.

The final SRSD instructional stage was *Independent Performance*. During this stage, the goal was for the students to use strategies independently and write a good story without assistance. At the beginning of this stage, the nonprofessional adult asked the

students about their goals and transfer efforts, and answered questions about graphing, but all writing was done independently. The nonprofessional adults no longer provided the graphic organizer, but encouraged students to make their own. Instruction ended when the student demonstrated competence in the following areas. First, the student demonstrated that he/she could plan a good story. Without using any prompts (e.g., graphic organizer, self-statement charts) or receiving any assistance from the nonprofessional adult, the student used POW and WWW, What = 2, How = 2 to generate ideas for a story. Second, the student wrote a story that included the seven essential story parts (e.g., what a character wanted). During the stage, the nonprofessional adults also encouraged the students to use covert self-statements (e.g., saying self-statements quietly to self as opposed to talking about them out loud). The nonprofessional adults continued to praise students for their efforts and their ability to work independently.

The students in this study progressed through the six SRSD stages (eight lessons, see Appendix B) in 12 to 17, 30-min lessons. Prior to the intervention, the researcher provided adequate training related to correctly and successfully administering SRSD instruction (see Appendix B). This training is described next.

Preparation of nonprofessional adults. Training was provided prior to the start of the study (pre-baseline phase). The goal was to have the nonprofessional adults feel well prepared to teach the story-writing lessons. The training was nearly identical (positive behavior support tickets not provided in the current study) to that described in previous SRSD studies (see Lane et al., 2008). The training differed slightly in that there was more of an emphasis on aspects of good instruction (Brophy, 2001 as cited in Gall, Gall, & Borg, 2003), as the current writing tutors did not have teaching backgrounds or any

formal teacher training. Specifically, the preparation/training covered the following topics: (a) SRSD, (b) good instruction, and (c) story-writing lesson plans.

During the intervention phase of the study, three nonprofessional adults delivered the SRSD story-writing instruction (see lessons 1a to 7 in Appendix B). The writing instruction was referred to as the *Writing Academy*. The researcher trained the nonprofessional adults one week before they began teaching students. As each instructor began instruction at different times (intervention was time-lagged due to study design), all training sessions were provided on an individual basis. Total pre-treatment training time varied across tutors: James (8 hours), Jessica (6 hours), Phyllis (6 hours). Training occurred in various locations depending on tutor availability. I trained James and Jessica at my home. I trained Phyllis at the school in which the study occurred. Total training time and training procedures were similar to those reported in other SRSD studies (see Lane et al., 2008).

During the training sessions, the nonprofessional adults were first presented with information related to SRSD instruction (what it is) as well as its successful effects on improving students writing. This was done through lecture, discussion, and by watching a video of its application. Second, participants engaged in a discussion about elements of good teaching (e.g. enthusiasm, being supportive, providing an opportunity to learn) and were told that part of their weekly feedback would pertain to these elements. Finally, the researcher trained the nonprofessional adults to teach the story-writing lessons, giving them opportunity to practice to mastery and providing both verbal and written feedback. An example of how this was accomplished is provided next.

Instruction related to teaching the story-writing lessons occurred during the last

phase of the training. During this final phase, participants were given a binder with the seven story-writing lessons. The notebooks contained directions for all teaching activities as well as a checklist for each lesson. The checklists covered each component of instruction and were provided so that the nonprofessionals could check them off during instruction. Although each lesson was scripted, the nonprofessional adults were told that they could have some flexibility in teaching individual students. More specifically, the nonprofessional adults were told to cover the areas in each lesson related to the 6-stage SRSD instruction (e.g., Memorize it; testing students at the beginning of the lesson and reminding them of another test at the end of the lesson), but told that they did not have to read the script verbatim. This afforded the nonprofessional adults some flexibility in teaching individual students. The adults were able to reorder or repeat steps as necessary. They were also able to adjust their feedback and levels of support to meet each child's unique needs.

Specific training in the story-writing lessons began with the researcher modeling lesson 1a. The nonprofessional adult was then given a chance to model this lesson with the researcher acting as a "student". This was done to familiarize the nonprofessional with the lesson and help him/her have a better idea of appropriate pacing. After the nonprofessional delivered the lesson, the researcher provided immediate feedback on their performance. In providing feedback, the researcher addressed the elements of good teaching used and the areas covered and/or omitted from the lesson plans. These steps were repeated for each of the lessons. This area of training was discontinued when the adult participant was able to demonstrate mastery for each lesson. Mastery was defined as including at least 90% of all lesson components (e.g., practicing finding the story parts in

a read story), demonstration of a supportive and enthusiastic learning environment (e.g., smiles and uses encouraging statements), and identification of the essential writing and story writing elements (POW and *W-W-W, What = 2, How = 2*). All tutors successfully demonstrated 100% of all lesson components and displayed supportive and enthusiastic behaviors. At the end of training, the nonprofessional adults were asked to rate their perception of preparation. All tutors stated they felt prepared, but expressed concerns about actual lesson delivery, as they did not have teaching experience.

The nonprofessional adult volunteers were also provided with rich support during the intervention phase of the study. That support is delineated in the Results section (see Table 10). The high levels of support were provided in part to improve procedural fidelity.

***Procedural fidelity*.** All sessions were audiotaped and reviewed to determine if lessons were executed as planned. Additionally, the researcher observed at least 25% of instructional lessons delivered by each nonprofessional adult. A checklist containing step-by-step instruction for completing each lesson was used to determine presence or absence of each instructional component (see Appendix C). The degree of fidelity was determined for each lesson by dividing the total number of steps completed divided by the total number of steps possible. For each participant I provide the mean, number, and range of these percentages. Interrater reliability was conducted on at least 25% of each student's audiotapes and was calculated by dividing the components observed by the components possible and multiplying the quotient by 100. Interrater reliability for each student ranged from .92 to 1.00 with an overall average of .98.

Measures

There were several measures collected during the pre-baseline and treatment phases (baseline, treatment, post-treatment, maintenance). First, during the pre-baseline phase, the previously described screening instruments (DIBELS and TOWL-3) were used to describe and identify student participants, respectively. During this phase, the current classroom writing practices were also evaluated using responses given on the Teacher Survey of Classroom Writing Practices (see Appendix A). Second, during the treatment phase of the study, the researcher collected both writing outcomes and social validity measures.

Pre-baseline Measures. The pre-baseline measures were DIBELS, TOWL-3, and the Teacher Survey of Classroom Writing Practices.

Dynamic Indicators of Basic Early Literacy Skills. The Dynamic Indicators of Basic Early Literacy Skills assessment (DIBELS; Good & Kaminski, 2007) were delivered to all third grade students as part of school-wide quarterly testing procedures. The most recent DIBELS assessment occurred one month prior to study onset. DIBELS is a brief standardized assessment measure designed, in part, to assess oral reading fluency of connected text. The student reads from a standardized set of passages (allowed one minute per passage) with outcomes typically used for identification of students who may need more instructional support in reading and to monitor progress toward instructional goals (DIBELS is typically administered multiple times during the school year). The number of correct words per minute from the passages (typically the middle score from three passages) is the oral reading fluency rate. The measure has high concurrent validity with coefficients ranging from .91 to .96 and high alternate-form

reliability ranging from .89 to .96 across the 8 reading passages (Good & Kaminski, 2007). This assessment was used to better describe the student's ability.

Test of Written Language-3. The classroom teachers administered the Story Construction Subtest from the TOWL-3 to all students who had parent permission to participate in the study. This measure is appropriate for this study because it assesses the ability to successfully write a complete and interesting story (as determined by inclusion of specific thematic elements). Students were given a sheet of paper with a picture that served as the story prompt, were encouraged to pre-plan, and were provided 15 min to compose their story (see Lane et al., 2008). The Story Construction Subtest is reliable at the third-grade level (r = .89; Hammill & Larsen, 1996), is moderately correlated to other measures of writing, and has been shown to differentiate between strong and weak writers.

***Teacher Survey of Classroom Writing Practices*.** After classroom teachers agreed to participate in the study, they were asked to complete a Teacher Survey of Classroom Writing Practices survey (see Appendix A). This survey was adapted from Cutler and Graham (2008) by Lane et al. (2008). In the first section of the survey (41 questions), the teachers were asked to rate their overall use of writing activities and practices that were commonly recommended in teaching writing to young students (e.g., "Circle how often you conference with students about their writing"). The teachers rated their use of these activities using a 7-point Likert-type scale with responses ranging from 0 (*never*) to 7 (*several times a day* or *always*; see Appendix A). In the second section, the teachers were asked to write a brief description of their writing program. They were asked to identify the approach that best described their writing instruction (e.g., traditional skills

approach). This information was used to better describe the writing practices that were occurring before treatment began.

Survey answers indicated the three teachers addressed the following writing skills from *several times a week* (score of 5) to *several times a day* (score of 7): sentence construction, handwriting, spelling, grammar, punctuation, and capitalization. Additionally, each teacher indicated they *almost always* or *always* had multiple instructional goals for writing and had students use a graphic organizer when writing. The third-grade teachers were also unanimous in *never* (score of 0) addressing the following behaviors or addressing them only *several times a year (*score of 1*)*: assigning writing homework, using rubrics to evaluate writing, writing at home with parents help, asking parents to listen to something their child wrote at school, allowing students to dictate compositions to someone else, and using computers during the writing period.

Of particular importance to this study was the frequency study participants were taught how to (a) *plan and draft* (Questions 3 and 11), (b) *use strategies* related to writing skills or processes (Questions 18 and 21), (c) write via *teacher modeling* (Questions 19 and 20), and (d) *evaluate* writing using similar procedures as were used in the current study (Questions 25-28 and 41). Across the classrooms, students were specifically taught how to *plan before* writing and *plan and draft* approximately *several times a month* (n = 6, mean = 3.2, SD = .98). The frequency ranged from weekly (Cecilia and Justin's teacher), to several times a month (Michelle and Lucas's teacher), to monthly (Alexis and Nick's teacher). The teachers taught students how to plan and draft at least weekly (Cecilia, Justin, Alexis, and Nick's teachers) to monthly (Michelle and Lucas's teacher).

On average, students were taught *strategies* related to writing skills or processes

on a weekly basis (n = 6, mean = 3.7, SD = 2.3). Answers varied across classrooms. The frequency of this instruction ranged from several times a year / monthly (Alexis and Nick's teacher) to several times a week / daily (Cecilia, Justin, Michelle, and Lucas's teachers). Michelle and Lucas's teacher indicated she taught these skills and strategies, but rarely re-taught them (monthly).

The average frequency with which *modeling* of writing strategies and excitement about writing was used in the classroom was several times a month (n = 6, mean = 3.3, SD = 2.3). The frequency of modeling varied across classrooms. The participants' teachers modeled writing strategies several times a year (Alexis and Nick's teacher), weekly (Michelle and Lucas's teacher), or several times a week (Cecilia and Justin's teacher). They modeled excitement about writing never (Alexis and Nick's teacher), weekly (Michelle and Lucas's teacher), or daily (Cecilia and Justin's teacher).

Participants' teachers used similar *evaluation procedures* (e.g., reviewed previous writing) approximately several times a month (n = 12, M = 2.8, SD = 2.5). The frequency with which teachers used rubrics to evaluate writing was low and ranged from never (Justin and Cecilia's teacher) to several times a year. Students were encouraged to monitor their own progress weekly (Alexis and Nick's teacher) or daily (Cecilia and Justin's teacher). Alexis and Nick's teacher infrequently monitored students' writing progress. The other two teachers did this weekly (Michelle and Lucas's teacher) or daily (Cecilia and Justin's teacher). Only Michelle and Lucas's teacher used writing portfolios and had the students review previous writing samples. The other teachers indicated this was never done.

Students' teachers did use procedures that were generally similar to the ones in

the SRSD treatment (e.g., teaching writing strategies, planning and drafting, modeling, and evaluations). In discussion with teachers, however, there was no evidence that they used the exact same procedures as employed by the nonprofessional tutors.

Finally, all third-grade teachers indicated they used a traditional skills approach combined with process writing when elaborating on her writing program. Cecilia and Justin's teacher wrote: "For most writing assignments, I provide prompts. I model the product I am looking for, stressing structure and providing ideas. First, the students brainstorm ideas and then develop a graphic organizer (Day 1). The next day they begin a rough draft. After completing the rough draft (day 2 and 3) they revise and share with a buddy. By the end of the (typically) 4 or 5 days, the students have their final drafts. I make corrections at that point and conference with each student. Usually the final draft is then transferred by the student to a paper for display or to be types in the computer lab.

Alexis and Nick's teacher wrote: "I usually provide a prompt for the students. They are required to brainstorm and produce an organizer to show their process and procedures. Students write rough drafts. I check over, make notations of changes and mistakes, then talk over the rough draft with the student. They then rewrite final drafts."

Michelle and Lucas's teacher wrote: "The children having writing folders where they keep all writing, prewriting and drafts. Students list ideas in their folders. Students are taught to pre-write, draft, edit and finalize. If they finish the class assignment, they are to write a *free choice* story. We model writing on the ELMO, students share stories, and we also do shared writing. Students practice commenting on one another's stories in a *constructive* format."

Treatment measures (Baseline, Treatment, Post-treatment, Maintenance). Writing prompts were collected during the baseline through maintenance phases. Social validity was assessed during instruction (nonprofessional adults) and after the treatment had ended (students).

***Writing prompts*.** Writing prompts were collected during all four phases of the intervention (baseline, treatment, post-treatment, and maintenance). Identical procedures for collecting the writing probes were used during all phases of the study. The researcher, unfamiliar to each student, collected all writing prompts. The writing prompts and procedures for administering were based on procedures used in previous studies (Graham et al., 2005; Harris et al., 2006; Lane et al., 2008). All of the prompts administered were black and white line drawings that depicted children or animals engaged in an activity that was of interest to third-grade students (e.g., an alien landing in the front yard, a child pulling a dog in a cart). The order of the writing prompts was randomized for use in this study, and the prompts have been shown to be equivalent in terms of the writing generated and overall writing quality produced (Olinghouse, 2008). Third-grade teachers and students have also validated the writing prompts as interesting and appropriate for young children (Lane et al., 2008).

The third-grade students in the study were given as much time as needed to complete the writing responses and were individually tested using conditions that were identical to baseline conditions. Testing occurred in a quiet area of the school such as an empty classroom or library. Students were shown the writing prompt and were given three pieces of lined paper. They were told to look at the picture and write their best story about the picture. They were told to plan first on one of the pieces of paper. Finally they

were told to take as much time as needed and to start whenever they were ready. If students asked the researcher a question, they were told that no assistance could be provided. If the student indicated frustration in the writing task, he or she was encouraged to do their best. After the student completed the writing task, the researcher asked the students to read their story out loud. This provided an opportunity for the student to have an audience for their work and an opportunity for the researcher to "identify correct words that were misspelled or difficult to read because of illegibility" (Lane et al., 2008, p. 244).

During the treatment phase, writing prompts were administered after certain SRSD story–writing lessons (e.g., Lesson 1b, see Appendix B). The first writing prompt was delivered after students had been taught *lesson 1b*. (SRSD stages covered: Develop Background Knowledge and Discuss It), The second writing prompt was delivered after *Lesson 2* (SRSD stages covered: Discuss It, Model It and Memorize It). The third writing prompt was delivered after *Lesson 4* (SRSD stages covered: Discuss It, Model It, Memorize It, and Support It) . The fourth writing prompt was delivered after *Lesson 5* (SRSD stages covered: more of an emphasis on the Support It stage) . The fifth writing prompt was delivered after *Lesson 6* (SRSD stages covered: heading toward Independent Performance as less Support provided), The sixth writing prompt *varied* with some students revisiting earlier lessons (e.g., Cheyenne = Lesson 2; Nick = Lesson 6) and others completing lesson 7. The seventh writing prompt was only administered if students had not yet progressed through all eight SRSD story-writing lessons (had not yet reached the independent performance stage of instruction).

The stories were typed and spelling, punctuation, and capitalization errors were

corrected. Stories were scored for total length, number of story elements, and overall writing quality. First, the total amount written was recorded. The number of words included all written words in a story. This was scored for each story using a computerized word counting program.

Second, a story grammar scale was used to measure the number of story elements included in each participant's writing (see Lane et al., 2008). Each story was assessed for the following structural elements: main character(s); setting; time (or locale); what the main character(s) want to do (goals); action to satisfy the goal(s); consequence of actions; and character's reactions (Lane et al., 2008). Each element was given a *0* if the element was omitted, a 1 if the element was present, and a *2* if additional detail or elaboration was provided. These scores resulted in a total possible score of *14*. Similar versions of this story grammar scale have been validated and are highly correlated with other story-structure measures (Graves, Montague, & Wong, 1990), performance on standardized tests (Graham & Harris, 1989), and overall writing quality (Graham & Harris, 1989; MacArthur & Graham, 1987). The researcher and one graduate student unfamiliar with the design and study purpose scored all papers independently. The researcher first trained the graduate student using other third-grade writing samples (not samples from current study). Training continued until overall interrater reliability on these probes was above .80. The interrater reliability on scored papers was .91. The scores from the individual unfamiliar with the study were reported. These data are provided in the results section.

Finally, the stories were evaluated for overall writing quality. A holistic scoring method (Harris et al., 2006) that is based on an overall impression of writing quality gained from reading the written product was determined. Procedures for evaluating the

writing quality were drawn from similar studies (Graham et al., 2005; Harris et al., 2006; Lane et al., 2008). Two trained graduate students, blind to the design and purpose of the study, scored all papers for writing quality at the end of the study. Papers received a score of 0 to 8 with the high numbers representing papers that were of higher quality. Before scoring the papers for this study, the graduate students received training.

Training for scoring writing quality included providing graduate students with papers that represented several anchor points (scores of 1, 3, 5, and 7). These anchor points had been used in several studies (c.f., Graham et al., 2005; Harris et al., 2006; Lane et al., 2008) and were representative of third-grade writing. The graduate students were told to attentively (rather than laboriously) read the papers. They were shown how overall quality is determined based on several writing factors (no single factor should receive undue weight): ideation, organization, sentence structure, grammar, and appropriateness of word choice used in papers (see Lane et al., 2008). The graduate students were then given stories written by other third-grade students. The graduate students assigned each paper a score of 0 to 8 (using anchor points provided during training). The interrater reliability for each pair of papers was calculated. Training continued until overall interrater reliability reached .90. Interrater reliability for quality scores on the current study's writing prompts was .72. The researcher randomly selected which raters' set of scores to use for reporting overall writing quality.

Social validity. The nonprofessional adults and participating students participated in social validity interviews to assess how valuable and meaningful they felt the story-writing SRSD intervention was for them. For students, this was assessed after treatment had ended. The researcher completed these interviews to reduce the likelihood of students

giving inflated appraisals to their tutors who had spent much time working with them in a one-to-one situation. The student interview included the following questions:

1. Should the POW trick be taught to other students?
2. Should the W-W-W, What = 2, and How = 2 trick be taught to other students?
3. What did you like about how your teacher taught you POW and WWW…?
4. What did you dislike?
5. How would you change the way you were taught?
6. Have you used anything you learned at home or in another class?
7. Tell me what you did.

These results are provided in the results section (see Table 11).

For nonprofessional adults, social validity was assessed throughout the study. Throughout the treatment phase, the nonprofessional adults were asked to keep a journal. I provided guiding questions (adhered to front of journal) and asked the nonprofessional adults to address at least two to three in their weekly writing: (a) Was what you did with the student today effective? Why or why not?, (b) What do you feel is working well or not working well? (Be specific and tell why.), (c) Is there anything that is happening while you are teaching that is making you less effective? (Be specific.), (d) Rate the effectiveness of your instruction this week from "4" (very effective) to "1" (not effective). All of the journals were reviewed and relevant themes identified. These results are provided in the results section.

Procedures

The single subject design used in this study allowed me to determine whether the SRSD writing intervention had reliable and veridical effects, even though a large group of participants was not available for testing (Horner et al., 2005; Kennedy, 2005). A multiple-probe design across subjects was used to determine if SRSD story-writing instruction delivered by nonprofessional adults had positive effects on the writing behavior of six struggling third-grade writers. In the next section, I first detail procedures used to establish internal validity (consistency and study design) and then detail the procedures used during the pre-baseline and treatment (baseline, treatment, post-treatment, and maintenance) phases.

Internal validity. The multiple-probe design controlled for threats to the internal validity of the results in multiple ways. First, I implemented consistent procedures across all phases of the study. Second, I adhered to the rigors of the design, collecting sufficient data points and documenting stability before introducing the intervention in a time-lagged fashion across participants (Kennedy, 2005).

***Consistency*.** Consistency of procedures across phases maximized internal validity by minimizing that observed behavior changes occurred for reasons other than the introduction of the SRSD story-writing instruction (Kennedy, 2005; Lane et al. 2007). I implemented the intervention and administered writing probes across six students using consistent procedures. First, one student was paired with only one nonprofessional adult for the entire length of the study. Second, the student received instruction at the same time of the day (between 9:15 and 10:45 AM) and in the same location. The tutoring sessions occurred three to four times per week, for 30-40 min per session in the same quiet area outside of the student's classroom (an empty third-grade classroom).

Second, probe administration procedures remained consistent across phases. The same person (researcher), who was unfamiliar to all children in the study, administered functionally equivalent writing prompts and used identical writing protocols across all treatment phases. Others have found that functionally identical writing prompts were equivalent among third-grade writers (Olinghouse, 2004 as cited in Harris et al., 2006). Additionally, I maximized internal validity by adhering to the rigors of the multiple-probe design.

***Study design*.** During implementation (baseline, treatment, post-treatment, and maintenance), the design allowed me to establish that changes occurred because of the intervention and not due to history, maturation, instability, and intervening when change was likely to occur (Kennedy, 2005). For example, the treatment was not introduced if baseline data showed three ascending data points as another increased data point would seem likely.

I assessed whether the SRSD story-writing intervention had reliable and veridical effects on the outcome measure (the primary variable assessed story elements with scores ranging from 0 to 14) by administering ample writing prompts using consistent procedures. I was then able to assess the consistency of the effects across one demonstration and five replications. The multiple-probe design allowed me to control for effects that could have been due to history or maturation. First. I collected sufficient (at least 3) baseline data with the first pair of students, documenting stability (no more than 20 percent variance across mean and median phase lines) before introducing the intervention. These procedures were replicated with two more pairs of students. Although multiple studies have shown that 3nd-grade story-writing does not improve without

specific instruction (Adkins, 2005; Lane et al., 2008; Lienemann et al., 2006; Saddler et al., 2004), I continued to collect ample baseline data points for students in tier 2 (n = 5) and tier 3 (n = 8; see Figures 2, 3, and 4).

Second, I used the multiple-probe design to analyze the data, not moving to the intervention phase if there was instability with either students' data (within each tier) or if change was likely to occur (see how this is defined below). As recommended by Kennedy (2005), a minimum of three data points were required to establish a level of stability or trend. Stability was defined as data that varied by no more than 20 percent (based on the measure's total possible score) around the mean and median levels.

Stability was calculated using the following three steps. First, I multiplied .2 and the measure's highest score. For example, twenty percent variability around 14 (total possible on story element's measure) is 2.8. This number should be rounded to closest whole number (in this case 3). Second, student's mean and median data within a phase were calculated. For example, the female student in tier 1 had baseline scores of 3, 0, and 2 on the story element's measure. The mean and median of these numbers is 2 (1.7 is rounded to 2). Third, I determined whether these data varied by no more than twenty percent (in this case 3 points) around the mean and median (a score of 2). In this case, if there was a score of 6 or higher (mean and median score of 3 plus 3 points determined by multiplying .20 and 14) during baseline, data would not have been stable and more data points would have been needed. These data (3, 0, and 2) were stable and the intervention was introduced

I defined *change that was likely to occur* as 3 increasing data points. For example, if a student scored 3, 4, and 5 during baseline, they would have three increasing data

points. The intervention would not have been introduced until the data stabilized. This was done to ensure that the change that occurred during intervention could only be attributed to the introduction of the intervention and not improvement that had already had begun before the treatment was introduced.

In the current study, a multiple-probe design with two students (total of six students) in each tier was used to establish experimental control. Students were advanced to the treatment phase when their baseline data (number of essential story elements included) for both students within a tier were stable. Again, stability was defined as data that had no more than 20 percent variability around mean and median levels and no more than two accelerating data points. In addition to stable baseline data, students in tiers two and three were only advanced to the treatment phase when students in the proceeding tier demonstrated adequate improvement. Adequate improvement was defined as including at least 50% more story elements (compared to baseline condition) on writing assessments, over two adjacent data points. This 50% increase was only used to determine when students in the next tier should begin treatment and was not used to determine whether students within a tier would stop receiving SRSD story-writing lessons. Students continued in the SRSD treatment until they demonstrated the ability to write independently during tutoring sessions and included more story elements on writing assessments. The four instructional phases, plus a pre-baseline condition, are described next.

Treatment phases. There were five main treatment phases: pre-baseline, basline, treatment, post-treatment, and maintenance.

Pre-baseline. The nonprofessional adults met with students one time prior to the study, at a similar time, and in the same location, to build rapport prior to starting the study. The nonprofessional adults were not allowed to work with familiar students. Two of the nonprofessional adults were unfamiliar with the school and its occupants. The third nonprofessional adult did have a student in the third-grade, but did not have a previous relationship with either of her tutees. This tutor stated that she had seen the students in the cafeteria, but had only had brief conversations with one of them. In pre-baseline, as in the remaining four conditions, one nonprofessional adult worked individually with the same student in the same location. The work area was an unused classroom in the same hallway as the other third-grade classrooms. This room was equipped with tables, desks, chairs, and whiteboards.

During the pre-baseline phase, a writing survey was also administered to the three participating third-grade teachers. This survey was used to describe the writing instruction that students' teachers currently use. These data are important in understanding the results of the study, were provided earlier, and will be used in the discussion section

Baseline. During the baseline phase, the third-grade students' pre-intervention response rates were established for writing stories (see Adkins, 2005). The primary dependent variable used to determine the rate was the number of essential story elements (scores ranged from 0 to 14). Pre-intervention data for this measure were collected until stability or decreasing trends were established for each student (see earlier description). A minimum of three data points was required to establish data stability and trend (Kennedy, 2005). The researcher administered all writing prompts during the study in a quiet area

outside the classroom. The same procedures were used to administer the story writing prompts during all phases of the intervention (see earlier description of writing prompt administration).

Treatment. The nonprofessional adults provided individual instruction to two students, two to four days per week. Instruction was typically delivered three days per week with both students in the tier receiving instruction on the same day. However some changes occurred due to school cancellations (snow days) and student absences. The nonprofessional adults met individually with the researcher for a brief meeting before delivering that day's writing lesson. Although each nonprofessional adult was equipped with his/her training manual, the researcher duplicated all necessary lesson documents (e.g., copy of lesson, picture prompt, lined paper, previously written story) and organized each tutee's folder. During this pre-treatment meeting, the researcher also supplied the nonprofessional adult with a tape player and a blank audiotape. The researcher asked the nonprofessional adult if he/she had any questions and if not observing, would excuse herself from the session before the student arrived.

The researcher administered all writing prompts to students when the nonprofessional adult tutors were not present. The writing prompts were delivered after students had completed lesson 1 (2 to 3 days of instruction), lesson 2 (4 to 6 days of instruction), lesson 4 (5 to 8 days of instruction), lesson 5 (7 to 10 days of instruction), lesson 6 (9 to 13 days of instruction), and lesson 7 (11 to 17 days of instruction; see Appendix B for individual lesson descriptions). The researcher delivered these writing prompts using procedures identical to baseline condition. Treatment data, in part, were used to determine when to make phase changes for students in subsequent tiers. As stated

earlier, students began treatment after demonstrating baseline stability and observing adequate improvement for students in the previous tier. Adequate improvement was defined as including at least 50% more story elements (compared to baseline condition) on writing assessments, over two adjacent data points.

Treatment data, together with students' progress with the nonprofessional adult, were also used to determine if students had reached criterion and could graduate from the Writing Academy. First, the conclusion of the writing treatment was based on student's ability to plan (WWW) and write a good story with all seven story parts independently when working with the nonprofessional adult (see lesson 7, Appendix B). Second, the conclusion of the writing treatment was also based on the number of story elements included in stories written during the testing sessions. In order for instruction to end, the student had to have included more story elements on at least two adjacent data points than were included in their highest baseline probe. For example, if the student's highest baseline score on the story elements scale was 3, the student would had to have a score of 4 or higher on at least two adjacent treatment probes (and successfully demonstrate the ability to plan and write independently) to conclude the writing treatment.

The number of sessions required to meet criterion for each participant varied. Harris et al. (2003) stated that when instruction is delivered in 30-minute sessions, elementary students typically reach the independent performance stage in 8 to 12 lessons (see Harris et al., 2003) The students in this study reached criteria in 11 (Mary), 12 (Nick, Alexis, and Lucas), 16 (Jacob), and 17 (Cecilia) lessons. The total time in instruction for each student was as follows: Mary (5.5 hours), Nick, Alexis, and Lucas (6.0 hours), Jacob (8.0 hours), and Cecilia (8.5 hours). Treatment occurred over 4 to 6 weeks.

***Post-treatment*.** Post-treatment data collection procedures exactly replicated those used during baseline and treatment phases. The same researcher administered a minimum of three post-treatment story-writing probes within four weeks after treatment had ended. Again, the primary outcome measure used was the number of essential story elements (scores ranged from 0 to 14). Post-treatment probes continued (minimum of three) until stability had been reached. Again, stability was defined as no more than a 20 percent variance around the mean and median levels.

***Maintenance*.** Maintenance procedures again replicated previous conditions. The same researcher collected maintenance data at a minimum of six weeks after instruction had ended. The school year ended before maintenance data could be collected for students in the final group. These data are used to determine if participations maintained instructional effects.

Data Analysis

Treatment results. The purpose of the current study was to assess the effectiveness of a supplemental writing treatment delivered by nonprofessional adult tutors on improving the writing outcomes of third-grade struggling writers. Experimental control was assessed using a multiple-probe across students design. Visual analysis was the primary method for analyzing the single-subject graphed data. Visual analysis of the story element data occurred frequently throughout the study and was used to make key decisions about phase changes. The researcher used these procedures to examine the data and assess changes related to level, trend, and stability (Kennedy, 2005).

Experimental control (did the intervention and the intervention alone cause observed changes to occur) was established when data demonstrated a predicted pattern of change across participants, but remained stable in untreated legs/tiers. Experimental control was only achieved when predicted co-variation between introduction of treatment and changes in the dependent measures were demonstrated across three students' data points.

Level. Level was assessed by calculating the change in mean and median phase lines and by calculating the percentage of nonoverlapping data (Kennedy, 2005). First, I calculated the mean and median value of the data for each dependent measure (story elements, quality, number of words written). I calculated the mean phase line by first finding the average score within a phase and then drawing a line at that point. The median phase line was only calculated when two out of three data points were the same and it varied from the mean. These procedures were repeated for each participant's baseline, treatment, and post-treatment phases as well as for student's maintenance phases when such data were available (see Table 9). Post-treatment and maintenance data were not available for Alexis (tier 3) due to an extended school absence. Maintenance data were not available for Lucas (tier 3) as the end of the school year occurred before these data could be collected.

Second, I also calculated the percentages of nonoverlapping data (PND) by drawing a line through the highest data points for each student's baseline phase and then determining the percentages of points in the adjacent conditions that exceeded those levels (Kennedy, 2005). I did not repeat these steps to compare students' treatment and post-treatment data and post-treatment and maintenance data as it was not assumed that

students would continue to accelerate during the post-treatment and maintenance phases. I calculated PNDs for all graphed outcome measures (story elements, holistic writing quality, and number of written words, see Figures 2, 3, and 4). These data are reported in the results section.

Slope. I also assessed the slope or the direction the data were moving over time (Kennedy, 2005). The slopes, or trends, of the data were calculated using a computerized program (Microsoft excel). To calculate the slope, I first entered the series of data for each treatment phase (baseline phase). I then selected the data and used a slope command to calculate the slope using the number of probes (1, 2, 3…) and the corresponding data points (3, 5, 8…). I calculated slopes for available data. These 21 slopes are reported in the results section (see Table 9).

Stability. Finally, I assessed the predictability of the data series (Kennedy, 2005). I assessed how much data moved within each phase versus between each phase. Specifically, I calculated six phase lines for each condition: baseline, treatment, post-treatment, and maintenance. The percentage of data that lay outside the 20 percent envelope (around the mean phase line) was reported. This information was used to assess within and between phase changes and to determine whether these changes were consistent across replications. These data are also reported in the results section (see Table 9).

CHAPTER IV

RESULTS

Three trained nonprofessional adult school volunteers taught six struggling third-grade writers a story-writing strategy (i.e., WWW, What = 2, How = 2) using the six SRSD instructional stages (*Develop Background Knowledge; Discuss It; Model It; Memorize It; Support It*; and *Independent Performance*). Four research questions guided this investigation. The first question pertained to procedural fidelity. Did the nonprofessional adults demonstrate mastery of the SRSD material during the pre-baseline training sessions? I observed tutors conducting each lesson during the training sessions and recorded the percentage of lesson components completed during each lesson. Additionally, did the nonprofessional adults complete all sessions with a high degree of fidelity? I assessed fidelity during implementation by determining the (a) percentage of lesson components completed during each student's lesson and (b) the quality of instruction. I predicted that tutors would cover at least 90% of each lesson's components for each child taught and would provide quality instruction.

The second and third questions related to the effect of the story-writing treatment on improving writing outcomes for struggling third-grade writers. First, did the students include more essential story elements, write stories of improved quality, and compose longer stories (*M* and *Mdn*) during the treatment phase? Second, did students maintain improved writing performance during post-treatment (within a four week period after treatment ended) and maintenance phases (six weeks after treatment ended)? I anticipated

that students' post-treatment data would be higher than their baseline data and that their maintenance data would decrease by no more than 20% from post-treatment levels (mean and median).

The fourth question pertained to social validity. Did students and nonprofessional adults report high social validity for the SRSD story-writing instruction? This was assessed through student and nonprofessional adult interviews as well as by analyzing nonprofessional adult journal entries.

Procedural Fidelity

Lesson components. Three types of procedural fidelity were collected and assessed: percentage of lesson components tutors successfully applied during *training sessions*, percentage of lesson components covered during *tutoring sessions*, and teaching quality observed during tutoring sessions. First, during training sessions, I observed each tutor conduct each lesson in order to assess each tutor's ability to successfully demonstrate the SRSD lessons. During these training observations, I used a form (identical to those used during treatment) to record the percentage of lesson components successfully completed. Each training session was completed on an individual basis one week prior to the onset of tutoring sessions. All tutors could successfully implement 100% of all lesson components during these training session observations.

Second, all tutoring sessions with students were audiotaped and evaluated for the number of lesson components successfully completed during each lesson (see Table 8). The number of components for each lesson ranged from 6 to 13 and included items such as practicing WWW, What = 2, How =2, modeling planning, and lesson wrap-up. An

example of a procedural fidelity form can be found in Appendix C. The procedural fidelity forms assessed whether the SRSD lessons were conducted as planned. A trained graduate student listened to each audiotape and completed the corresponding procedural fidelity form for each lesson (see Appendix B for form examples). I randomly reviewed between 25% (Alexis = 3/12 lessons) and 45% (Michelle = 5/11 lessons) of each student's audiotapes to document interrater reliability with the graduate student.

Table 8. Procedural Fidelity: Percentage of Lesson Components Covered and Teaching Quality Scores

Tutor	Percentage of Components Covered for Lessons 1a-7			Teaching Quality Scores	
Student	*M*	*N*	range	Lessons Assessed	Average Quality Score (range 0 to 8)
James					
Cecelia	97%	17	75% to 100%	5/17 lessons (29.4%)	7.2 (5.0 to 8.0)
Justin	95%	16	70% to 100%	5/16 lessons (31.3%)	6.8 (5.0 to 8.0)
Jessica					
Michelle	98%	11	80% to 100%	5/11 lessons (45.5%)	7.8 (7.0 to 8.0)
Nick	99%	12	92% to 100%	4/12 lessons (33.3%)	7.5 (7.0 to 8.0)
Phyllis					
Alexis	92%	12	75% to 100%	4/12 lessons (33.3%)	7.5 (7.0 to 8.0)
Lucas	93%	12	80% to 100%	4/12 lessons (33.3%)	7.3 (7.0 to 8.0)

Notes. M = mean. *N* = total number of lessons. Quality scores ranged from 0 to 8 and were measured across 8 measures: preparedness, attitude in general, attitude about SRSD, rapport, behavior management, progress monitoring, lesson pacing, and lesson ending. Beh. = behavior.

Interrater reliability was calculated by dividing the components observed by the components possible and multiplying the quotient by 100. Interrater reliability was .98 (range .92 to 1.0). As anticipated, each tutor covered at least 92% of lesson components with scores ranging from 70% to 100% (see Table 8).

The first tutor, James, omitted the following lesson components at some point when teaching: lesson wrap up, adequately discussing self-statements, and modeling the writing process. The second tutor, Jessica's, primary omission was the lesson wrap up. She stated the reason for an abrupt ending was student sickness. The third tutor, Phyllis, most frequently omitted (although this was only occasionally) the lesson wrap up. Phyllis stated that she often covered the component, but the activity was not captured on the recorder. Phyllis also omitted the introductions on several occasions. Phyllis stated that she forgot to capture the component on the audio recorder or simply forgot it altogether due to her excitement to teach the students the writing strategies (WWW, What = 2, How = 2). All of the lesson components omitted throughout the Writing Academy are provided in Table 9. Omission *patterns* were not consistently observed for any tutor and if a lesson component was skipped (and important to overall effectiveness of lesson), it was added in the next lesson.

Teaching quality. I observed approximately 33 percent of each student's lessons and evaluated overall teaching quality using a 0 to 8 point scale (see Table 8). An evaluation of these data is important because the participating tutors were not trained teachers and did not have any writing tutoring experience. The first tutor, James is a retired army chaplain. The second tutor, Jessica, has a nursing degree. The third tutor,

Table 9. Lesson Components Omitted Throughout the Writing Academy

Lesson Component	Description	Relevance to SRSD instructional stage	Frequencies of Omissions (Total Sessions)	Re-taught
Make introductions	Build rapport with tutees and foreshadow "writing tricks".	*Develop Background Knowledge*	*James = 1* *Phyllis* = 2	No Yes
Discuss good stories	Discuss student's favorite stories and reasons students like them.	*Develop Background Knowledge* & *Discuss It*	*James = 2*	Yes
Introduce and Discuss POW	Emphasize POW is a trick good writers often use for many things they write. Discuss importance.	*Develop Background Knowledge* & *Discuss It*	*Phyllis = 4*	Yes
Discuss MDW	Discuss importance of using MDWs in writing to make stories more interesting to read.	*Develop Background Knowledge* & *Discuss It*	*James = 1*	Yes
Discuss Transfer	Goal for POW and WWW strategies: use them in other writing tasks.	*Develop Background Knowledge* & *Discuss It*	*James = 3*	Yes
Discuss Self-statements	Discuss importance of positive self-talk. Use statements that are similar to those used by student. Discuss letting mind be free and clear.	*Develop Background Knowledge* & *Discuss It; Self-regulation*	*James = 1*	Yes
Adds to self-statement sheet	Help student identify new statements such as "I know tricks for writing good stories."	*Discuss It and Model It; Self-regulation*	*Phyllis = 1*	Yes
Model Writing	Model planning and writing a story that has all 7 parts, is fun to read, and makes sense.	*Discuss It and Model It*	*James = 2* *Phyllis = 1*	Yes
Graph Data	Ask students if their writing has all 7 parts and MDWs.	*Discuss It and Model It; Self-regulation*	*James = 1* *Phyllis* = 1	Yes
Lesson Wrap Up	Remind student about "test", organize materials, discuss transfer	*Discuss It and Memorize It*	*James = 2* *Jessica = 1* *Phyllis = 4*	No

Phyllis, has a marketing degree. The teaching quality data were used to determine the quality of instruction provided to students.

Tutors received a score of 1 for displaying each of the following behaviors: (a) preparedness (Was the tutor well prepared for SRSD instruction when I arrived?), (b) positive attitude in general (Was the tutor positive and enthusiastic with the student?), (c) positive attitude about SRSD (Did the tutor appear positive and enthusiastic about POW and WWW, What = 2, How = 2?) (d) building / establishing rapport (Did the tutor seem to have rapport, or was working on building rapport, with the student?), (e) behavior management (Was the tutor able to redirect off-task behavior so that they did not interfere with instruction?), (f) progress monitoring (Did the tutor carefully monitor and support the student and actively involve them in instruction?), (g) lesson pacing (Was the lesson appropriately paced?), and (h) lesson ending (Did the tutor end the lesson well and set goals for the next lesson?). Data related to quality scores are reported in Table 8.

The nonprofessional adults (James, Jessica, and Phyllis) provided high quality instruction. On average, the tutors met at least 7 out of 8 quality indicators across all teaching sessions (see Table 8). On average, the first tutor, James, exhibited 6.8 out of 8.0 quality behaviors during Justin's lessons and 7.2 out of 8.0 quality behaviors during Cecelia's lessons. The quality indictors that were not met generally differed by student. James did not exhibit the following in one or more of Cecilia's lessons: preparedness, progress monitoring, lesson pacing, and lesson ending. James did not exhibit the following in one or more of Jacob's lessons: preparedness, positive attitude (in general), behavior management, and lesson ending. James was the first tutor to deliver instruction in this study. As quality scores indicated, lesson preparedness was an initial problem. I

used this information to alter tutor preparation and lesson procedures. Specifically, I assumed responsibility for preparing students' daily lesson folders. I copied the appropriate materials (e.g., teaching script, prompts, stories) and discussed the material with each tutor immediately before each tutoring session began. Identical procedures (i.e., preparing students' folders) were repeated with Jessica and Phyllis.

I also provided specific feedback for all tutors. The specific feedback to James related to progress monitoring, lesson pacing, behavior management, and lesson ending. James indicated that modeling was most challenging for him (e.g., demonstrating how one picks an idea, organizes notes, and writes and says more). James stated that he felt he was lecturing too much and consequently losing student attention. I suggested that James check for student understanding more frequently to ensure that students were attending and more accurately gauge lesson effectiveness. Prior to this suggestion, James would talk for up to 5 minutes without interacting with his students. Upon receiving feedback, James was more successful in actively engaging each student. James solved many of the problems related to *lesson pacing* and *lesson ending* by setting time management goals (e.g., the first part of the lesson should not take longer than 5 minutes) or using a watch to monitor the time. Suggestions for effective behavior management strategies were discussed after each tutoring session. I suggested that James set clear goals for each lesson and remain positive and enthusiastic with each student. James also dedicated much of his journal writing to the topic of behavior management and his struggles with certain issues. These data are provided in the social validity section.

Jessica and Phyllis, the second and third tutor, began their tutoring sessions after James and benefited from James' tutoring experience. First, their training sessions

included specific guidance for problematic issues that had come to light during James' tutoring sessions (e.g., time management, behavior management). Second, James shared his experience directly with each tutor. Each tutor stated that they found these discussions helpful.

Jessica and Phyllis met almost all of the quality indicators when teaching. On average, Jessica met 7.8 (Michelle) and 7.5 (Nick) of the quality indicators. Jessica occasionally omitted the following behaviors: positive attitude about SRSD instruction and lesson ending. These were only omitted on one occasion and feedback was provided. Jessica successfully exhibited these teaching qualities on every observed session thereafter. On average, Phyllis met 7.5 (Alexis) and 7.3 (Lucas) of the quality indicators. On one occasion, she did not display the following behaviors: preparedness, lesson ending, and behavior management. Feedback was provided and future lessons improved.

The high levels of procedural fidelity (lesson components covered and teaching quality) were most likely influenced by the rich amount of support provided to each tutor during the treatment phase (see Table 10). There were seven main types of support provided: providing transportation to and from the school (James and Jessica), preparing student folders, delivering and removing tape recorders and audio tapes, providing general feedback after each tutoring session, observing regularly during treatment delivery, providing specific feedback related to lesson component omissions and teaching quality indicators, and providing a journal for tutors to record lesson effectiveness and comments. The importance of providing these supports and the impact on procedural fidelity levels are further addressed in the discussion section.

Table 10. Support Provided to Nonprofessional Adults During the Treatment Phase

Type of Support	Definition	Frequency of Support	Goal
Transportation	I provided round-trip transportation for James and Jessica. Tutors were either picked up at their homes or met me at my house. Each way took approximately 25 minutes. There were several occasions in which Jessica and James share the same ride.	*James* 16 occasions for a total of 13.3 hours *Jessica* 11 occasions for a total of 9.2 hours	1. Make volunteering more manageable and reduce associated costs (i.e., gas and total mileage) 2. Increase opportunities for feedback 3. Increase tutor attendance 4. Improve procedural fidelity
Folder Preparation	I prepared each student's folder to ensure that only the correct materials were present. I gave to tutor before each tutoring session began. This took approximately 10 minutes per folder.	*James* 13/16 & 14/17 lessons for a total of 4.5 hours *Jessica* All lessons for a total of 3.8 hours *Phyllis* All lessons for a total of 4.0 hours	1. Ease workload for tutor 2. Improve procedural fidelity
Tape Recorder and Audio Tape	I was in charge of delivering and removing the tape recorder and tape each day.	Every lesson. By the end of the treatment, I had collected 17, 12, and 12 tapes for James, Jessica, and Phyllis (respectively).	1. Ease workload for tutor 2. Improve procedural fidelity
General Feedback	I was in the school during each tutoring session. If not observing in the room, I would wait in an unused room (e.g., teacher's lounge) until sessions were over. After sessions were completed, tutors would relay how each session went.	Every treatment session.	1. Ensure tutors questions were answered 2. Improve procedural fidelity

Table 10 continued. Support Provided to Nonprofessional Adults During the Treatment Phase

Type of Support	Definition	Frequency of Support	Goal
Observations	I observed nonprofessional adults delivering instruction. Observations lasted for entire lesson (approximately 30 minutes). I recorded lesson components covered and evaluated teaching quality.	*James* 10 observations - total of 5.0 hours *Jessica* 9 observations - total of 4.5 hours *Phyllis* 8 observations – total of 4.0 hours	1. Monitor procedural fidelity 2. Improve procedural fidelity
Specific Feedback	After observing lesson, I provided specific feedback related to lesson component omissions and teaching quality. Each feedback session lasted approximately 15 to 20 minutes.	*James* 1.3 to 1.7 hours *Jessica* 1.3 to 1.7 hours *Phyllis* 1.0 to 1.3 hours	1. Improve procedural fidelity 2. Provide opportunities for problem solving and tutor comments.
Tutor Journaling	Tutors rated lesson effectiveness (1 to 4 scale) and provided written comments about each tutoring session. Journals turned in at end of treatment phase.	11 (Phyllis), 12 (Jessica), and 17 (James) journal entries	1. Assess social validity 2. Improve procedural fidelity

In summary, the tutors taught each SRSD story-writing lesson with a high degree of fidelity (as indicated by the high percentage of lesson components covered by each tutor for each student) and provided high quality of instruction (as indicated by high average quality teaching scores). Problems identified during classroom observations, tape reviews, and discussions were used to improve training and instructional procedures (see

Table 10). Next, the effects of the nonprofessional adult-led SRSD writing instruction on multiple writing outcomes are provided. Results are provided for the number of essential story elements, holistic writing quality, and number of words written. The effects of the nonprofessional adult-led SRSD writing instruction on multiple writing outcomes are presented. Results are provided for number of essential story elements included, holistic writing quality, and total words written.

Writing Outcomes

Data for the three writing outcomes (essential story elements, holistic writing quality, and number of written words) are presented in Table 11. The writing outcome measures are also depicted on three multiple baseline graphs (see Figures 2, 3, and 4). James' tutees, Cecilia and Justin, are always represented on the first tier of each graph. Jessica's tutees, Michelle and Nick, are always represented on the second tier in each graph. *Baseline, treatment, post-treatment, and maintenance* data were collected for these four students. Lastly, Phyllis's tutees, Alexis and Lucas, are always represented in the third tier of each graph. *Baseline and treatment* data were collected for Alexis and Lucas. Post-treatment data were also collected for Lucas. Alexis's post-treatment data were not collected due to an extended school absence. Maintenance data were not collected for Lucas as the end of the school year occurred before these data could be collected.

Each student is consistently represented with the same figure (e.g., shaded square) on each multiple baseline graph (see Figures 2, 3, and 4). Justin and Lucas's data were represented with blackened squares. Cecilia and Nick's data were represented with open diamond shapes. Michelle and Alexis's data were represented with either a blackened

Table 11. Dependent Variables Data: Level (M and PND), Slope, and Variability

Student	**Number of Essential Story Elements** (0 to 14)				**Holistic Quality** (0 to 8)				**Number of Written Words**			
	Base.	Tx	Post-Tx	Maint.	Base.	Tx	Post-Tx	Maint.	Base.	Tx	Post-Tx	Maint.
Cecelia												
M (*SD*)	2.0 (1.0)	3.6 (1.8)	7.3 (1.2)	8.0 (0.0)	1.7 (1.2)	2.3 (0.8)	2.0 (1.0)	2.7 (0.6)	22.0 (6.0)	31.3 (5.2)	32.3 (2.1)	30.0 (3.5)
PND		42.9%				0.0%				71.4%		
Slope	-0.5	0.7	-1.0	0.0	-1.0	0.1	-0.5	0.5	-5.5	1.6	0.5	0
Variability	0.0%	14.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	14.3%	0.0%	0.0%
Justin												
M (*SD*)	1.7 (2.1)	4.4 (3.3)	5.3 (3.1)	3.5 (0.6)	0.3 (0.6)	1.6 (1.1)	1.7 (1.2)	1.7 (0.6)	16.7 (12.2)	23.9 (10.1)	24.3 (20.8)	29.0 (1.7)
PND		57.1%				42.9%				28.6%		
Slope	0.5	1.3	1.0	0.0	0.0	0.4	1.0	0.0	-12.0	3.2	16.0	-1.5
Variability	0.0%	42.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	66.7%	57.1%	66.7%	0.0%

Table 11 continued. Dependent Variables Data: Level, Slope, and Stability

Student	**Number of Essential Story Elements** (0 to 14)				**Holistic Quality** (0 to 8)				**Number of Written Words**			
	Base.	Tx	Post-Tx	Maint.	Base.	Tx	Post-Tx	Maint.	Base.	Tx.	Post-Tx.	Maint.
Michelle												
M (*SD*)	2.2 (1.1)	5.4 (2.4)	6.7 (1.5)	8.0 (1.2)	1.8 (0.8)	2.6 (1.3)	4.0 (1.0)	4.0 (1.0)	44.0 (26.5)	41.3 (18.4)	64.3 (6.1)	80.3 (23.7)
PND		57.1%				14.3%				14.3%		
Slope	0.3	1.1	1.0	1.0	-0.4	0.5	-0.5	0.5	-14.0	6.1	2.0	20.0
Variability	0.0%	14.3%	0.0%	0.0%	0.0%	14.3%	0.0%	0.0%	60.0%	14.3%	0.0%	33.3%
Nick												
M (*SD*)	2.8 (1.3)	4.6 (3.0)	6.3 (0.6)	8.5 (1.0)	1.6 (1.3)	2.4 (1.0)	3.3 (0.6)	4.0 (1.0)	48.4 (18.0)	38.9 (13.8)	50.3 (8.1)	82.3 (58.6)
PND		42.9%				0.0%				0.0%		
Slope	0.2	1.2	0.6	-1.0	0.0	0.3	0.5	-0.5	1.7	3.9	7.0	-0.5
Variability	0.0%	28.6%	0.0%	0.0%	20.0%	0.0%	0.0%	0.0	0.0%	0.0%	0.0%	100.0%

Table 11 continued. Dependent Variables Data: Level, Slope, and Stability

Student	**Number of Essential Story Elements** (0 to 14)				**Holistic Quality** (0 to 8)				**Number of Written Words**			
	Base.	Tx	Post-Tx	Maint.	Base.	Tx	Post-Tx	Maint.	Base.	Tx.	Post-Tx.	Maint.
Alexis												
M (SD)	5.0 (1.6)	6.7 (1.6)	ND	ND	2.9 (1.2)	3.3 (1.3)	ND	ND	44.5 (15.8)	62.8 (19.2)	ND	ND
PND		16.7%				0.0%				33.3%		
Slope	0.1	0.6	ND	ND	0.3	0.1	ND	ND	2.8	5.8	ND	ND
Variability	12.5%	0.0%	ND	ND	25.0%	33.3%	ND	ND	12.5%	50.0%	ND	ND
Lucas												
M (*SD*)	2.1 (1.2)	4.2 (2.3)	7.0 (1.0)	ND	1.1 (0.6)	1.8 (1.5)	2.7 (1.2)	ND	30.4 (5.6)	26.7 (8.2)	32.7 (7.2)	ND
PND		50.0%				33.3%				16.7%		
Slope	-0.1	1.0	0.5	ND	0.0	0.2	-1.0	ND	-0.6	2.1	-6.5	ND
Variability	0.0%	33.3%	0.0%	ND	0.0%	33.3%	0.0%	ND	0.0%	33.3%	33.3%	ND

Notes. Base. = baseline phase; Tx = treatment phase; Post-Tx = post-treatment phase (within four weeks after treatment had ended); Maint. = maintenance phase (six weeks after treatment had ended). SD = standard deviation. PND = percentage of nonoverlapping data calculated by drawing a line through the highest baseline data point and determining the number of treatment points that exceeded those levels. ND = no data. Slope = calculated by using a computerized program and used to assess the direction data were moving over time. Variability was calculated by first multiplying .2 by student's highest score (number of words written) or by highest possible score (story elements = 14; holistic quality = 8). This number was then added and subtracted to each mean. The percentage of data that lay outside those numbers was used to describe variability.

circle (Michelle) or an open circle (Alexis). A booster session was provided to both students in the first tier and is depicted on each graph. I provided a short, 10-min, individual booster session for Justin after observing his decreased writing performance on his second post-treatment probe (see Figure 2). Cecilia had been absent for two weeks, so she was also given a booster session. Booster sessions included a discussion related to good stories (e.g., What makes a story good?) and a review of the general and specific story-writing tricks (i.e., POW and WWW, What = 2, How = 2). In the sections that follow, data are presented according to the four phases of the study: baseline, treatment, post-treatment, and maintenance. Data changes are described in relation to changes made to level, slope, and stability (see Table 11) and are always in relation to the adjacent condition (e.g., baseline to treatment, treatment to post-treatment, post-treatment to maintenance), unless otherwise specified. An evaluation of experimental control is provided at the end of each section.

Experimental control was established for the treatment phase when at least three students' performance on the dependent measures was stable during baseline condition and consistently changed in the desired direction after the treatment was introduced (Kennedy, 2005). Additionally, experimental control was established for post-treatment and maintenance levels if (a) experimental control had first been established during the treatment phase, (b) data were stable, and (c) data were consistent or improved when compared to the previous phase (level, slope, or variability). The data used to evaluate experimental control are provided in Table 12. It is important to remember that number of *essential story elements* (see Figure 2) was used to establish stability and make decisions about phase changes.

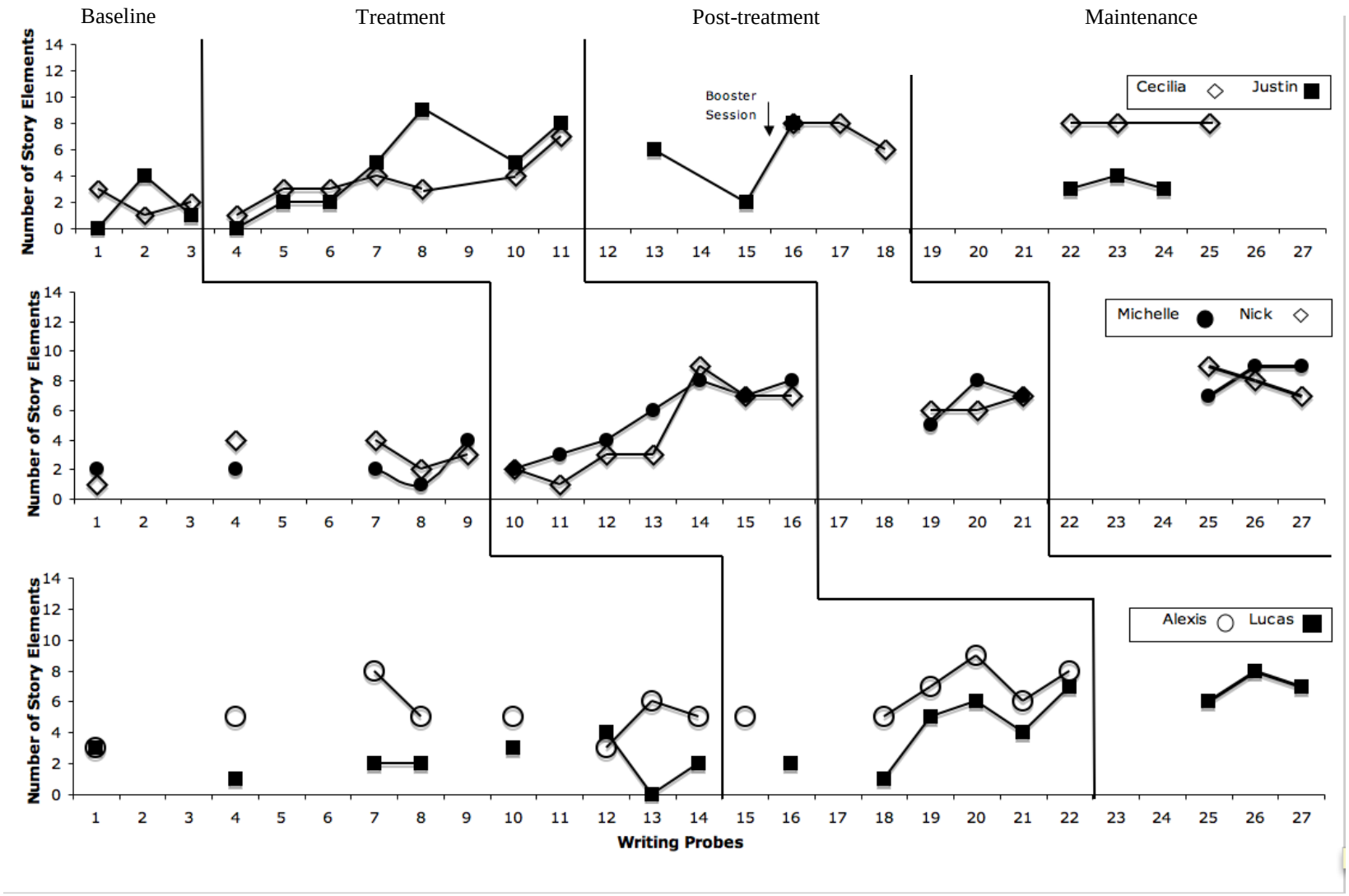
The two basic questions concerning essential story elements were: Did the students include more essential story elements (mean and median) in their stories during the treatment phase? Did the students maintain improved writing performance during post-treatment (within four weeks after treatment ended) and maintenance phases (six weeks after treatment ended)? I anticipated that students' post-treatment data would be higher than their baseline data and that their maintenance data would decrease by no more than 20% from post-treatment levels (mean and median). Results for story writing elements are provided first.

Essential story elements. The data related to essential story elements (level, slope, variability) are provided in Table 11. Individual data are also graphed in Figure 2. Story elements scores ranged from 0 to 14 with students receiving up to two points for the following: character, time, location, stated goal, action to achieve those goals, feelings, and ending.

In considering essential story elements (and all other writing measures), I first describe if data were low and stable during the baseline phase (before the treatment was introduced). Next, I describe data changes (mean/median scores, PND, slopes, and variability) that occurred during treatment, post-treatment and maintenance phases. I then provide data related to experimental control. Finally, I provide a summary of the findings as they relate to research questions two and three.

The number of essential story elements included in students' stories during baseline was low and stable for all six participants (see Table 11 and Figure 2). The average baseline story elements' score across all participants was 2.6 with scores ranging from 1.7 (Justin) to 5.0 (Alexis). Four of the students (Cecilia, Justin, Mary, and Lucas)

Figure 2. *Effect of SRSD Story-Writing Instruction on Story Elements*



had mean and median baseline scores that did not surpass 2.2. Nick had slightly higher mean/median baseline scores of 2.8 and 3.0, respectively. Alexis's mean/median baseline scores were 5.0 with a range of 3.0 to 8.0. The variability of baseline data across five of the six participants (exception Alexis) was 0.0%. Alexis had one data point that fell more than 20% away from her baseline mean (variability = 12.5%). However, her last three data points were stable. The baseline slopes were slightly positive for most students (n =4), ranging from 0.1 (Alexis) to 0.5 (Justin). The remaining two students had slightly negative baseline slopes of -0.1 (Lucas) and -0.5 (Cecilia).

In every case, story elements scores improved during the treatment condition (level and slope). The overall mean story element score across all participants was 4.8 (median = 4.7) with mean scores ranging from 3.6 (Cecilia; median = 3.0) to 6.7 (Alexis, median = 6.5). A comparison of treatment to baseline scores (means and medians) revealed that, on average, students included 2.18 (median = 2.17) more story elements in their stories during the treatment phase. The increased number of story elements ranged from 1.8 for Nick (median = 0) to 3.2 for Michelle (median = 4.0). The percentages of nonoverlapping data (PND) between baseline and treatment phases ranged from 16.7% to 57.1% and were consistent across five out of six students (see Table 11). The PND scores were as follows: 42.9% (Cecilia, 3/7), 57.1% (Justin, 4/7), 42.9% (Nick, 3/7), 16.7% (Alexis, 1/6), and 50% (Lucas, 3/6).

The PND scores met my expectation as data were gathered during instruction with the first data points occurring as early as lesson 1b (see Appendix B). *Immediate* improvement and higher PND scores were not expected. I anticipated that each student would start showing improvement after having the opportunity in treatment to plan and

draft using pre-made graphic organizers *without* pictures (see Lesson 5, Appendix B, treatment prompt number 4). Four students' data met these expectations: Cecilia, Justin, Michelle, and Alexis (see Figure 2). Lucas started showing improvement after he reviewed his previous writing (see Lesson 4, treatment prompt number 3). Nick started showing improvement after writing without the use of any pre-made graphic organizer (see Lesson 6, treatment prompt number 5).

The data also revealed that each student's slope increased positively during the treatment phase (see Table 11 and Figure 2). The slopes ranged from nearly 1.0 (Alexis = 0.6; Cecilia = 0.7) to almost 1.5 (Justin = 1.3). This upward trend resulted in most students having a greater percentage of data fall further than 20% from their mean score. Treatment data were more variable for 83.3% (6/7) of the students. For these students, variability ranged from 14.3% (Cecilia and Michelle) to 42.9% (Justin). Alexis was the only student whose data became less variable during treatment. Her variability scores went from 12.5% to 0.0% (see Table 11).

The students continued to improve during the post-treatment phase (Alexis was not evaluated due to school absences). The overall mean and median story elements score during post-treatment was 6.5 with scores ranging from 5.3 (Justin; median = 6.0) to 7.0 (Lucas; median = 7.0). A comparison of treatment to post-treatment mean scores revealed improvements of 0.9 (Justin) to 3.7 (Cecilia). A similar comparison using median scores revealed improvements of 0.5 (Lucas) to 5.0 (Cecilia). Most students' PND scores between baseline and post-treatment phases were 100% (Cecilia, Michelle, Nick, and Lucas). Justin had one post-treatment data point that equaled his highest baseline score resulting in a PND of 66.7% (average PND across five students = 93.3%). The

comparison between baseline and treatment phase was made in order to make a comparison with previous examinations of SRSD's effectiveness. The current study's average PND score of 93.3% at post-treatment was similar to the >90% PND scores reported in 27 other SRSD single subject studies (see Rogers and Graham, 2008).

When considering treatment slopes, only one post-treatment slope decreased (-1.0, Cecilia). Finally, a comparison of treatment to post-treatment variability revealed an increase in data stability when all participants were considered together. Post-treatment variability was 0.0% across all participants.

Maintenance data revealed that most students' mean and median scores were very consistent with post-treatment scores (did not decline by more than 20%). All students' maintenance scores (mean and median) surpassed their baseline scores. The average maintenance score across the four students was 7.0 (median = 7.0) and ranged from 3.5 (Justin; median = 3.0) to 8.5 (Nick; median = 8.0). The decrease in Justin's mean maintenance story elements' score represented more than a 20% decline from his post-treatment mean score. However, his mean maintenance score was still higher than his mean baseline score (mean = 1.7, median = 1.0). During maintenance, most students' slopes remained similar to post-treatment slopes and ranged from 0.0 (Cecilia and Justin) to 1.0 (Justin and Michelle). Nick's maintenance slope declined going from 0.6 (post-treatment) to -1.0 (maintenance) All students had stable data during the maintenance phase (variability = 0.0%).

I used visual analysis procedures to evaluate whether experimental control had been established for the number of essential story elements included in writing (see Table 12). Baseline data were low and stable for every student. Experimental control was

Table 12. *Data Changes Used to Determine Experimental Control: Baseline Stability and Anticipated Level, Slope, and Variability Changes*

Writing Outcome Measure	Baseline	Treatment				Post-treatment			Maintenance		
Student	Low and Stable	> *M*	Improvement = Prompt	> Slope	> Var.	> *M*	Positive Slope	Stable	> *M*	Positive Slope	Stable
Story Elements (0 to 14)											
Cecilia	Yes	**1.6***	**4***	**1.2***	**14.3***	**3.7***	No	Yes	**0.7***	Yes	Yes
Justin	Yes	**2.7***	**4***	**0.8***	**42.9***	**0.9***	**Yes***	Yes	-1.8	Yes	Yes
Michelle	Yes	**3.2***	**4***	**0.8***	**14.3***	**1.3***	**Yes***	Yes	**1.3***	Yes	Yes
Nick	Yes	**1.8***	5	**1***	**28.6***	**1.7***	**Yes***	Yes	**2.2***	No	Yes
Alexis	Yes	**1.7***	3	**0.5***	0.0	ND	ND	ND	ND	ND	ND
Lucas	Yes	**2.1***	3	**1.1***	**33.3***	**2.8***	**Yes***	Yes	ND	ND	ND
Holistic Writing Quality (0 to 8)											
Cecilia	Yes	0.6	PND = 0.0%	-0.6	0.0	-0.3	No	Yes	0.7	Yes	Yes
Justin	Yes	**1.3***	4	**0.4***	0.0	**0.1***	Yes	Yes	0	Yes	Yes
Michelle	Yes	**0.8***	7	**0.9***	14.3	**1.4***	No	Yes	0	Yes	Yes
Nick	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Alexis	No	NC	NC	NC	NC	ND	ND	ND	ND	ND	ND
Lucas	Yes	**0.7***	3	**0.2***	33.3	**0.7***	No	Yes	ND	ND	ND

Table 12 continued. *Data Changes Used to Determine Experimental Control: Baseline Stability and Anticipated Level, Slope, and Variability Changes*

Writing Outcome Measure / Student	Baseline	Treatment				Post-treatment			Maintenance		
	Low and Stable	> *M*	Improvement = Prompt	> Slope	> Var.	> *M*	Positive Slope	Stable	> *M*	Positive Slope	Stable
Written Words											
Cecilia	Yes	9.3	4	**7.1***	14.3	1	Yes	Yes	-2.3	Yes	Yes
Justin	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Michelle	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Nick	Yes	-9.5	PND = 0.0%	**2.2***	0	11.4	Yes	Yes	32	No	No
Alexis	No	NC	NC	NC	NC	ND	ND	ND	ND	ND	ND
Lucas	Yes	-3.7	3	**2.7***	33.3	6	No	No	ND	ND	ND

Notes. * = indicates that experimental control was established. Experimental control was defined as at least three demonstrations of the desired change, occurring only after baseline stability had been established. Comparisons were always made to adjacent condition. Post-treatment data were collected within four weeks of treatment ending. Maintenance data were collected six weeks after treatment ended. *M* = Mean. Prompt = The corresponding prompt in which student's data exceeded highest baseline data. For written words, data had to improve by more than one word. Slope = Did treatment slopes increase? Did post-treatment and maintenance slopes maintain or increase? Var. = variability was the percentage of data that lay outside the mean phase line. Variability was calculated by first multiplying .2 by student's highest score (number of words written) or by highest possible score (story elements = 14; holistic quality = 8). This number was then added and subtracted to each mean. The percentage of data that lay outside those numbers was used to describe variability. NC = Not calculated due to data instability. ND = No data.

established during treatment for positive changes to level, slope, and variability for every student. The students consistently began to show improvement between treatment prompts three and four, and continued to improve or maintain improvements throughout the treatment phase. Experimental control was also established during post-treatment for maintaining post-treatment levels (5 students), improving stability (5 students), and maintaining positive slopes (4 students, see Table 12). Experimental control was also established during maintenance for maintaining post-treatment levels (Cecilia, Michelle, and Nick) and maintaining stability. These results are meaningful and consistent with effects observed with typical writers (see Tracey, Reid, and Graham, 2009).

Data were also analyzed to determine what essential story elements improved in students' writing: frequency of use of each essential story element and elaboration (see Table 13). During baseline, students included few essential elements in their story writing, and did so inconsistently. The total number of story elements in baseline writing was: three (Cecilia), four (Justin and Lucas), five (Michelle), six (Nick), and seven (Alexis, see Table 13). The most frequently included element, across all students was what happened. The frequency this element was included in baseline writing ranged from 67% (Justin) to 100% (Cecilia, Michelle, and Alexis). All of the students also included a *where* in their baseline writing samples, but did so less frequently (33% to 80%). Additionally, during baseline, only Alexis elaborated (received a score of 2 for that story element) on the following essential story elements: what happened (2/8; 25%), when (1/8; 13%), and where (1/8; 13%).

During treatment through maintenance phases, all students improved in the frequency of use and elaboration of the seven essential story elements (see Table 13). The

Table 13. Essential Story Element Parts Included in Students' Writing: Frequency of Use and Elaboration

	Baseline Phase		**Treatment through Maintenance Phases***	
Student	**Story Elements** (Frequency Included in Writing)	**Elaboration**	**Story Elements** (Frequency Included in Writing)	**Elaboration**
Cecilia	What Hap. (3/3; 100%) Where (2/3; 67%) Who (1/3; 33%)	None	Who (12/13; 92%) What Hap. (11/13; 85%) Feelings (10/13; 77%) Where (9/13; 69%) Ending (9/13; 69%) When (7/13; 54%) What Wants (6/13; 46%)	What hap. (5/13; 38%) When (1/13; 8%) What wants (1/13; 8%)
Justin	What Hap. (2/3; 67%) Who (1/3; 33%) When (1/3; 33%) Where (1/3; 33%)	None	What Hap. (12/13; 92%) When (10/13; 77%) Where (6/13; 46%) What Wants (6/13; 46%) Who (5/13; 38%) Ending (5/13; 38%) Feelings (5/13; 38%)	When (6/13; 46%) Where (1/13; 8%) What hap. (1/13; 8%)
Michelle	What Hap. (5/5; 100%) Who (2/5; 40%) Where (2/5; 40%) Ending (1/5; 20%) Feeling (1/5; 20%)	None	Where (13/13; 100%) What hap. (12/13; 92%) When (11/13; 85%) Ending (11/13; 85%) Feelings (11/13; 85%) Who (9/13; 69%) What Wants (7/13; 54%)	What hap. (4/13; 31%) When (3/13; 23%) Where (2/13; 15%)
Nick	What Hap. (4/5; 80%) Where (4/5; 80%) Who (2/5; 40%) When (2/5; 40%) What Wants (1/5; 20%) Ending (1/5; 20%)	None	What Hap. (12/13; 92%) Who (10/13; 77%) Ending (10/13; 77%) When (9/13; 69%) Where (9/13; 69%) What Wants (9/13; 69%) Feelings (9/13 69%)	What hap. (5/13; 38%) When (1/13; 8%) What Wants (1/13; 8%)
Alexis	What Hap. (8/8; 100%) Who (8/8; 100%) When (7/8; 88%) Where (6/8; 75%) Ending (4/8; 50%) What Wants (2/8; 25%) Feelings (1/8; 13%)	What Hap. (2/8; 25%) When (1/8; 13%) Where (1/8; 13%)	What Hap. (8/8; 100%) Where (6/6; 100%) Ending (6/6; 100%) Who (5/6; 83%) When (5/6; 83%) What Wants (3/6; 50%) Feelings (3/6; 50%)	When (2/6; 33%) Where (2/6; 33%) What hap. (1/6; 17%) What wants (1/6; 17%)
Lucas	What Hap. (7/8; 88%) Where (5/8; 63%) Ending (3/8; 38%) Feelings (1/8; 13%)	None	What Hap. (9/9; 100%) Who (8/9; 89%) When (5/9; 56%) Where (5/9; 56%) Ending (5/9; 56%) What Wants (5/9; 56%) Feelings (5/9; 56%)	What hap. (4/9; 44%)

Notes. * = Alexis' data reflect *treatment* writing prompts. Lucas' data reflect *treatment and post-treatment* writing prompts. The remaining four students' data reflect treatment, post-treatment, and maintenance writing prompts. Elaboration = Scored two points on a 0-2-point scale. Hap. = Happened.

what happened was still the most frequently included story element during these phases, however the students consistently added more, and often elaborated on those elements. As can be seen in Table 13, every student included all seven story elements, at some point. Students (n = 5) who had not elaborated at all in their writing during baseline, now elaborated on at least one (Lucas) to three (Cecilia, Justin, Michelle, and Nick) story elements. Alexis also improved. During the treatment phase, Alexis elaborated on an additional story element (when) and more frequently included the seven parts in her writing. During treatment through maintenance phases, all six students failed to elaborate on the follow story elements: *who*, *ending*, and *feelings*. These results are discussed further in the recommendations for future research later in the paper.

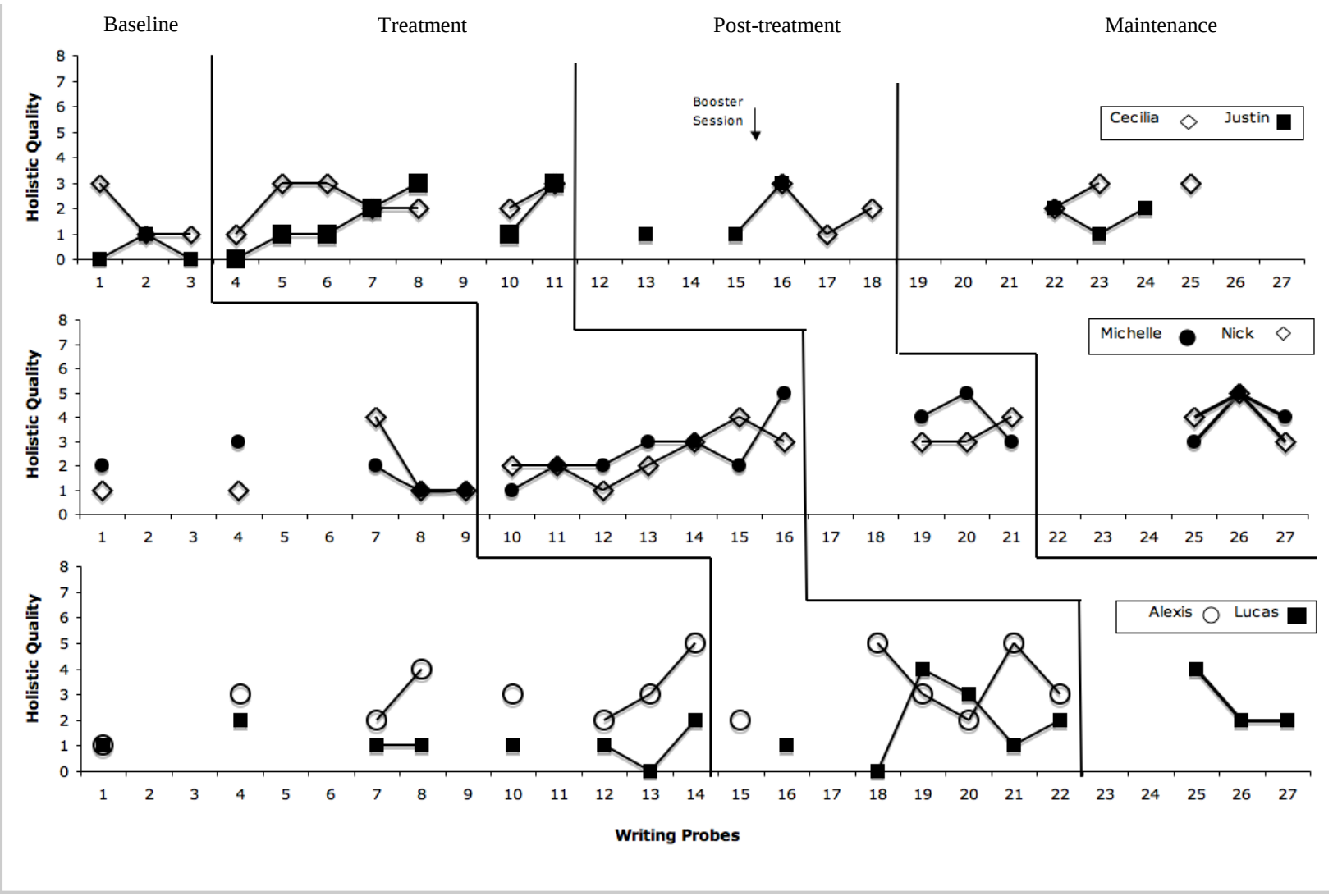
In summary, the results of this study indicated consistently positive effects of the nonprofessional adult-led SRSD story-writing instruction on improving the number of essential story elements. All six students showed similar rates of improvement during treatment (mean and median changes, PND scores, and slopes). The students' data continued to improve during post-treatment. The five students with post-treatment data on average included between 0.9 (Justin) and 3.7 (Cecilia) more story elements in their writing than were included during the treatment phase. Three of the students' maintenance data revealed consistently high story element scores with only one student's data decreasing by more than 20%. Gains were made in treatment (6/6 students) and post-treatment (5/5 students) phases and were maintained (3/4) by the majority of the students. Additionally, experimental control for essential story elements was established during treatment (Cecilia, Justin, Michelle, Nick, and Lucas), post-treatment (Cecilia, Justin, Michelle, Nick, and Lucas), and maintenance phases (Cecilia, Michelle, and Nick).

Finally, all students included more essential elements as a result of the writing treatment, and consistently improved in elaboration of between one and three story elements. Next, I describe the results for holistic writing quality scores.

Holistic writing quality. I examined if students wrote stories of improved quality (mean and median) during the treatment phase. Additionally, did the students maintain improved writing performance during post-treatment (within four weeks after treatment ended) and maintenance phases (six weeks after treatment ended)? I anticipated that students' post-treatment data would be higher than their baseline data and that their maintenance data would decrease by no more than 20% from post-treatment levels (mean and median). Data for holistic writing quality are displayed in Table 11 and Figure 3. Data are analyzed as they were for story elements.

Baseline data were low and stable (no data points fell further than 20% away mean scores) for four of the six students (see Table 11). The average baseline writing quality score across the six students was 1.6 with a range of 0.3 (Justin) to 2.9 (Alexis). Cecilia and Justin's baseline data were low and stable. Cecelia's baseline mean was 1.7 (median = 1) and stable (variability = 0.0%). Justin's baseline data were low (mean = .03; median = 0.0) with little variability (0.0%). In the second tier, Michelle's baseline quality scores were also low (mean = 1.8; median = 2.0) and stable (variability = 0.0%). In the third tier, Lucas's baseline data were also low (mean = 1.1; median = 1.0) and stable (variability = 0.0%). Additionally, these four students' baseline slopes were either negative or flat (see Figure 3). The remaining two students, Nick and Alexis, had a variability score of 20% and 25%, respectively. Both of these student's slopes were relatively level, however (Nick = 0.0; Alexis = 0.3).

Figure 3. *Effect of SRSD Story-Writing Instruction on Holistic Writing Quality*



The students' quality scores improved only slightly during the treatment condition (mean and slope changes). The average writing quality score received across all participants during treatment was 2.3 with a range of 1.6 (Justin) to 3.3 (Alexis). All six students had mean quality scores that were slightly higher than their baseline data, but for most students this improvement meant less than a 1.0 gain. The exception to this was Justin who had a mean quality score of 0.3 during baseline and 1.6 during treatment. Additionally, the PND score between baseline and treatment phases were consistently low. The PND scores were 0.0% (Cecilia, Nick, and Alexis), 14.3% (Michelle, 1/7), 33.3% (Lucas, 2/6), and 42.9% (Justin, 3/7). Three students started showing evidence of improvement after lesson 4 (Lucas, prompt 3), Lesson 5 (Justin, prompt 4), and Lesson 7 (Michelle, prompt 7). Alexis began showing improvement in her holistic quality scores prior to the intervention starting. The PND scores for quality were lower than expected and will be discussed further in the discussion section.

The majority of students' slopes also showed small improvements during the treatment phase. Five students had negative (Cecilia = -1.0; Michelle = -0.4) or flat (Justin, Nick, and Lucas = 0.0) slopes during the baseline phase and positive slopes during the treatment phase (see Table 11). There were also some small changes in variability during the treatment condition. The data for students in the first tier (Cecilia and Justin) remained consistent with baseline data (0.0% to 0.0%). Data for Michele, Lucas, and Alexis show an increase in the variability (Michelle = 0.0% to 14.0%; Lucas = 0.0% to 33.3%; Alexis = 25.0% to 33.3%; see Table 11). Nick's data became less variable (20.0% to 0.0%).

Post-treatment data for quality scores revealed an improvement of approximately 1.0 point when scores (mean and median) were compared to the proceeding treatment phase for three students (Michelle, Nick, and Lucas). Michelle's post-treatment quality scores (*M* and *Mdn* = 4.0) were an improvement from treatment (*M* = 2.6; *Mdn* = 2.0) and baseline data. Nick's post-treatment quality scores (*M* = 3.3; *Mdn*= 3.0) were an improvement from treatment (*M* = 2.4; Mdn = 2.0) and baseline data. Lucas's post-treatment quality scores (*M* = 2.7, *Mdn* = 2.0) were an improvement from his treatment (*M* = 1.8, *Mdn* = 1.5) and baseline data. Cecilia and Justin had post-treatment data (Cecilia's *M* and *Mdn* = 2.0; Justin's *M* and *Mdn*= 1.7 and 1.0) that were relatively consistent to their treatment data (Cecilia's *M* and *Mdn* = 2.3 and 2.0; Justin's *M* and *Mdn* = 1.6 and 1.0). Finally, three students with positive slopes during treatment (Cecilia = 0.1, Michelle = 0.5, Lucas = 0.2) had negative slopes during post-treatment (see Table 11). All students' data were stable.

The quality of student's stories during the maintenance phase showed sustained (*M* and *Mdn*; Justin and Michelle) or continued improvement (*M* and *Mdn*; Cecilia and Nick). Justin maintained an average quality score of 1.7 (*Mdn* = 2.0) during the maintenance phase. This average was consistent with his treatment and post-treatment data and an overall mean increase of 1.3 points from his baseline data (*Mdn* change = 2.0). Michelle maintained the quality in her writing as well (*M* and *Mdn* = 4.0). Her overall mean increase from baseline condition was 2.2 (*Mdn* change = 2.0). Cecilia improved slightly during the maintenance phase (*M* = 2.7; *Mdn* = 1.0), achieving a 1.0 difference from her baseline mean score (*Mdn* change = 2.0). Nick also continued to write stories of higher quality during the maintenance phase (*M* and *Mdn* = 4.0) with an

overall baseline to maintenance mean increase of 2.5 (*Mdn* change = 3.0). Students' slopes during the maintenance phase were inconsistent (Justin = 0.0; Cecilia and Michelle = 0.5; Nick = -0.5). All students' data were stable (see Table 11).

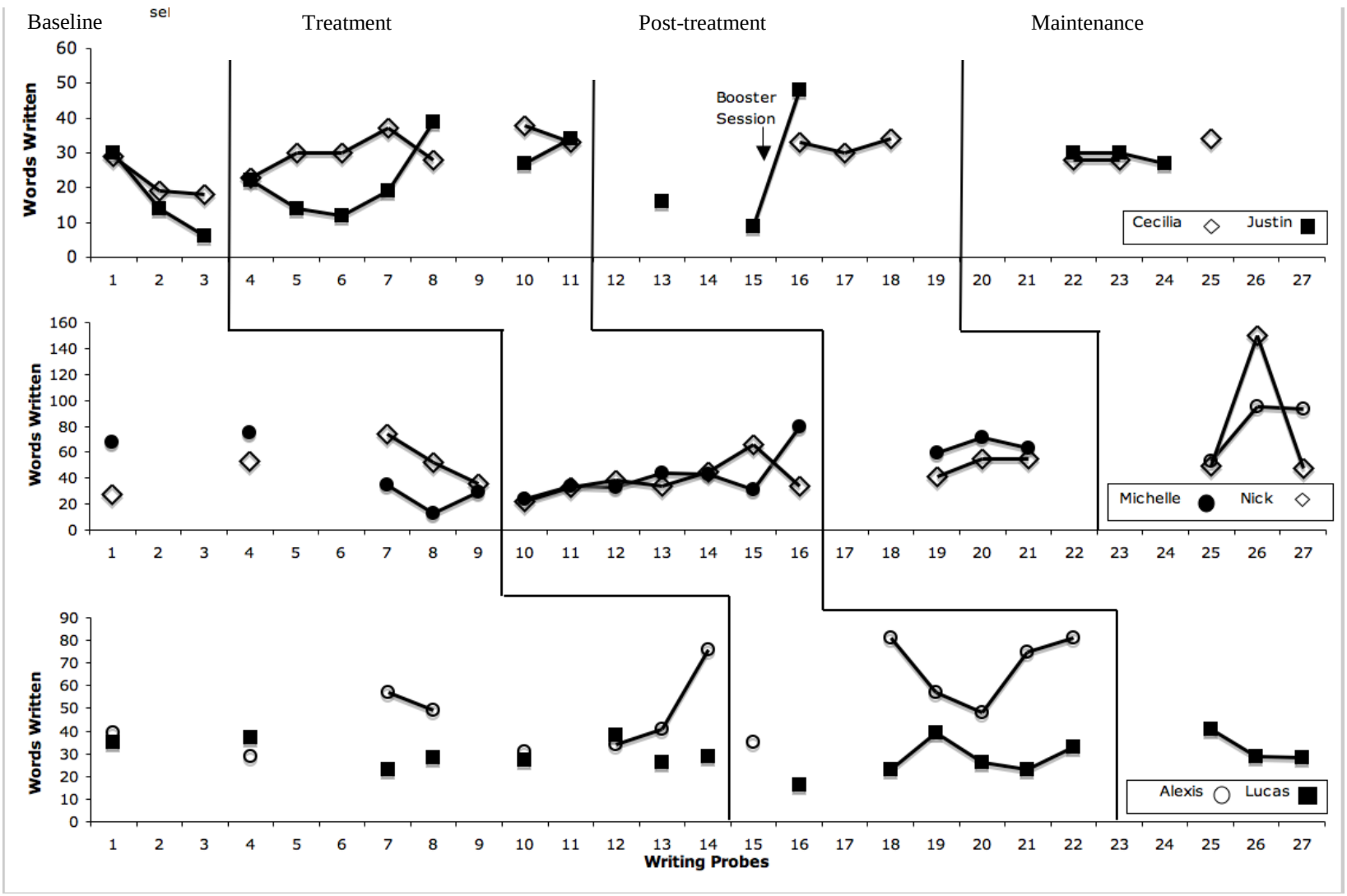
I used visual analysis procedures to evaluate whether experimental control for holistic writing quality had been achieved (see Figure 3 and Table 11). Experimental control was established during treatment for positive changes that occurred to levels and slopes for Justin, Michelle, and Lucas. The increase of the mean holistic quality scores was lower than expected and will be addressed further in the discussion section. The PND scores (comparing baseline to treatment phases) exceeded 0.0% for Justin (42.9%, 3/7), Lucas (33.3%, 2/6), and Michelle (14.3%, 1/7). Baseline data for these students were also stable. Justin, Michelle, and Lucas's mean quality scores from baseline to treatment ranged from 0.6 (Lucas) to 0.8 (Michelle) to 1.3 (Justin). These students' slopes also improved during treatment phase, increasing by 0.4 (Justin), 0.9 (Michelle), and 0.2 (Lucas). Experimental control was established during the post-treatment phase as well for positive changes that occurred to overall levels (Justin, Michelle, and Lucas). Experimental control was not established during the maintenance phase as only Justin and Michelle had these data (Cecilia's improvement was not included as she had not demonstrated gains during treatment).

In summary, the results of this study indicated small, but positive results during treatment for the effect of the nonprofessional adult-led SRSD story-writing instruction on improving writing quality for struggling third-grade writers (research question two and three). All six students showed an improvement during the treatment phase in their slope and mean/median data. The slopes ranged from -1.0 to 0.3 during baseline and from

0.1 to 0.5 during treatment. Second, the students maintained improved quality scores during post-treatment and maintenance phases. One student's data revealed post-treatment scores that were consistent with treatment scores and higher than baseline scores. Three other students' data revealed continued improvements during the post-treatment phase. Maintenance data revealed maintained or improved quality scores. Although small gains were made across each student, experimental control for holistic writing quality was established at treatment and post-treatment for Justin, Michelle, and Lucas (James, Jessica, and Phyllis's tutees, respectively).

Number of words written. Data on the total number of words written is displayed in Table 11 and Figure 4. Data are described as they were for story elements and holistic quality. Baseline productivity data were stable for three students (Cecilia, Nick, and Lucas) and highly variable for the remaining students (Justin, Michelle, and Alexis; see Table 11 and Figure 4). On average, students wrote approximately 34 words per prompt during the baseline phase. The students with stable baseline data were spread across the three tiers of instruction. Cecilia wrote an average of 22.0 (median = 19.0) words on her baseline prompts. Nick wrote an average of 48.4 (median = 52.0) words on his writing prompts. Lucas wrote an average of 30.4 (median = 28.5) on his writing prompts. The remaining three students' baseline data ranged 16.7 (Justin, median = 14.0) to 44.0 (Michelle and Alexis, medians = 35.0 and 40.0). Baseline data for Cecelia, Justin, Michelle, and Lucas show a negative slope indicating that the number of words written decreased during baseline. Nick and Alexis had positive slopes indicating that an upward trend in their baseline data had occurred.

Figure 4. *Effect of SRSD Story-Writing Instruction on Number of Words Written*



During the treatment phase, the average number of words written ranged from 23.9 (Justin, median = 22.0) to 62.8 (Alexis, median = 66.0). These averages represented a mean increase for Cecilia (22.0 to 31.3), Justin (16.7 to 23.9), and Alexis (44.5 to 62.8). Similar positive changes occurred when median scores were used for analysis. The PND scores between baseline and treatment phases were consistently greater than 0.0%, with most percentages reflecting one to two treatment data points that exceeded the highest baseline data points. The PND scores for number of words written were 0.0% (Nick), 14.3% (Michelle, 1/7), 16.7% (Lucas, 1/6), 28.6% (Justin, 2/7), 33.3% (Alexis, 2/6), and 71.4% (Cecilia, 5/7). The students started showing evidence of improvement after Lesson 2 (Alexis), Lesson 4 (Lucas), Lesson 5 (Cecilia), Lesson 6 (Justin), and Lesson 7 (Michelle). The PND scores and the variability in the onset of improvement for number of written words were as expected, as the goal of the intervention was not necessarily to have students write more, but rather to include more essential story elements in their writing. These results will be discussed further in the discussion section.

During treatment, the trend or slope of the treatment data improved (became more positive) for all six students: Cecilia (-5.5 to 1.6), Justin (-12.0 to 3.2), Michelle (-14.0 to 6.1), Nick (1.7 to 3.9), Alexis (2.8 to 5.8) and Lucas (-0.6 to 2.1). Some students' (Cecilia, Alexis, and Lucas) data became more variable during the treatment phase indicating that more data points fell further than 20% away from the student's mean post-treatment score. One student's data became less variable during the treatment phase (Michelle) whereas the variability data for two other students (Justin and Nick) remained relatively unchanged.

During the post-treatment phase, the mean/median scores for total words written ranged from 24.3/16 (Justin) to 64.3/63 (Michelle). When compared to treatment data, most students' post-treatment mean and median scores did not show a large change (negative or positive). For example, Cecilia went from writing an average of 31.3 words (*Mdn* = 30.0) during treatment to 32.3 words (*Mdn* = 33) during post-treatment. Justin and Lucas mean and median score changes showed similar results. The slopes across students' post-treatment productivity scores showed a general tendency to decrease from treatment slopes (Cecilia, Michelle, and Lucas). Two students showed an increase in their productivity slope data (Justin and Nick). Most of the students' variability data remained consistent with their treatment data (Justin, Nick, and Lucas). Two other students' post-treatment data became less variable (Cecilia and Michelle, see Table 11).

Maintenance data were collected six weeks after the treatment ended. Again, for the most part, students' data were very similar to their post-treatment data, with the exception of becoming more stable for one student (Justin) and increasing means for two students (Michelle and Nick).

I used visual analysis procedures to determine whether experimental control for number of words written had been established. Three students had baseline data that were low and stable: Cecilia, Nick, and Lucas. Experimental control was established during treatment for improved data slopes for Cecilia, Nick, and Lucas (see Table 11). Experimental control was not established for positive mean changes. When comparing baseline and treatment data, Cecilia was the only student to write more. Additionally, no experimental control was established for post-treatment and maintenance phases.

In summary, the results of the nonprofessional adult-led story-writing instruction on improving the number of words written by students showed some positive changes, but these changes were not consistent across all students. Experimental control was established for improving slopes during the treatment condition (Cecilia, Nick, and Lucas). The students' baseline slopes ranged from -14.0 (Michelle) to 2.8 (Alexis). Most of these slopes were negative (Cecilia = -5.5, Justin = -12.0, Michelle = -14.0, and Lucas = -0.6). The students' treatment slopes ranged from 1.6 (Cecilia) to 6.1 (Michelle). Experimental control was not established for effecting positive and consistent changes to level or variability during treatment, post-treatment, or maintenance phases. The results are not surprising as the goal of the SRSD instruction was to increase the number of essential story elements and not, necessarily, to increase total words written.

Social Validity

The fourth research question examined the overall acceptability of the SRSD story-writing treatment for both students and nonprofessional adults. Specifically, did students and/or nonprofessional adults report high social validity for the SRSD story-writing instruction? I predicted that both nonprofessional adults and students would find value in the SRSD story-writing instruction and would recommend its continued use. These results are provided next.

Students. Social validity was assessed from both students' and nonprofessional adults' perspectives. I describe the students' social validity data first (see Table 14). I met with each student individually after the writing treatment had ended. The students were asked and answered a series of seven questions (see Table 14). I hypothesized that

Table 14. *Social Validity Responses from Student Participants*

Question	Responses
Should POW be taught to other students?	YES (n = 6) Michelle said, "Yes, because they could be story writers when they grow up." She went on to say that she is using POW to help her with her song writing.
Should WWW, What = 2, How = 2 be taught to other students?	YES (n = 5) Justin said he liked the WWW trick better than POW. Lucas was the only student to say, "No". When asked why it shouldn't be taught he said, "Because. It's too long. I couldn't remember it."
What did you like about how your teacher taught you POW and WWW?	POSITIVE COMMENTS (n = 6) Nick and Alexis said they liked the testing. They said the flashcards were a fun way to start each lesson. Nick and Michelle stated they liked the individual help that Ms. Jessica provided. They said she helped them "get it". Nick said he especially liked the way Ms. Jessica explained things. Cecilia and Justin both said they liked how Mr. James taught the tricks. When asked further about this they stated that they couldn't really explain. They just liked it. Lucas said that his teacher, Ms. Phyllis, was very fun. When asked if he liked it when she put a sticky note with the word "what" on her forehead, he started to chuckle. Additionally, Alexis said she liked looking at and writing about all the different pictures and going to the counselor's room to get million dollar word examples.
What did you dislike?	NEGATIVE COMMENT (n = 1) Most students stated "nothing" (n = 4). When I continued to ask, most students repeated, "nothing". When queried further on this issue, Cecilia looked at me and stated. "The only thing I don't like is that the Writing Academy ended." Justin stated that he did not like missing class.
How would you change the way you were taught?	SHORTER LESSONS/ TRICKS (n = 3) Most students at first said, "nothing". When I continued to ask about this issue, Nick suggested that we shorten each lesson. He said that sometimes he got tired. Michelle stated that sometimes her hand hurt from writing so much. Lucas suggested that we shorten the tricks that they have to memorize. Cecilia suggested we let students come up with their own tricks. MORE GAMES (n = 1, Michelle). Michelle said it would be fun to play more games during the lesson. When asked about this she did not have any specific examples.

Table 14 continued. *Social Validity Responses from Student Participants*

Question	Responses
How would you change the way you were taught? *(continued)*	NO IDEAS (n = 2) Justin said he didn't know. Alexis said she wouldn't change anything.
Have you used anything you learned at home or in another class?	YES (n = 5) Justin said he didn't know.
Tell me what you did.	HOME (n = 3) Lucas wrote about popcorn. Michelle wrote songs and letters to her grandmother. Cecilia said she wrote about her dog and hamster. SCHOOL (n = 2) Nick wrote a story in school during free time and used the who, when, and where. Alexis said she's transferred a lot from writing letters to book reports to school reports to personal journals. NOTHING (n = 1) Justin said he had not done anything, but still might.
Thumbs up?	YES (n = 6). Alexis said the Writing Academy changed her life. She said she couldn't spell, but now could write good stories.

students would report high social validity for the SRSD story-writing instruction. The data are consistent with this prediction.

As can be seen in Table 14, students expressed positive feelings about the writing treatment and the tricks learned. All students stated that POW should be taught to other students. Michelle stated, "Yes, because they could be story writers when they grow up." The majority of students (5/6; 83.3%) of the students stated that the WWW, What = 2, How = 2 trick should also be taught to other students. However, Lucas stated that this

trick should not be taught because it was too long (see Table 14). Students were also willing to discuss what they liked about *how* their teacher taught the POW and WWW trick. Nick and Alexis liked the testing that occurred at the beginning of every lesson. They stated that they enjoyed the rapid fire drill with flashcards. Nick and Michelle stated they liked the individual help Ms. Jessica provided. They said she helped them "get it". Nick said he especially liked the way Ms. Jessica
explained things. Cecilia and Justin both liked how Mrs. James taught the tricks (no elaboration was provided). Lucas said he liked how Ms. Phyllis was fun. When asked if he liked when Ms. Phyllis put a sticky note with the word "what" on her forehead, he began to chuckle. Finally, Alexis said she liked looking at and writing about all the different pictures and going to the counselor's room to get examples of million dollar words.

Most students stated nothing negative about how they were taught. One student expressed a dislike about being out of class. Justin stated he did not like missing class or getting behind in schoolwork as a result of attending Writing Academy lessons. Cecilia stated that the only thing disliked was the Writing Academy ending.

Students came up with several suggestions to improve the Writing Academy lessons. Three of the students suggested shortening the lessons or shortening and changing the tricks that they were taught. For example, Nick suggested that total time spent in each lesson decrease. He stated he sometimes got tired at the end of the lesson. Michelle stated her hand sometimes hurt at the end of the lesson. Lucas suggested that we shorten the tricks, because the WWW trick was too long and he had a difficult time remembering all of the parts. Cecilia suggested we let the students come up with their

own mnemonics or tricks. Michelle also suggested we include more games in the lessons (specific examples were not provided). Justin said he had no idea how he would change it. Alexis said that she would not change a thing.

Five of the six students stated they transferred the writing tricks to other writing situations. Three of the students said they were writing more at home. Two students said they were using the tricks in school. Justin said he was not using these tricks, but said he still might. Overall, all students gave the SRSD story-writing treatment a "thumbs up". Alexis stated the Writing Academy changed her life.

Nonprofessional adults. During the treatment phase the nonprofessional adults were asked to maintain a journal and were asked to address at least two to three of the following questions after every lesson: (a) Was what you did with the student today effective? Why or why not?, (b) What do you feel is working well or not working well? (Be specific and tell why.), and (c) Is there anything that is happening while you are teaching that is making you less effective? (Be specific.). They were further asked to rate the effectiveness of their instruction that day from "4" (very effective) to "1" (not effective).

After the treatment ended, the journal entries were typed into a word document. I read through the approximately 4,400-word 18-page document several times until the following themes were identified: *student behavior* (e.g., focused, distracted, engaging), *tutor feelings/effort* (e.g., feeling ineffective, expressing excitement, lesson modification), *student ability* (e.g., capable), and *student performance* (e.g., improving, not improving, knew all of the tricks).

Each tutor maintained his/her journal throughout the study, writing an entry for each lesson taught. Jessica and Phyllis provided an effectiveness rating for each lesson taught. James wrote entries for each lesson taught, but did not begin assigning an effectiveness rating of 1 to 4 until the fourth day of treatment. Overall effectiveness was rated as high as 3.8 for Michelle by Jessica (the tutor) and as low as 2.2 for Cecilia by James (the tutor). Phyllis rated effectiveness as 3.1 for Alexis and 2.5 for Lucas. These effective ratings are congruent with journal comments and students' overall performance during treatment (discussed at end of social validity section). The effectiveness ratings for James' students, Cecilia and Justin, are displayed in Figure 5. The effectiveness ratings for Jessica's students are displayed in Figure 6. The effectiveness ratings for Phyllis's students are displayed in Figure 7. Students' data are represented with same geometric shapes as in Figures 2, 3, and 4: James: Cecilia (open diamond), Justin (closed square); Jessica: Michelle (closed circle), Nick (open diamond); Phyllis: Alexis (open circle), Lucas (closed square). These data will be further addressed in the discussion section.

James's journal entries were the longest of the three tutors. James wrote approximately 2300 words in 18 journal entries. He indicated tentative approval of the treatment, stating: "Ultimately, she knows the tricks. Only time will tell if she improves" and "Hopefully he draws from all of the tricks". Jessica indicated high approval stating "I really enjoyed working with these students. I feel this project has been most beneficial to the students." These seem to accurately reflect the 2.2 and 2.3 rating for James and the 3.6 and the 3.8 rating for Jessica. Phyllis frequently indicated an overall excitement in her

Figure 5. Overall Effectiveness (0 to 4) of Each SRSD Session: James' Journal Ratings

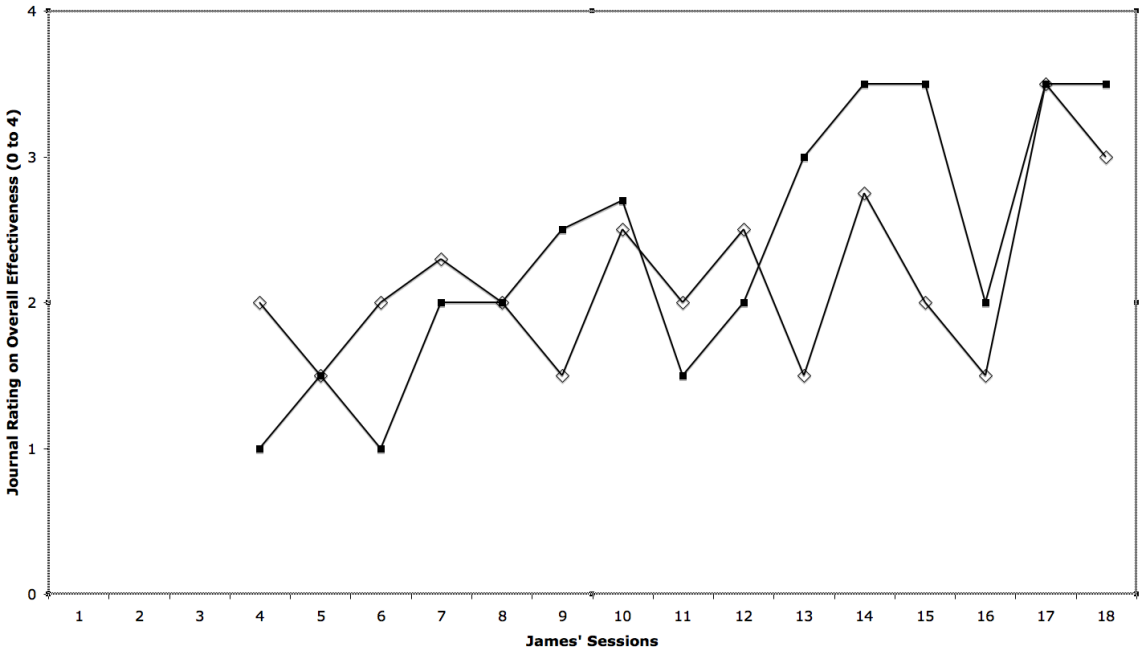


Figure 6. Overall Effectiveness (0 to 4) of Each SRSD Session: Jessica's Journal Ratings

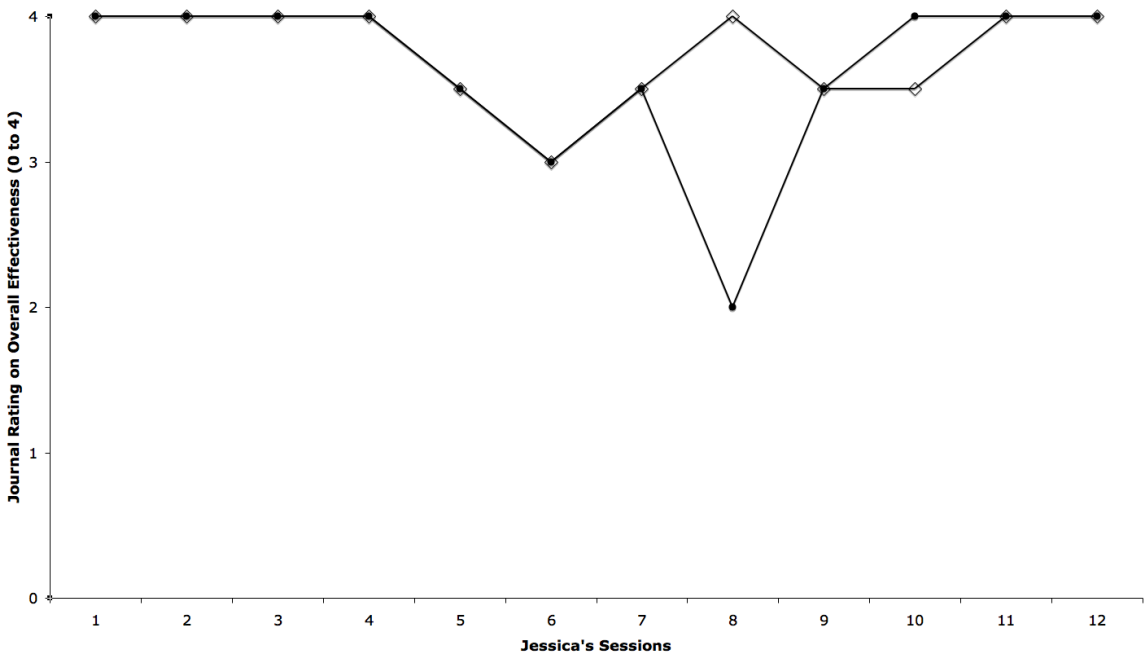
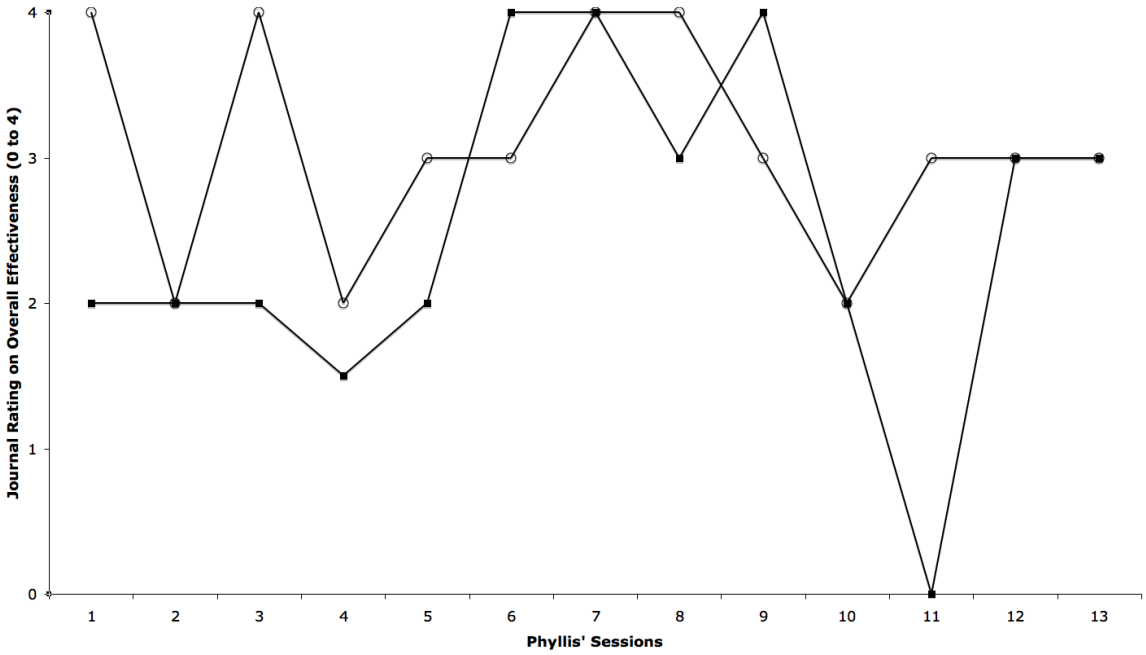


Figure 7. Overall Effectiveness (0 to 4) of Each SRSD Session: Phyllis' Journal Ratings



journal entries (e.g., "It was SO exciting!!"), but this excitement seemed to be tempered by the difficulty engaging Lucas and helping Alexis transfer notes into a story (e.g., "I hoped she would transfer her notes a little better…her thoughts didn't seem to flow and make sense"). Next, I more fully describe each tutor's journal entries. Information is presented first for James and his tutees, Cecilia and Justin. Next I describe the entries related to Jessica and her tutees, Michelle and Nick. Finally, I describe the entries related to Phyllis and her tutees, Alexis and Lucas.

James primarily addressed the following in Cecelia's journal entries: (a) tutor feelings or effort (14/18 entries), (b) behavior/affect (13/18 entries), (c) overall performance (13/18 entries for each), and (d) student ability (11/18 entries). James described changes he made (or needed to make) to lessons to make them more effective: "For lesson to be better will work together (her doing most of work) on collaborative

writing" and "Changing roles today brought on some improvement". Second, James wrote about the positive impact of Cecelia's *motivation* on overall lesson effectiveness: Cecilia "wanted to learn about writing", "tries hard to work on material", and "is willing to learn". Additionally, James commented on the positive impact of Cecilia's engagement: Cecilia was "engaging", "as engaging as yesterday", "very engaging and tries hard to work on material", and "easy to work with and willing to learn". As the lessons progressed, James wrote less about Cecelia's willingness to learn and wrote more about her lack of participation: "drifted a tad when I was lecturing", "goes off focus", and "is not always trying to giver her best effort and will drift".

When writing about Cecilia's overall performance. James wrote that Cecilia knew all of the tricks by day 7 and specifically wrote about her improvement (3 days), or lack of improvement (3 days) during the next 11 lessons: "She is showing portions of improvement" and "A definite improvement". He also wrote, "her writing has not improved" and she "fails to write well on her own". James further wrote about Cecelia's abilities (i.e., her difficulties with writing and spelling) and how some challenges impacted overall lesson effectiveness. Beginning on day 5, James wrote the following type of comments: " her writing is poor", "ultimately she has great potential, but needs stronger framework in spelling, grammar, and penmanship".

James also wrote extensively about Justin's instruction. James primarily wrote about Justin's behavior/affect (16/18 entries), tutor feelings or effort (14/18 entries), overall performance (11/18 entries), and student ability (5/18 entries). James frequently wrote about the negative impact Justin's behavior had on lesson effectiveness. James identified the following behaviors as problematic: lack of effort, being reserved,

mumbling or having poor posture, appearing bored, and working slowly. For example, James frequently wrote comments like, “seems unwilling to put forth extra effort”, “Had pep talk on how this wasn’t for me but for him and how he can have pride with better effort”, and “is unwilling to put forth much effort and is lacking in enthusiasm”.

James also wrote about the positive impact Justin’s behavior had on lesson effectiveness. On 33% of the days (6/18), James only described the positive behaviors exhibited by Justin and how they were beneficial to overall lesson effectiveness. For example, James wrote that Justin ”was somewhat focused” and “worked hard today”. There were three days when both positive and negative behaviors were described in the same journal entry.

The next major theme identified in James’ journal entries for Justin was related to James’ feelings and his effort to make lessons better. These entries were primarily related to feelings of effectiveness. The entries were typically negative with most entries related to the difficulty in engaging Justin: “Had to pull answers from him at times”, “although Justin was there, he did bare minimum”, and “from beginning to end, he gives very little input.” Other entries were mostly positive. At least on three occasions James wrote something to the effect, “has developed enthusiasm. Have hope for him.” James also frequently wrote about the need to change the lesson to make it more effective for Justin: “will try and break memorization drill quickly”, “had to put him in a pass-fail mode”.

James also wrote about Justin’s overall performance. These comments were included in well over half of the entries. James wrote that Justin could recite all the tricks (WWW) by memory by the fifth day in treatment. Beginning on that day and forward, James wrote that he felt Justin was showing improvement, indicating only positive

performance in every entry thereafter. The fourth identified theme was related to Justin's abilities. James wrote that Justin had the ability to succeed. He wrote, "Smart kid", "memory was quite good", "shows good memory", and "the kid is smart!"

Jessica's journal entries were typically in reference to both her students. She primarily wrote positive comments about overall lesson effectiveness (8/12 entries) and expressed enthusiasm (5 entries). Some comments related to overall effectiveness were "both students were able to verbalize understanding of the tricks", "lesson was effective in that students *got it*", and "I really enjoyed working with these students. I feel this project has been most beneficial to the students." Jessica also indicated that some lessons were not as effective due to problems with engaging the students: "taking the ideas/tricks and planning and then writing and saying more is time consuming for students" and she "did not feel well…". Some comments related to enthusiasm for program were "The repetition is oh so very good." and "This is a fun portion!".

Jessica also addressed Michelle and Nick's behavior in her journal entries. Michelle was described as "quite focused" on one day and as bored/tired, distracted, and not feeling well on three other days. Jessica did not indicate when students memorized tricks. Nick's journal entries mirrored Michelle's entries, with the exception of entries related to student behavior. Nick's four entries were evenly split between him showing good focus and participation (2/12) and seeming distracted, bored, or tired (2/12 entries).

Phyllis, like James, described the effectiveness separately for each student. Her entries for Alexis were primarily related to her excitement about Alexis's progress (tutor feelings or effort; 10/13 entries). For example, Phyllis wrote, "I'm excited about seeing her stories blossom", "at one point she closed her eyes to picture some words on her

classroom wall. I was so thrilled!", and "she lit up like a Christmas tree when she saw the difference between her original story and the rewritten one…It was SO exciting!". Phyllis also described her effort to alter lessons to fit Alexis's individual needs (5/13 entries). For example, she wrote, "I didn't need to model much for her", "I felt it important for her to pay closer attention to her "what happens", and "we took her to the guidance room to see lots of descriptive words on the wall". There were also some tentative comments made about overall progress (4/13 entries) such as "we will see if she remembers next time", "I hoped she would transfer her notes a little better", and "hopefully she will make that a new habit" (when discussing crossing off notes with a crayon).

Phyllis's other entries for Alexis dealt with student behavior/affect and how satisfactory participation had a positive effect on overall lesson effectiveness (6/13 entries). For example, Phyllis wrote, "It was easy to work with Alexis – easy to build rapport", "she did a great job making better effort for a strong beginning", and "she was so excited about beefing up her story…". Finally, Phyllis's entries for Alexis related to student performance (6/13 entries). For example, beginning on day 5, Phyllis wrote that Alexis knew all the tricks. She also wrote, "she definitely knows POW" and "she did a great job planning … wrote a fabulous story!"

Most of Phyllis's entries for Lucas were related to her own feelings or effort (13/13), Lucas's overall performance (8/13), and Lucas's behavior (4/13). First, Phyllis's entries were primarily related to her feelings of excitement or reserved excitement. For example, Phyllis frequently wrote the following types of comments: "Made me feel great!", "He was so relaxed today and he completely filled out the organization sheet by

himself! Yes!" There were some statements that were more cautiously optimistic: "I'm looking forward to a better day tomorrow" and "I hope he retains the tricks he learned today, he doesn't give much feedback." Phyllis indicated that she often felt ineffective (5/13 lessons) when teaching Lucas. This was mostly categorized as a problem engaging the student. Phyllis indicated that several other lessons were effective (4/13 lessons).

Phyllis also wrote about Lucas's performance (8/13 entries). She wrote that Lucas knew POW or parts of the tricks by day two and that he was improving by day five. Phyllis indicated that Lucas was not improving on one entry (day three). In all, the comments related to Lucas's performance were mostly positive with comments such as "he did a great job planning and wrote a good story" and "he was so proud of his story". Finally, Phyllis wrote about Lucas's behavior. Phyllis wrote that Lucas did participate although at times seemed nervous or fidgety.

Summary. In summary, the opinion of the nonprofessional adult volunteers about the effectiveness and social validity of the SRSD story-writing treatment were assessed in multiple ways. First, the tutors wrote comments related to lesson effectiveness after each tutoring session. These journals were typed and analyzed according to types of comments made (positive, negative) and themes addressed (e.g., impact of student behavior on lesson effectiveness). Tutors provided either unwavering support of the writing treatment (e.g., Jessica wrote, "I feel this project has been most beneficial to the students. I really enjoyed working with these students") or a tentative approval (James wrote, "Only time will tell if she improves"). All of the tutors commented on the positive *and negative* impact of certain student behaviors or affect (e.g., engagement) and overall lesson

effectiveness. Each tutor also wrote about his/her feelings in the journal and about their efforts to make each lesson better.

Social validity was also measured by calculating the average effectiveness rating provided by each tutor for each tutee's lessons (scores ranged from 0 to 4). Phyllis and Jessica provided effectiveness ratings for all entries. James provided effectiveness ratings for all but the first three entries for his tutees, Cecilia and Justin. Jessica reported the highest overall rating of 3.7 (Michelle = 3.6; Nick = 3.8). Phyllis's average effective rating was the next highest at 2.8 (Alexis = 3.1; Lucas = 2.5). James's average effective rating was the lowest at 2.3 (Cecilia= 2.2; Justin = 2.3).

Finally, I asked the three tutors to summarize their overall experience. All three individuals stated the experience had been very positive and that they would be happy to teach the lessons again. Phyllis also told me that she approached a third-grade teacher and expressed her enthusiasm about the SRSD story-writing instruction. Phyllis reported that she told the third-grade teacher to ask how we can use the program for all third-grade students. Following this conversation, an in-service about SRSD instruction was planned at the request of the third-grade teachers. Additionally, Phyllis stated she would be willing to help lead training sessions for future Writing Academy (SRSD story-writing intervention) instructors at the current school.

Did the nonprofessional adults overall impressions accurately match up with student outcomes? James's overall impressions were the lowest of the three tutors. His tutees, Cecilia and Justin, did take the longest to progress through the SRSD story-writing lessons (17 lessons as compared to 11 and 12 for Janet and Phyllis's tutees). During treatment his tutees also had the lowest mean scores for story elements and holistic

writing quality. Cecilia and Justin had a mean treatment story element score of 4.0 (3.6 and 4.4) which was lower than Jessica's tutees' mean score (5.0; 5.4 and 4.6) and Phyllis' tutees' mean score (5.5; 6.7 and 4.2). Cecilia and Justin's mean treatment holistic quality score of 2.0 (2.3 and 1.6) was also lower than Jessica's tutees' mean score (2.5; 2.6 and 2.4) and Phyllis's tutees' mean score (3.9; 3.3 and 1.8).

After the treatment ended, James' tutees continued to improve yet their average story element score was still below Jessica and Phyllis' tutees' mean story elements scores. The average post-treatment and maintenance story elements scores were 6.0 for Cecilia (M = 7.7) and Justin (*M* = 4.3), 7.3 for Michelle (*M* = 7.5) and Nick (*M* = 7.2), and 7.0 for Lucas. Similarly, overall writing quality scores during post-treatment and maintenance were much higher for Jessica's tutees with means ranging from 3.3 (Nick) to 4.0 (Michelle) as compared to James's tutees who had mean quality scores ranging from 1.7 (Justin) to 2.0 (Cecilia).

CHAPTER V

DISCUSSION

In this chapter, I evaluate and interpret the results of the study. Three trained nonprofessional adult volunteers taught six struggling third-grade writers how to plan and draft better stories using the SRSD model of instruction. The research was guided by four research questions: one question addressed *procedural fidelity* (Were the tutors able to successfully demonstrate each lesson during training sessions? Did the tutors cover at least 90% of each lesson component for each lesson during treatment? Was the tutoring provided during treatment of high quality?), two questions addressed *effects on writing outcomes, during and after treatment* (Did the SRSD story-writing instruction result in improved writing output, as measured by number of essential story elements and number of words written, and improved writing quality?), and one question addressed *social validity* (Did the students and tutors believe the SRSD story-writing instruction was effective. Would they recommend its continued use?).

I predicted that the nonprofessional adult volunteers would demonstrate mastery of lesson material during training sessions and would complete at least 90% of lesson components (for each lesson) during the intervention and that the teaching quality would be high. I also predicted that the SRSD writing treatment would result in improved writing output (number of story elements and total words written) and improve writing quality. I anticipated that these positive results would be maintained for four weeks and six weeks post instruction (post-treatment and maintenance phases). I anticipated that

students' post-treatment data (mean and median scores) would be higher than their baseline data. I also anticipated that student's maintenance data (mean and median scores) would decrease only slightly when compared to their post-treatment data (mean and median; no more than 20%). Finally, I hypothesized that both nonprofessional adults and students would find value in the SRSD story-writing instruction and would recommend its continued use.

In the sections that follow, I first discuss the findings in light of the results obtained in this study. Second, I describe how the current study extends previous research related to SRSD instruction and nonprofessional adult-led literacy interventions. Third, I describe the study limitations. Finally, I summarize and make suggestions for future research and practice.

Results

Procedural Fidelity: Lesson components. Procedural fidelity was 100% during training sessions. Previous examinations of SRSD's effectiveness have highlighted the importance of providing sufficient pre-treatment training that includes opportunities for future tutors to successfully master each lesson (Harris, Graham, and Mason, 2003). The nonprofessional adults in this study successfully applied each lesson component for all lessons during the training sessions. As the individuals in this study were volunteers from the community, the current study demonstrates that including pre-tutoring training sessions, although lengthy (6 to 8 hours), was feasible and successful in the current study (as demonstrated by each tutors ability to apply each lesson component for all lessons). Replication of these procedures may be challenging for schools, as a tutor trainer must be

recruited. In the current study, the last tutor, Phyllis, stated she would be willing to fill this role in the future. The use of nonprofessional adult tutors provides one possible remedy for staffing problems. In addition, trained nonprofessional tutors could train others to provide these services (although such an approach requires further testing). Tutor training (procedures used and total time required) are further addressed under recommendations for future research.

Although the percentage of lesson components successfully covered during every tutoring session did not always meet the pre-specified criteria of 90%, this goal was met when all lessons were averaged together. Each tutor covered at least 92% of lesson components across all tutoring sessions with scores ranging from 70% to 100%. Ten lesson components were occasionally omitted during treatment (see Table 9), but if important to the overall effectiveness of the lesson, were included in future lessons.

It should be noted that early in the experiment I changed procedures and began organizing student folders for each tutor. I gave the folders to tutors before each tutoring session began. I changed these procedures (tutors had this responsibility initially) after observing the challenges that James experienced with organization of tutee materials. This procedural change occurred *after* James tutoring session began, but *before* Jessica and Phyllis's tutoring sessions began and may have resulted in higher percentages of lesson components completed for Jessica and Phyllis. These and other tutor supports (e.g., frequent feedback) that may have had a positive impact on overall high procedural fidelity levels are addressed later under study limitations and recommendations for future research.

Procedural fidelity: Teaching quality. The teaching quality for the nonprofessional-led SRSD story-writing instruction sessions was high. I anticipated high teaching quality scores, as this aspect of instruction was addressed during the pre-treatment tutoring sessions. On average, the tutors met at least 7.0 out of 8.0 quality indicators across all tutoring sessions (see Table 8). James' scores were the lowest at 6.8 for Justin and 7.2 for Cecilia (Jessica's tutees: Michelle = 7.8, Nick = 7.5; Phyllis's tutees: Alexis = 7.5, Lucas = 7.3), but were still high. I provided specific feedback related to omissions (e.g., importance of showing excitement about SRSD tricks or frequently monitoring student progress) during the experiment. Interpretation of the findings for quality must be tempered by the fact that reliability was not established for this measure.

Writing Outcomes. In this section, I discuss findings as they relate to treatment conditions: pre-baseline, baseline, treatment, post-treatment, and maintenance. This structure highlights the findings as students data were mostly low and stable during baseline condition, consistently improved during treatment (slope), and in some cases showed continued areas of improvement during post-treatment and maintenance phases.

Pre-baseline. In discussing pre-baseline data, I first describe the writing practices used across the three third-grade classrooms using the responses provided in the Teacher Survey of Classroom Writing Practices. Next, I describe the baseline data for each writing outcome measure.

The third-grade teachers indicated they used a traditional skills approach to teaching writing combined with process writing. Each teacher spent the majority of her time on the following skills: sentence construction, handwriting, spelling, grammar, punctuation, and capitalization. Additionally they stated they almost always had multiple

instructional goals for writing and had students use graphic organizers when writing. Teachers used procedures that were generally similar to the ones used in the current study (e.g., teaching writing strategies, planning and drafting, modeling, and evaluations). However, in discussion with teachers, there was no evidence they used the exact same procedures as employed by the nonprofessional adult tutors. Thus, the writing programs implemented across classrooms were similar, but dissimilar to the procedures used in this study, and should not have impacted overall results.

Baseline. Baseline data were collected over three (Cecilia and Justin), five (Michelle and Nick), and eight (Alexis and Lucas) writing probes. This number is sufficient to demonstrate data stability (Kennedy, 2005). There were at least three demonstrations of low and stable baseline data for each of the writing outcome measures (essential story elements, holistic writing quality, number of written words; see Table 11). The most consistent data were observed on the essential story elements outcome measure. All students‘ essential story elements’ baseline data were low and stable with means ranging from 1.7 (Justin) to 5.0 (Alexis). The baseline data for holistic writing quality were also low and stable for four students: Cecilia, Justin, Michelle, and Lucas. Their baseline holistic quality data ranged from 0.3 (Justin) to 1.8 (Michelle). The baseline data for number of words written was less consistent across students with only three demonstrations of stability: Cecilia, Nick, and Lucas. These students, on average, wrote between 22.0 words (Cecilia) and 48.4 words (Nick) during the baseline phase.

During the study, I analyzed the story elements’ data to make decisions about baseline stability and phase changes. When baseline stability was demonstrated across students’ baseline story elements’ scores, the SRSD story-writing treatment was

introduced in a time-lagged fashion across each tier. I continued to document baseline stability in the untreated tiers to document that no change occurred prior to the introduction of the intervention. Across each writing outcome, there was at least one student in each tier whose baseline data remained stable, even when the intervention was introduced in a proceeding tier. In the next section, I discuss the consistency of findings during the treatment sessions for these students.

Treatment. Students showed improvement in the trend or slope of their data across all the writing outcome measures (number of essential story elements, holistic writing quality, and number of written words; see Table 11). The slope improvements from baseline to treatment phases were the largest and most consistent for the story elements' outcome measure. Every student increased their slope by approximately 1.0 point (range 0.5 to 1.2). Three students improved the trend of their treatment data on the holistic writing quality measure: Justin, Michelle, and Lucas. The average improvement on the holistic writing quality measure was 0.5 points with scores ranging from 0.2 (Lucas) to 0.9 (Michelle). Three students showed an increase in their treatment slopes for the number of written words: Cecilia, Nick, and Lucas. Nevertheless, the average increase of words was very small: 4.0 words with a range of 2.2 words (Nick) to 7.1 words (Cecilia). Visual analysis of the graphed data (see Figures 2 to 4) also supports these findings.

When did improvements begin to manifest during treatment? I anticipated that improvements would not occur until the strategy was modeled. Improvements for most students occurred after the tutors modeled *and* the students were provided an opportunity to write collaboratively with their tutor. For example, three students (Cecilia, Justin, and

Michelle) began to show improvements in number and quality of story elements on the fourth writing prompt. This prompt was given after the students had an opportunity to write using a graphic organizer without pictures (Lesson 5, see Appendix B). Two other students (Alexis and Lucas) showed improvement after the third writing prompt (given after students evaluated one of their earlier writing products, Lesson 4, see Appendix B) and Nick showed improvement on the fifth prompt (after he wrote without a graphic organizer, Lesson 6, see Appendix B).

For the three students whose mean level performance on the holistic writing quality improved during the treatment phase, improvement for Justin and Lucas occurred at the same time their story elements score increased (Prompts 3 and 4, respectively). Michelle did not begin to show improvement on the holistic quality score until prompt 7 (after she had reached the independent performance stage in the tutoring sessions).

The SRSD story-writing treatment also influenced variability of students' story elements scores. During baseline, five students' data were low and stable: Cecilia, Justin, Michelle, Nick, and Lucas. When the treatment was introduced, these students' data began to improve (as indicated by their positive slope changes) and continued to improve throughout the treatment. This steady improvement resulted in many students early treatment data points being 20% below their mean level and their ending data points exceeding 20% from their mean levels. Overall variability ranged from 14.3% (Cecilia) to 42.9% (Justin). Such changes in variability were not observed for the holistic writing quality or number of words measures.

In conclusion, my second research question was favorably answered. The nonprofessional adult-led SRSD story-writing instruction resulted in improved writing

for struggling third-grade writers during treatment. Implications of these findings are addressed further under suggestions for future research.

Post-treatment. There were three data points collected for five students during the post-treatment phase. Post-treatment data were not collected for Alexis due to an extended school absence. Post-treatment data were collected within four weeks of the treatment ending. All of the students with post-treatment data continued to show improvements during the post-treatment phase: Cecilia, Justin, Michelle, Nick, and Lucas. Their story elements scores went from being highly variable in the treatment condition (variability = 14.3% to 42.9%) to stable in the post-treatment condition (stability for all students = 100%). In addition, each of these students' mean level of performance improved during the post-treatment condition by a score of one (0.9 for Justin) to almost four (3.7 for Cecilia).

It should be noted that Justin was given a booster session after his second post-treatment probe as his score fell to baseline levels. Cecilia was also given a booster session before any of her post-treatment prompts as she had been absent from school for two weeks due to an illness. Even if these students were removed from the analysis, there would still be three demonstrations of improved post-treatment performance: Michelle = 1.3 more elements, Nick = 1.7 more elements, and Lucas = 2.8 more elements. Additionally, four students' slopes continued to improve during the post-treatment phase: Justin, Michelle, Nick, and Lucas. Taken together, these results indicate that students' data were not declining over the post-treatment phase.

The SRSD story-writing instruction also resulted in maintained improvements during the post-treatment phase on the holistic writing quality measure. Three students

made small to moderate improvements in their mean holistic quality scores from treatment to post-treatment phases: Justin (> 0.1 quality points), Michelle (>1.4 quality points), and Lucas (>0.7 quality points). There data were also stable (stability for all students = 100%). However, only one of these students' data had a positive slope (Justin), which indicated that Michelle and Lucas's quality scores showed a decreasing trend over this phase.

Maintenance. This discussion is based on three data points collected for each student available during the maintenance phase: Cecilia, Justin, Michelle, and Lucas. Data were not collected for Alexis and Lucas as the end of the school year occurred before such data could be gathered. Maintenance data were collected six weeks after the treatment ended. The SRSD story-writing instruction had a positive effect on maintaining or improving students' mean levels of performance on the story elements' measure during the maintenance phase. Three students' means improved from post-treatment condition: Cecilia (>0.7 story elements), Michelle (>1.3 story elements), and Nick (>2.2 story elements). All of these data continued to remain stable (stability for all students = 100%). Of the students who showed increased means, only Cecilia and Michelle continued to show a positive slope during the maintenance phase. Nick's maintenance slope for essential story elements was negative. Experimental control was not established for the quality measure during the maintenance phase, as there were only two demonstrations of maintained improvements: Justin and Michelle. Lucas had shown improvements in his writing quality during treatment and post-treatment, but the school year ended before his maintenance data could be collected.

In conclusion, the third research question that guided this study can be favorably answered. The objectives for the current study were met. The nonprofessional adult-led SRSD story-writing instruction resulted in improved writing output for struggling third-grade writers during post-treatment and maintenance phases in the following ways: essential story elements (post-treatment improvements: mean, slope, stability; maintenance improvements: mean); holistic writing quality (post-treatment improvements for mean performance). These results are meaningful as the improvements obtained with the six struggling third-grade writers in this study are similar to those obtained with other typical third-grade writers (see Tracey, Reid, and Graham, 2009). As in previous SRSD studies, the students in this brief study were not expected to elaborate on each of the seven essential story elements (and thus receive a 14 on the story elements scale). The growth that was observed in this study can be built on. This finding is discussed further in the recommendations for future research.

Positive effects of the SRSD story-writing treatment on number of written words were not observed at post-treatment or maintenance. It should be noted that increasing the total number of words written was not the primary intent of the current study.

These results suggest that most improvements were made on the writing outcome measure that tutors spent the most time focusing on during tutoring sessions: essential story elements. Treatment in the current study ended when students could successfully demonstrate the ability to plan and draft stories that included the seven essential story elements. Implications of these findings and how they direct future research and practice are addressed later in the chapter.

Social validity: Students. The students reported high levels of social validity for the nonprofessional adult-led story-writing instruction and recommended its continued use. The nonprofessional adults varied in their overall assessment, but each recommend its continued use. Students were unanimous in their like of learning POW (the trick for all writing) and provided examples of what they liked about their teacher teaching them the tricks. There was one negative comment about learning the WWW, What = 2, How = 2 trick. Lucas stated the WWW trick was too long and difficult to memorize. There was one negative comment related to an overall dislike of the SRSD writing treatment. Justin stated he did not like missing class. Students provided several examples of how to improve instruction from making each lesson shorter, to coming up with and using more games during sessions or having students generate their own mnemonics or writing tricks. When asked if they had used their writing tricks on other writing tasks, most students said yes. Students used the tricks when writing at home (Cecilia, Michelle, and Lucas) and at school (Nick and Alexis). Justin said he had not yet used the tricks, but that he still might. Finally, all students gave the Writing Academy a *thumbs up*. Alexis also said that the Writing Academy changed her life. Overall results were unanimous with students having equally high opinions of their tutors and the SRSD treatment.

Social validity: Nonprofessional adults. The nonprofessional adults maintained journals throughout the SRSD story-writing intervention. The overall effectiveness rating scores (1 to 4) across students ranged from 3.8 (Nick) to 2.2 (Cecilia). Jessica's overall ratings were the highest (M = 3.7). Phyllis's average effective rating was the next highest (M= 2.8). James average effective range was the lowest (M = 2.3). James' journal

comments reflected a tentative approval for the SRSD story-writing instruction, whereas Jessica and Phyllis's journal entries suggested a strong approval of the Writing Academy.

In general, the tutors' written comments accurately reflected their overall effectiveness ratings. There also seemed to be a connection between effectiveness ratings and student performance. The mean level of performance on number of essential story elements, holistic quality, and number of written words was consistently the highest for Jessica and Phyllis's tutees and lowest for James' tutees. James' tutees also took longer to progress to the independent performance stage during the SRSD lessons (17 lessons as compared to 11 and 12 for Janet and Phyllis's tutees).

When asked about their overall impression of the SRSD story-writing intervention and their willingness to teach again, each tutor answered favorably. All three said their overall experience had been positive and that they would happily teach again. The more subdued indications of overall effectiveness reported in the journals may have been a more accurate response, as it seemed less likely that these individuals would have expressed negative comments when speaking with me directly.

The use of adult journaling to document overall effectiveness and provide a space for tutors to reflect on each lesson seemed an effective way to capture honest responses, during each SRSD lesson. Journaling after each lesson allowed the tutors to recall the benefits and/or challenges for each lesson. The tutors used the journals after each lesson and provided information that was valuable in understanding the effectiveness of the SRSD story-writing treatment (see Figures 5, 6, and 7). For example, in the first four or five lessons, tutors gave slightly higher effectiveness ratings to students who were easier to work with. As lessons progressed, they gave higher scores to students who performed

better. Such information is very useful when conducting future training sessions, especially with nonprofessional adult volunteers. Journaling also appeared to have a positive impact on overall high procedural fidelity levels (see Table 8). The use of journaling as it relates to procedural fidelity is highlighted further in the recommendations for future research section.

Extension of Previous Research

SRSD story-writing literature. The current study extends what is known about the effectiveness of SRSD story-writing instruction, when that instruction is delivered by someone other than a researcher, experienced teacher (Rogers & Graham, 2008), or paraeducator with previous tutoring experience (Reid et al., 2009). Specifically, it demonstrated that a nonprofessional adult-led supplemental SRSD story-writing intervention produced positive effects for young struggling writers. Students in the current study improved during treatment (essential story elements, holistic writing quality, and number of written words) and maintained those improvements from four to six weeks after instruction ended (essential story elements). The results from this study also highlight the time-intensive nature of the SRSD story-writing treatment and provide several directions for future research. These suggestions for future research are discussed in the next section.

Supplemental nonprofessional adult-led literacy interventions. The current study also extended the supplemental nonprofessional adult-led literacy intervention literature.

Writing intervention. First, the current study was the first to evaluate the effects of a nonprofessional adult-led *writing* treatment on multiple student outcomes. Previous reviews have primarily involved reading interventions. Of course, these findings need to be replicated and additional strategies need to be tested with nonprofessional adult-led instruction.

Maintenance data. The current study also extended the supplemental nonprofessional adult-led literacy intervention research, by including maintenance data. In my review of this literature, only 32 percent (6/19) of the studies reviewed examined effects several weeks after the treatment had ended (Nielson, 1991; Meyer et al., 2002; Ramey, 1990, 1991; Vadasy et al., 1997a, 2000). In the majority of these cases (4/6; 67%), effects diminished at maintenance. The current study provided evidence for maintained effects of this supplemental writing treatment.

Program structure. The current study also extended the supplemental nonprofessional adult-led literacy intervention research by providing adequate details related to program structure. Wasik et al. (1993, 1998) suggested that programs led by knowledgeable coordinators that included individually paced lessons, active student participation, explicit modeling, and scaffolded instruction were superior to those supplemental programs that did not contain those attributes. The current study provides an example of such a program that is described in enough detail so that it can be replicated. This is not the case in most studies (see Chapter 2).

Treatment fidelity and social validity. The current study included treatment fidelity and social validity data. In my review of nonprofessional adult-led literacy interventions, less than 30 percent of the studies reported treatment fidelity. Additionally,

in the current study I gathered social validity data from students (at the end of the study) and nonprofessional adult tutors (throughout the study via journal entries). The treatment fidelity and social validity results of the current study demonstrate that the described SRSD treatment was delivered as described and was viewed favorably by participants.

Study Limitations

The results of this study must be interpreted in light of several study limitations. The limitations are related to volunteer identification, pre-baseline data, treatment procedures, and post-treatment and maintenance data. Those limitations are discussed next.

Volunteer identification. My intent was to select and train SRSD story-writing tutors from the school's community of volunteers. I was unable to identify more than one volunteer at the school using principal recommendations and the distribution of a flyer soliciting participation. With principal permission, I sought out individuals from a community 30 miles away. This is a study limitation because it may be difficult to replicate similar procedures in future studies. However, the problems that arose with tutor identification in this study may also arise in other studies or real-life situations. The results from this study would suggest that if similar identification problems arise (cannot locate volunteers from within a school), individuals should be sought from other communities.

Identifying tutors from a different community may also have biased the social validity results, although this is unlikely. I knew the tutors who were recruited from the other community (James and Jessica). Our previous associations may have affected their

effectiveness rating for the SRSD story-writing intervention. This seems doubtful, at least in part, as James gave the lowest marks and both James and Jessica appeared very honest in their journal responses.

Pre-baseline data. The students for this study were selected based on pre-treatment writing skills (performance on TOWL-3 and teacher recommendation). One of the third-grade classroom teachers was absent due to an unexpected medical leave when students were identified. The other third-grade classroom teachers provided valuable input, but the students (Alexis and Nick) may not have been best suited for the study. For example, Alexis wrote one sentence on her formal writing assessment, but wrote an average of 44.5 words per prompt (range 29 – 76) during the baseline phase. Additionally, Alexis was the only student to include all of the seven story elements, and elaborated on three of these elements, during baseline (see Table 13). A pre-baseline conference with Alexis's teacher may have revealed that she was a better writer than the formal writing assessment score implied. Alexis baseline data were superior to other students' baseline data (mean and median scores, frequency including elements in writing, and elaboration), however, her *improvement* was similar to that observed among the other students with lower baseline scores (see Tables 11 and 13). Additionally, Alexis expressed much satisfaction with the SRSD story writing instruction, stating, , "The Writing Academy changed my life".

Treatment procedures. I observed tutors every three sessions. During James' first observation, I noticed his lack of organization. His students' folders were overflowing with papers and this seemed to decrease James' success. For example, he had to say, "Wait a minute while I find that …" on several occasions. I talked with James

about this and suggested I give him the folders with only the information that was needed for each lesson. James immediately agreed. I continued to organize student folders for James from this point forward. I implemented identical procedures for Jessica and Phyllis from the first day of their instruction. This is a possible study limitation for two reasons.

One limitation is related to the degree to which the nonprofessional was fully responsible for each lesson. In preparing the folders, I took on a much more active role. This active role increased the amount of supervision required for the current study. The folder preparation required approximately 10 minutes per each student's lesson. Over the course of the study that extra time amounted to about 8 extra hours. The time requirements used in this controlled study may be difficult to sustain in a typical school situation. The second limitation relates to inconsistent procedures used across tutors. James' tutees did have the lowest means across essential story elements and holistic quality measures. James' initial experience with disorganization may have influenced these outcomes.

Conversely, these procedures may have contributed to the high levels of procedural fidelity observed in the current study. Nonprofessional adult tutors may need this level of support (having folders prepared for them) initially or throughout the study. This finding is discussed further under recommendations for future research.

Incomplete post-treatment and maintenance data. Another limitation resulted when I was unable to collect post-treatment data for Alexis and maintenance data for Alexis and Lucas. Alexis's absences during post-treatment and maintenance phases were unexpected and were not related to the SRSD story-writing intervention. Lucas's maintenance data could have been collected had the study begun earlier in the school

year. Collecting post-treatment and maintenance data for all students would have strengthened the overall findings.

Suggestions for Future Research and Practice

The results of this study provide a basis to continue the implementation of nonprofessional adult-led structured writing interventions with young struggling writers in school settings. I highlight three areas for future research based on the findings from the current study: (a) extending SRSD instruction to include other writing genres, (b) evaluating how much training and support are needed to replicate positive outcomes, and (c) examining the effectiveness of extending the SRSD story-writing instruction.

Other writing genres. First, the current study examined the effects of SRSD *story-writing* instruction on multiple outcome measures. Students' performance in other genres such as persuasive writing and personal narratives should also be examined.

Procedural fidelity. Second, more research is needed to see how much pre-treatment training and tutor support (direct support and tutor compensation) are needed to replicate the positive outcomes observed in the current study. These are important questions to ask as they address the feasibility of replications. The current study was a controlled / time-intensive treatment. I provided high levels of support from pre-treatment (training) through treatment (see Table 10). The amount of support provided in the current study (see Table 10) supported the high levels of fidelity, but may also make it difficult to replicate procedures (e.g., some potential applicants in this study declined participation due to time commitment).

Pre-treatment tutor training. One of the supports provided to the tutors in the current study was tutor training (6 to 8 hours). The amount of training was based on the nonprofessional adults' lack of previous training and procedures used in previous SRSD studies (Harris et al., 2003). The importance of providing adequate training has been well documented (Erlbaum et al., 2000; Harris et al., 2003; Wasik, 1998). Researchers who have examined the nonprofessional adult-led supplemental reading intervention literature have suggested that structured programs that require informed judgment on the part of the tutor should include adequate tutor training (Wasik, 1998). Wasik defined adequate training as having an understanding of the basic reading principles. Erlbaum et al. reported that studies that reported tutor training were associated with larger overall effects.

The SRSD story-writing treatment used in the current study was structured, requiring tutors to make informed judgments related to lesson progression (e.g., Have students developed the adequate background knowledge?). More research is needed to assess *how much* and *what type* of pre-treatment training is needed to replicate the positive results found in this study. In the current study, training was criterion-based versus time-based. Tutors were also paid a small stipend for participating in the training sessions. Future replications should continue to provide adequate, criterion-based training. More research is needed to assess whether total time needed to reach criterion varies based on tutor training experience and whether rates of nonprofessional adult volunteer participation are effected by varying levels (or omission of) tutor compensation.

Another recommendation for future research relates to *what* is covered in tutor training. An observation made in working with the tutors in the current study was tutors' expectations that students should become master writers during this relatively short writing treatment. Future pre-treatment training procedures should include showing tutors the *anticipated* rate of student progress. In other words, tutors should be shown previous student data and informed that progress does not usually occur until SRSD lessons 3 to 5. In other words, tutors should be informed that in previous research (i.e., the current study), young struggling writers did not demonstrate improved writing output (number of story elements) until the following SRSD stages had been covered: Develop Background Knowledge, Discuss It, Model It, Memorize It, and Support It. Along these lines, more research is needed to determine whether similar rates of progress will be observed.

Direct support provided during treatment. In the current study, tutors were also provided with much support during treatment. More research is needed to assess how much tutor support (direct support and compensation) is needed to replicate the positive outcomes observed in this study. In the current study, I provided direct support to the tutors during treatment in several ways (see Table 10). I provided transportation, prepared student folders, delivered and removed tape recorders and audio tapes, provided general feedback, observed weekly, provided specific feedback, and requested tutors to journal after each tutoring session.

Although future replication of these procedures (research or practice), may be less controlled and require less intensive support from a knowledgeable supervisor, certain elements should remain. For example, I believe that preparing student folders for tutors greatly improved the procedural fidelity of the treatment and resulted in high rates of

tutor attendance. I also believe the frequent observations and discussions related to student progress, and tutor journaling had a positive impact on the high procedural fidelity levels and positive writing outcomes observed in the current study. These procedures should remain in future replications.

Additionally, other procedures should be added based on observations made in the current study. For example, tutors should be shown student progress (graphed data) throughout the treatment phase. Tutor journal entries suggested an apprehension from some tutors regarding the rate at which students were progressing. Presenting these data throughout the treatment phase would allow tutors to visualize progress made.

Finally, other procedures should be examined to assess whether their inclusion is necessary to replicate the positive outcomes observed in this study. Previous reviews of nonprofessional adult supplemental reading programs have suggested that tutor compensation may be important to a program's overall success (Meyer, 2008; Wasik et al., 2002). The high levels of procedural fidelity may not have been as greatly impacted by tutor compensation as originally predicted. Participation rates may have been positively influenced by tutor compensation ($50 for attending training and $100 per student tutored), but other types of support (e.g., frequent observations and feedback) appeared to have more of an impact on the high procedural fidelity levels observed (see Table 10). Future research should examine the impact of the size or type of compensation on procedural fidelity levels.

Extending SRSD story-writing instruction. Third, more research is needed to assess the effectiveness of extending the SRSD story-writing instruction to elaboration of story elements missed entirely or elaborated on infrequently. In the current study, the

objectives were met. Every student wrote stories that included more essential story elements and were of higher quality as a direct result of the SRSD story-writing instruction. The growth observed in the current study can be built on. Future studies should examine the effectiveness of extending the SRSD story-writing instruction to focus on elaboration of essential story elements frequently missed or omitted altogether.

Conclusion

In conclusion, the results from this study suggest that nonprofessional adults can be used to effectively deliver supplemental structured writing intervention. This study demonstrated that the SRSD story-writing intervention delivered by nonprofessional adult volunteers had positive effects on six struggling writers' writing during treatment (improved slopes for essential story elements, holistic writing quality, and number of written words) and after the treatment ended (maintained higher means for number of essential story elements six weeks after instruction ended). Caution should be used when interpreting these findings as the tutors in the current study were provided with much support. More research is needed to see how much tutor support is needed to replicate these positive outcomes.

APPENDIX A

Teacher Survey of Classroom Writing Practices (see Lane et al., 2008)

Section 1

The following 36 questions will use the *Never* to *Several Times a Day* scale displayed below.

Never	**Several Times a Year**	**Monthly**	**Several Times a Month**	**Weekly**	**Several Times a Week**	**Daily**	**Several Times a Day**

1. Circle how often **you conference** with students about their writing.
2. Circle how often **students conference with their peers** about their writing.
3. Circle how often your students engage in **"planning and drafting"** before writing.
4. Circle how often your students **"revise"** their writing products.
5. Circle how often **students share their writing** with their **peers**.
6. Circle how often your students **"publish"** their writing. (Publish means to print or write it so that it can be shared with others.)
7. Circle how often your **students help their classmates** with their writing
8. Circle how often **you read your own writing** to your students.
9. Circle how often you teach **sentence construction skills**
10. Circle how often you teach students about ways of **organizing text** or how texts are organized.
11. Circle how often you teach students **strategies for planning and drafting**.
12. Circle how often you teach students **strategies for revising**.

13. Circle how often you teach students **handwriting skills**.

14. Circle how often you teach **spelling skills**.

15. Circle how often you teach **grammar skills.**

16. Circle how often you teach **punctuation skills.**

17. Circle how often you teach **capitalization skills**.

18. Circle how often you **provide mini-lessons** on writing skills or processes students need to know at this moment---skills, vocabulary, concepts, strategies, or other things.

19. Circle how often you overtly model **writing strategies**.

20. Circle how often you **model the enjoyment or love of writing** for students.

21. Circle how often you **re-teach** writing skills or strategies that you previously taught.

22. Circle how often you assign **writing homework** to students in your class.

23. Circle how often your students work at **writing centers**.

24. Circle how often you use a **writing prompt** (e.g., story starter, picture, physical object, etc.) to encourage student writing.

25. Circle how often **you monitor the writing progress** of your students in order to make decisions about writing instruction.

26. Circle how often you encourage **students to monitor their own writing progress**.

27. Circle how often students use **rubrics** to evaluate their writing.

28. Circle how often students in your classroom use **writing portfolios** (add material to a portfolio, look at material already in it, and so forth).

29. Circle how often you ask students to **write at home with parental help**.

30. Circle how often you ask **parents to listen** to something their child wrote at school.

31. Circle how often you **communicate with parents** about their child's writing progress.

32. Circle how often you allow one or more students in your classroom to write by **dictating** their compositions to someone else.

33. Circle how often you allow one or more students in your classroom to use **computers** during the writing period.

34. Circle how often students use **writing to support reading** (e.g., write about something they read).

35. Circle how often students use **reading to support writing** (e.g., read to inform their writing).

36. Circle how often your students use **writing in other content areas** such as social studies, science, and math.

Section 2: The next 5 questions will use the following scale:

0	1	2	3	4	5	6	7
Never			**Half The Time**				**Always**

1. Circle how often **students select their own writing topics**.
2. Circle how often students are allowed to **complete writing assignments at their own pace**.
3. Circle how often you encourage students to use "**invented spellings"** at any point during the writing process.
4. Circle how often your writing lessons have **multiple instructional goals**.
5. Circle how often your students use a **graphic organizer** (e.g., story map) when writing.

Section 2

Teachers were asked to write a **brief description of their writing program**.

Teachers will also be asked to check with of the following best describes his/her approach to **writing instruction**:

_____ Traditional skills approach combined with process writing

_____ Process writing approach

_____Traditional skills approach

APPENDIX B

Writing Academy: POW + WWW with Transfer
Lesson 1a

Instructor: _____ Date:_____
Student: _____

Purpose: Develop Background Knowledge, Discuss It

Objectives: Introduction to POW, story parts, and story parts reminder. Identification of story parts in story examples.

Materials needed: Mnemonic charts and story examples (Albert the fish), WWW graphic organizer WITH PICTURES, paper, pencils, scratch paper, student folder

____ I. Introduce Yourself
Introduce yourself as a writing academy teacher. Tell the student you're going to teach him/her some of the "tricks" for writing. First, we're going to learn a strategy, or trick, that good writers use **for everything** they write. Then we are going to learn the trick, or strategy, for writing good stories.

____II. Introduce the audio tape player.
Tell the student that the following, "I am going to use this tape player in order to see how well I am teaching. Before we start, I am going to test it and make sure that it is working. Can you help me? When I turn it on can you tell me today's date and read this number or would you like me to do it instead?"

_____III. Check the audio tape player.
Begin by saying your name into the tape player , the student's identification number, and the date. Play it back and make sure the tape recorder is working. If the student would also like to say these things, please allow him/her to do so. Replay the tape. Stop it. Push record to begin recording this lesson.

____ IV. Introduce POW

A. Put out the POW + WWW chart so that only POW shows.

B. Emphasize: POW is a trick good writers often use, for many things they write.

C. Go over parts of POW, discussing each. (P = Pick an idea to start with – this is an idea in our head; O = Organize my notes – I will teach you a trick for organizing your notes later; W = Write and say more– we will use our notes to write, and we can think of more ideas while we write). Describe and discuss the concept of notes. Use examples; "Your teacher uses notes when she creates a web on the board; your parents use notes when they write things on a calendar or a grocery list." Have the student generate some examples on their own. Emphasize that a good way to remember POW is to remember that it **gives them POWer for everything they write.**

D. Practice POW; Turn the chart over. Practice reviewing what each letter in POW stands for and why it is important (good writers use it often, for many things they write). Help as needed. Have the student write out POW on scratch paper and explain out loud what each letter stands for. Repeat until the student knows what POW stands for and why it is important.

____ V. Discuss Good Stories

Discuss good stories briefly – ask the student what makes a story good? Be sure to include:

A. Good stories are **fun for me to write and fun for others to read.**

B. Good stories **make sense and have several parts** – we will learn a trick for remembering the parts of a good story. This trick is the trick we will use to help us organize our notes.

____ VI. Introduce WWW

Introduce WWW – uncover more of the chart so that the WWW shows. "Let's find out what the parts of a good story are." Have the student view the chart. Briefly discuss each W. Use the word "character" for **Who**; for **When**, ask the student to tell you "how does a person tell you when in a story?" –Once upon a time….A long time ago…..Yesterday…..Wednesday afternoon at 4:00…..One night……and so on. Have the student generate examples. Next go over **Where**. Give examples such as Clarksvile, at school, in Iraq…have the student give examples.

____ VII. Find WWW in a story (Albert)

A. Say, "Now we are going to read a story to find out if the writer used **WWW** in the story." (Leave out the partially covered story parts reminder sheet where the student can see it.) Quickly review what the WWW stands for.

B. Give each student a copy of the story (Albert). Ask the student to follow along silently while you read the story out loud. Tell them to be listening for the who, when, and where in the story. Read the story a second time and ask the student to say when they hear a story part. Remind them that they might not hear the parts in that order. As the student identifies the parts, who, when, and where; write each part on the appropriate space on the graphic organizer. **Do not use complete sentences – do this in note form!** Be sure that the student knows you are writing in note form. Be explicit.

____ VIII. Introduce What = 2

A. Uncover each What=2. Explain briefly and discuss each what. Give examples of how a writer might tell each. (Use a story the student would know ~ 3 little pigs ~ what did the wolf want? What happens in the story?)

____ IX. Introduce How = 2

B. Uncover How=2. Explain briefly and discuss each how. Give examples of how a writer might tell each. (How does the story of the 3 little pigs end? How do the characters feel throughout the story ~ when the wolf knocks at the door?)

____ X. Find What=2 and How=2 in a story (Albert)

C. Tell the student that they are now looking for 2 whats and 2 hows. Briefly review what each means and reread the story. Stop and have the student name the parts. Write each part in note form on the graphic organizer. Point out that we might put more than one note in each part. A good story may have more than 2 "whats". Also, good writers tell how the characters feel in different parts of the story. If the student has not identified all the parts, go back over the story and help as needed. Be encouraging and positive throughout.

____ XI. Introduce and Find Million Dollar Words (MDW) in a story (Albert)

Explain briefly and discuss million dollar words (MDW). Million dollars words are good vocabulary words that help describe something. They make the story or sentence more interesting. Discuss some examples (i.e. frigid instead of cold). Tell student "Let's see if the writer included any million dollar words in our story about Albert". Work with student to find MDWs.

____ XII. Practice Story Parts Reminder

Turn over the WWW chart and the student's papers. Have the student practice telling you the 7 parts to a good story. Have the student write the reminder, WWW What=2 How=2 on scratch paper. Repeat several times till the student gets comfortable. If you have extra time, use POW cards for extra practice.

____XIII. Lesson Wrap Up

A. Announce test! (No grade-for fun!) next session. They will come and write out POW and the story parts reminder and tell what they mean from memory.

B. Give each student their own folder and a copy of the story parts reminder chart. Have student put today's work and their charts in their folder and give the folder back to you ~ explain you will bring the folder to every class.

Albert The Fish

On a warm, sunny day two years ago, there was a huge gray fish named Albert. He lived in a big, icy pond near the edge of town. Albert was swimming around the pond when he spotted a big, juicy worm on top of the water. Albert knew how wonderful worms tasted and wanted to eat this one for dinner. So he swam very close to the worm and bit into him. Suddenly, Albert was pulled through the water into a boat. He had been caught by a fisherman. Albert felt sad and wished he had been more careful.

Albert The Fish

On a warm, sunny day two years ago **(When)**, there was a huge gray fish named Albert **(Who)**. He lived in a big icy pond near the edge of town **(Where)**. Albert was swimming around the pond when he spotted a big, juicy worm on top of the water. Albert knew how wonderful worms tasted and wanted to eat this one for dinner **(What He Wanted To Do)**. So he swam very close to the worm and bit into him. Suddenly, Albert was pulled through the water into a boat **(What Happened)**. He had been caught by a fisherman **(Ending)**. Albert felt sad **(Feelings)** and wished he had been more careful.

Possible MDW: icy, huge, juicy, and wonderful

POW + W-W-W What = 2 How = 2

WHO

WHEN

WHERE

_____ _____ _____

_____ _____ _____

_____ _____ _____

Insert picture that is meaningful to student (based on discussion)

WHAT WANTS

1.
2.
3.

WHAT HAPPENS

Insert picture that is meaningful to student (based on discussion)

HOW FEELS

LAST PAGE

HOW END

_____ _____ _____ _____

_____ _____ _____ _____

_____ _____ _____ _____

Writing Academy: POW + WWW with Transfer
Lesson 1b

Instructor: _____ Date:_____
Student(s): _____

Purpose: Develop Background Knowledge, Discuss It

Objectives: Review and practice POW, story parts, and story part reminder; identification of story parts in story examples; establish concept of **transfer**

Materials Needed: Story example (The Lion and the Mouse), WWW graphic organizer WITH PICTURES (see earlier lesson), practice cards, "I transferred my strategies" chart, paper, pencils, scratch paper, student folder

____ I. Introduce tape player

____ II. Check tape player

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder.**

A. Give the student a piece of scratch paper. Ask the student to **write down POW** – then ask student what each letter stands for, and why it is important for writing stories. If student has trouble remembering POW, practice it using rapid fire with the cue cards.

Rapid Fire Practice

Give the student a set of cue cards (for WWW, start practice with cue cards with picture cues then wean the student to cards without picture cues). Say, "To help you remember the parts, we are going to do an exercise called **rapid fire**. We will take turns saying the parts. This is called rapid fire because you are trying to name the parts as rapidly as you can. If you need to look at the cue card, you may; however, don't rely on the card too much because I am going to put the card away after several rounds of rapid fire." Allow the student to paraphrase but be sure intended meaning is maintained. Do with cue cards and without. If response is correct, make brief positive comment. If incorrect, prompt by pointing to cue card.

B. Remind the student that **O** needs a trick for organizing. Ask the student what the trick is for organizing my notes for stories. Ask student to **write out the story parts reminder mnemonic/trick** on the scratch paper. The student should write: W-W-W; What = 2; How = 2. If the student has trouble, be supportive and prompt as needed.

C. Now **ask the student what each part of the story part reminder stands for**.

D. It is essential that the student memorize the reminder. If the student is having trouble with this, spend a few minutes practicing it using rapid fire with the cue cards.

E. Tell the student you will test him/her on it each day to make sure he/she has it. Remind the student that he/she can practice memorizing it.

____ IV. Find Parts in a Second Story (The Lion and the Mouse).

Leave out chart. As before, remind the student to raise his/her hand when they hear a part. Be sure each part is identified. As the student identifies who, when, and where; you write each in the appropriate space on the graphic organizer: **do not use full sentences – do this in note form.** Be sure that the student understands that you are writing in **note form!**

____ V. Introduce Transfer

Tell the student: **"We have a goal for our POW and WWW strategies."**

A. The first goal, **Goal 1** for the next time: **use all or parts of POW and/or WWW in other writing tasks**. Brainstorm together some classes or other writing tasks he/she could use both POW and WWW for, being sure to note that we should use POW with WWW whenever we use WWW. Other ideas could be: book reports, letters to friends, reports on special topics, writing for a school newsletter, writing about something that happened to you or a special event, and so on. Briefly note that for some tasks, like writing a report, all parts of the WWW trick might not be right to use – so what could we do? (Change WWW to fit the kind of report we need to write; **don't use all of WWW if it doesn't make sense** (i.e., don't use WWW to write directions for how to do something like make your lunch or bake a cake).

B. Tell him/her that it were their responsibility, once they learn to use the strategy, to come to every Writing Academy lesson with at least two examples of how they were able to transfer the writing strategies (POW/WWW) to other writing tasks (for example, the student might report **making notes** for a writing task before he/she wrote, this would count). Show him/her the "I transferred my strategies" chart and explain that **once a week** you will write down each time he/she tells you about using all or any part of POW/WWW outside of this class. Briefly discuss the word **"transfer"** – transfer means to move (like I transferred schools means that I moved from one school to another). Emphasize that you want him/her to **transfer** what they learn about POW and WWW from this class to other classes and other writing tasks.

____ VI. Lesson Wrap Up

A. Announce test! (no grade!) next session. He/she will come and write out POW and the story parts reminder and tell what they mean from memory.

B. **Remind the student to transfer the strategy, that you will ask him/her next time if he/she transferred, and that you were recording on their chart later in the week.

C. Give the student his/her folder, a copy of the story parts reminder chart, and a copy of the "I transferred my strategies" chart. **Have them put today's work and his/her chart in his /her folder and give the folder back to you – explain that you will bring the folder to every class.**

The Lion and the Mouse

One sunny day a long time ago a big strong lion was taking a walk in the forest near his home. He walked into a huge net. “Help!” he yelled. “I can’t get out. I am scared.” A cute little mouse came running along. She cried, “I’ll help you!” “Oh!” said the lion. “How could you help? You’re too little.” The mouse said, “I can too help! You’ll see.” And the mouse began biting the net into tiny bits. The lion was able to get out of the net. When the lion got out he grinned. He said, “You may be a little mouse. But you’re a big help.” The mouse felt proud that she had helped the lion.

The Lion and the Mouse

One sunny day a long time ago **(When)** a big strong lion **(Who)** was taking a walk in the forest near his home **(Where)**. He walked into a huge net. "Help!" he yelled. "I can't get out. I am scared **(Feelings)**." A cute little mouse **(Who)** came running along. She cried, "I'll help you!" **(What He Wanted to Do)** "Oh!" said the lion. "How could you help? You're too little." The mouse said, "I can too help! You'll see." And the mouse began biting the net into tiny bits **(What Happened)**. The lion was able to get out of the net **(Ending)**. When the lion got out he grinned. He said, "You may be a little mouse. But you're a big help." The mouse felt proud **(Feelings)** that she had helped the lion.

Possible MDW: huge, cute, tiny, grinned, and proud

I Transferred My Strategy!

Name: _____ Date:_____

Check all of the ways that you used the tricks (POW or WWW, What = 2, How =2) in your writing:

_____ ***Book Report*** _____ ***School Reports*** _____ ***School Newsletter***

_____ ***Letter*** _____***Personal Journal*** _____***School Journal***

Other examples:

_____

_____

_____

Did all of the WWW trick and/or POW trick work? If not, how did you change?

_____

_____

_____

Writing Academy: POW + WWW with Transfer
Lesson 2

Instructor: _____ Date:_____
Student(s): _____

Purpose: Review POW + Story Parts Reminder; Model; Record Self-Instructions

Objectives: Review POW and story parts reminder; model **self-statements**; have student establish personal self-statements; **introduce rockets**

Materials Needed: Mnemonic Chart, practice cards, 2-WWW graphic organizers- WITH PICTURES, scratch paper, pencils, lined paper, student folder, story (*Farmer's Story*), practice picture (*turtle*), self-statement sheets, one blank graph, student folder

____ I. Introduce tape player.

____ II. Check tape player.

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder** by writing the mnemonic out on a piece of scratch paper. Spend some time practicing the parts out loud. Use the rapid fire cards to play a game. Tell the student you will test them on it each day to make sure he/she has it. Be sure the student remembers that the **story parts reminder** is the trick for **O.**

____ IV. Find Parts in a Story

Practice finding parts of a story (Farmer's Story) and taking notes on the graphic organizer. Point out to the student how and why you are taking notes. Give the student opportunities to orally state the parts in note form.

____ V. Model Using Self-Statements for "P" in POW

Have a copy of your self-statement sheet available. Use problem definition, planning and drafting, self-evaluation, **self-reinforcement**, and coping statements as you work. Use statements that are similar to those employed by the student. Ask the student to help you with ideas, but be sure you are in charge of the process. Say:

"Remember that the first letter in POW is P – pick my idea. Today we are going to practice how to think of a good story idea and come up with good story parts. To do this we have to let our minds be free and creative."

A. Look at the practice picture: turtle. Model things you might say to yourself when you want to think of a good idea. For example, "Take my time and a good idea will come to me." "What ideas can I see in this picture?" You can also start with a negative statement and model how a coping statement can help you get back on track. For example, "I can't think of anything to write! Ok, if I just take my time, a good idea will come to me." Explain to the student that things you say to yourself out loud and in your head help you get through the writing process. I might think in my head, "**What is it I have to do? I have to write a good story. A good story makes sense and has all 7 parts."**

B. Ask the student to come up with things he/she might say in his/her head to help him/her think of good story ideas and good parts. If the student is having trouble, help him/her create a statement or let him/her "borrow" one of yours until he/she come up with his/her own. Have student record 1-2 things they can say to help think of good ideas on their self-statement sheet.

____ VI. Discuss Using "O" in POW

Tell the student the second letter in POW is O –ORGANIZE my notes. Explain that you are going to write a story today with his/her help. I need a trick for O. The trick is my story part reminder WWW What = 2 How = 2. Put out your graphic organizer and your story reminder sheet. Briefly review the 7 parts to a good story and point out their places on the graphic organizer. Review, what your goals should be – Write a good story, with all 7 parts, that makes sense, is fun to read, and fun to write.

Now I can do O in POW – Organize my Notes. I can **write down story part ideas** for each part. I can write ideas down in different parts of this page as I think of ideas (be sure to **model moving out of order during your planning and drafting)**. **What ideas do I see in this picture?** (Now – talk out and fill in notes for **who, when where)**. For "who" I see…For "when" I can write…Let's see, for "where" – it's …Good! I like these parts! Now I better figure out the **2 whats and 2 hows. Let my mind be free, think of new, fun ideas.** (Now talk out and briefly write notes for the 2 whats and 2 hows – not in full sentences - **use coping statements at least twice.**) **. For example, "I have so many ideas for where the story could be taking place, but I don't know which one to pick. Alright, I know what I'll do. I'll put one idea down and if I get through with all of my other parts I'll add more there later. Ooh, I feel good about that. Now I can move on and write a good story with all seven parts.**) Let's see, for the story question of "what does the main character want to do "I think…For the next "what" question, "what happens when she tries to do it" I think…**I can add more action by** writing about…For the "ending" I can say…For the "feeling" story part I can write about…(After generating notes for all the story parts say – Now I can **look back at my notes and see if I can add more notes for my story parts – actually do this** – model it – use coping statements). I can also **look for ideas for good word choice or million dollar words – do this.**

____VII. Model Writing a Story Using POW and WWW.

A. Keep the POW and story parts graphic out; also the student's self statement sheet

B. Model the entire process: writing an actual story as you go (using the practice picture and your graphic organizer). (Please print so student can easily follow)

Now I can do W in POW – write and say more. I can **write my story and think of more ideas or million dollar words as I write.** Now – talk yourself through writing the story; the student can help. Use a clean piece of paper and print. **Start by saying "How shall I start? I need to tell who, when, and where."** Then pause and think, then write out sentences. Do be sure to **add 1-2 more ideas and million dollar words** on your plan as you write. For example, you may say, "Wait, I just thought of a MDW that would make that sentence so much more fun to read!" Don't hurry, but don't slow it down unnaturally. Also, at least 2 times, **ask yourself, "Am I using good parts and, am I using all my parts so far?"** As you write and include ideas from your plan, model checking yourself as you write by checking off the story parts that you have used. This is also a good opportunity to use encouraging and positive self-statements. **Be sure to use**

coping statements. Also ask yourself, "Does my story make sense?" When story is done, say **"Good work, I'm done. It'll be fun to share my story with others."**

____VIII. Self-Statements for Story Writing

Add to student's self-statements' lists. Ask the student if they can remember: 1) The things you said to yourself to get started? 2) The things you said while you worked (try to get some creativity statements, coping statements, statements about remembering the parts, and self-evaluation statements) 3) The things you said to yourself when you finished. (Tell him/her if he/she can't remember and discuss each as you go). Make sure each student adds these to his/her list:

- what to say to **get started.** This must be along same lines as **"What is it I have to do? I have to write a good story with good parts, and with all 7 parts."** – but in student's own words.

- 1-2 things to say **while you work**: self-evaluation, coping, self reinforcement, and any others he/she likes (in student's own words).

- Note that we don't always have to think these things out loud; once we learn them we can think in our heads or whisper to ourselves.

____IX. Introduce Graphing Sheet/Graph the Story

Ask student: does this story have all 7 parts? Find each part and fill in graph. Use stars; circle or fill in a star around this rocket for each million dollar word used.

____X. Lesson wrap-up

A. Keep your story and graph.

B. Remind of POW and story parts reminder **test** again next time.

C. ** If appropriate remind student to transfer the strategy, that you will ask them next time if they transferred, and if so he/she will fill in the transfer chart.

The Farmer's Story

Many years ago there was an old farmer who lived near the woods. He owned a stubborn donkey. The farmer wanted to put his donkey in the barn. First he pushed him, but the donkey would not move. Next, the farmer tried to frighten the donkey into the barn. So he asked his dog to bark at the donkey, but the lazy dog refused. Then the farmer thought that his cat could get the dog to bark. So he asked the cat to scratch the dog. The dog began to bark angrily. The barking frightened the donkey and he jumped into the barn. The farmer was very proud of himself.

The Farmer's Story

Many years ago **(When)** there was an old farmer **(Who)** who lived near the woods **(Where)**. He owned a stubborn donkey. The farmer wanted to put his donkey in the barn **(What He Wanted To Do)**. First he pushed him, but the donkey would not move. Next, the farmer tried to frighten the donkey into the barn. So he asked his dog to bark at the donkey, but the lazy dog refused. Then the farmer thought that his cat could get the dog to bark. So he asked the cat to scratch the dog. The dog began to bark angrily **(What Happened)**. The barking frightened the donkey **(Feelings)** and he jumped into the barn **(Ending)**. The farmer was very proud of himself **(Feelings)**.

Possible MDW: stubborn, frighten, lazy, and angrily

My Self-statement Sheet

Name:_____ Date:_____

What can I say to myself to help me get **STARTED** with my writing?

1. _____

2. _____

3. _____

4. _____

What can I say to myself **WHILE** I am writing to keep me motivated?

1. _____

2. _____

3. _____

4. _____

What can I say to myself **AFTER** I finished writing?

1. _____

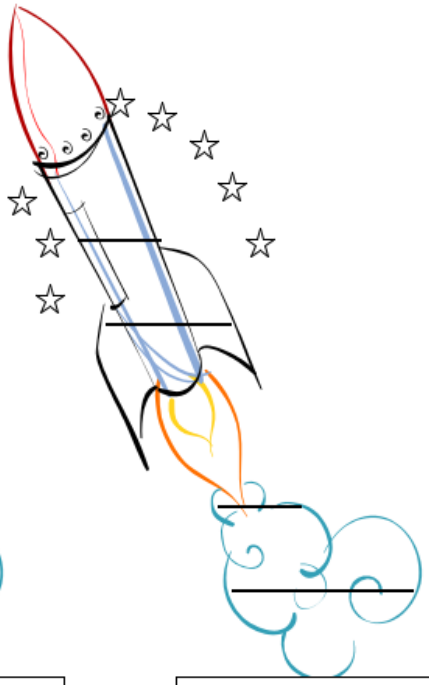
2. _____

3. _____

4. _____

The Writing Academy 2009-2010
Rocket Graphing Sheet

Date:
Did I blast off my rocket? _____
Did my story make sense? _____
Did I use million dollar words? _____



Date:
Did I blast off my rocket? _____
Did my story make sense? _____
Did I use million dollar words? _____

Writing Academy: POW + WWW with Transfer
Lesson 3

Instructor: _____ **Date: _____**
Student(s): _____

Purpose: Review POW & Story Parts Reminder, Self-Instructions, Collaborative Writing

Lesson Overview: The student and teacher will collaboratively write a story using **POW + WWW, What=2, How=2**. The teacher will need to provide the support needed to insure that student is successful in writing a story that has all 7 parts. The teacher should reinforce the student's use of self-instructions, good word choice, a story that makes sense, and "million dollar" words.

Objectives: Review and practice POW, story parts, and story part reminder; identification of story parts in story examples; reinforce transfer and write collaboratively

Materials Needed: Mnemonic charts and story example (Smokey), WWW graphic organizers-WITH PICTURES, Transfer Sheet, Self-Statements Sheet, Rocket Graphing Sheet, story prompt (boy on alligator), paper, pencils, scratch paper, student folder

____ I. Introduce tape player

____ II. Check tape player

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder.**

A. Test to see if the student remembers **POW + WWW, What=2, How=2**. Do it out loud to save time. It is essential that the student memorize these. If student has trouble, practice using rapid fire cue cards. Tell the student you will test them on it each day to make sure they have it.

____ IV. Find Parts in a Story (if needed); practice MDWs

Practice finding parts of a story (Smokey) and taking notes on the graphic organizer. Point out to the student how and why you are taking notes. Give the student opportunities to orally state the parts in note form. **Have the student change some of the words in Smokey to MDWs.**

____ V. Transfer

Review the meaning of transfer briefly. Ask the student to orally report back one time he/she used or could have used all or parts of POW and/or WWW in other classes or for other kinds of writing tasks. If necessary, brainstorm together again some classes or other writing tasks they could use both POW and WWW for, being sure to note that we should use POW with WWW whenever we use WWW. Other writing tasks could be: book reports, letters to friends, reports on special topics, writing for a school newsletter, writing about something that happened to you or a special event, and so on. Briefly remind the student that for some tasks, liking writing a report, all parts of the WWW trick might not be right to use – so what could we do? (Change WWW to fit the kind of report we need to write).

____ VI. Collaborative Writing

Give student a blank graphic organizer and ask them to take out their self-statements list. Put out the boy on the alligator practice picture. This time let the student lead as much as possible, but prompt and help as much as needed. It should be a collaborative process.

1. Say, "Remember that the first letter in **POW** is **P** - PICK my IDEA." Refer student to his self-statements for creativity or thinking free. Help the student get an idea.

2. Say, "The second letter in **POW** is **O** - ORGANIZE my NOTES. Remind the student to use the story parts reminder to help - the "trick" for organizing notes when writing a story. Encourage them to say, "I will use this page to make my notes and organize my notes." Review – "What should our goal be?" "We want to write a good story - a good story has all seven parts, makes sense, is fun for me to write and for others to read." After you have both generated notes for all the story parts (have student write as much as possible), say – "Remember to look back at our notes and see if we can add more detail or description" - help them actually do this. Make sure all the parts are filled in on the notes sheet. Identify at least 2 things the student did really well.

3. The last letter in **POW** is **W** - WRITE and SAY MORE. Encourage and remind the student to start by saying "What is it I have to do here? I have to write a good story - a good story has all 7 parts and makes sense. I can write my story and think of more good ideas or million dollar words as I write." Help student as much as they need to do this, but try to let them do as much as they can alone. Encourage them to use other self-statements of their choice while they write. If student do not finish writing today, they can continue at the next lesson.

____ VII. Graph Story Parts

Begin a Rocket Graphing Sheet for the student. Have the student shade in the graph to equal the number of story parts they included – have the student determine- does the story have all 7 parts - then fill in graph. Reinforce the student for reaching 7. Tell the student, "You blasted your rocket!" Use the stars; fill in or circle a star around that rocket for each million-dollar word used.

____ VIII. Lesson Wrap-Up

A. Have the student put her/his work and charts in the folder.

B. Remind the student that he/she will fill in the transfer chart again next time.

C. Remind student of the POW + WWW, What=2, How=2 test again next time.

Smokey

Smokey was an old gray horse. Lisa used to ride Smokey, but now Smokey stays in his field on the farm. He was happy. One hot summer day Lisa came to see Smokey. She brought him red apples. Smokey liked the red apples. Lisa liked to run through the meadow and fields. Lisa thought Smokey would like to run so she opened the gate. But Smokey didn't go out because he didn't want to run. Lisa said, "You don't have to run with me. You stay here and I will give you an apple everyday." And she gave him an apple everyday from that day on. Both Lisa and Smokey were happy.

Smokey

Smokey was an old gray horse **(Who)**. Lisa **(Who)** used to ride Smokey, but now Smokey stays in his field on the farm **(Where)**. He was happy **(Feeling)**. One hot summer day **(When)** Lisa came to see Smokey. She brought him red apples. Smokey liked the red apples. Lisa liked to run through the meadow and fields. Lisa thought Smokey would like to run so she opened the gate **(What she wanted to do)**. But Smokey didn't go out because he didn't want to run **(What happened next)**. Lisa said, "You don't have to run with me. You stay here and I will give you an apple everyday." And she gave him an apple everyday from that day on **(Ending)**. Both Lisa and Smokey were happy **(Feeling)**.

Writing Academy: POW + WWW with Transfer
Lesson 4

Instructor: _____ Date: _____
Student(s): _____

Purpose: Review POW & Story Parts Reminder, Compare Prior Performance to Current Writing Behavior

Objectives: Review and practice POW, story parts, and story part reminder; reinforce transfer, discuss pretest story and compare to current writing

Materials Needed: Mnemonic charts, WWW graphic organizer – NO PICTURES IF READY, Transfer Sheet, Self-Instructions Sheet, Rocket Graphing Sheet, pretest story, collaborative story, pencil, scratch paper, student folder.

____ I. Introduce tape player

____ II. Check tape player

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder.**

A. Test to see if the student remembers **POW + WWW, What=2, How=2**. Do it out loud to save time. It is essential that the student memorize these. If student has trouble, practice using rapid fire cue cards. Tell the student you will test them on it each day to make sure they have it.

____ IV. Transfer

Review the meaning of transfer briefly. Ask the student to orally report back one time he/she used or could have used all or parts of POW and/or WWW in other classes or for other kinds of writing tasks. If necessary, brainstorm together again some classes or other writing tasks they could use both POW and WWW for, being sure to note that we should use POW with WWW whenever we use WWW. Other writing tasks could be: book reports, letters to friends, reports on special topics, writing for a school newsletter, writing about something that happened to you or a special event, and so on. Briefly remind the student that for some tasks, liking writing a report, all parts of the WWW trick might not be right to use – so what could we do? (Change WWW to fit the kind of report we need to write).

____ V. Establish Prior Performance

Say, “Remember the stories you wrote before we learned POW and WWW?” Pull out a story the student wrote during pre-testing/baseline.

Have the student read his/her story and identify which parts he/she has. (You need to have worked out ahead of time what parts the student had and which ones the student didn't have.)

Briefly note with the student which parts he/she has and which he/she doesn't. Emphasize with the student that he/she wrote this story before learning the "tricks" for writing. Now that he/she knows the "tricks" his/her writing has already greatly improved. Compare the pretest story to the collaborative story and talk about what the student has learned about good writing. If the student is exhibiting frustration or is upset about his/her pretest story, encourage him/her to use a self-statement.

Spend some time talking about how to improve the pretest story and if the student would like, and time allows, give him the opportunity to redo the story or to do a graphic organizer for the story, now that he/she knows the "tricks" for writing a good story. Help the student make a commitment to use the strategies (tricks) to write better stories.

Set a goal to continue writing better stories. Remind them that good stories: are fun to write and for others to read, have all 7 parts, that each part is well done, and that good stories make sense.

Say, "Our goal is to have all of the parts and 'better' parts the next time we write a story."

____ VI. Lesson Wrap-Up

****If this lesson goes fast and you have time, use an extra picture and do a graphic organizer****

C. Have the student put his/her work and charts in his/her folder.

D. Remind the student that they will fill in the transfer chart again next time.

C. Remind student of the POW + WWW, What=2, How=2 test again next time.

Writing Academy: POW + WWW with Transfer
Lesson 5

Instructor: _____ Date: _____
Student: _____

Purpose: Review POW & Story Parts Reminder, Collaborative Practice; Review Self-Instructions

Objectives: Review and practice POW, story parts, and story part reminder; reinforce transfer, individual collaborative practice

Materials Needed: Mnemonic charts, WWW graphic organizers – NO PICTURES, Transfer Sheet, Self-Instructions Sheet, Rocket Graphing Sheet, story picture prompt (boat and sea monster), pencil, paper, student folder.

____ I. Introduce tape player

____ II. Check tape player

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder.**

A. Test to see if the student remembers **POW + WWW, What=2, How=2**. Do it out loud to save time. It is essential that the student memorize these. If student has trouble, practice using rapid fire cue cards. Tell the student you will test them on it each day to make sure they have it.

____ IV. Transfer

Review the meaning of transfer briefly. Ask the student to orally report back one time he/she used or could have used all or parts of POW and/or WWW in other classes or for other kinds of writing tasks. If necessary, brainstorm together again some classes or other writing tasks they could use both POW and WWW for, being sure to note that we should use POW with WWW whenever we use WWW. Other writing tasks could be: book reports, letters to friends, reports on special topics, writing for a school newsletter, writing about something that happened to you or a special event, and so on. Briefly remind the student that for some tasks, liking writing a report, all parts of the WWW trick might not be right to use – so what could we do? (Change WWW to fit the kind of report we need to write).

____ V. Individual Collaborative Writing

Give student a blank graphic organizer and ask them to take out their self-statements list. Put out the picture prompt (boat and sea monster). This time let the student lead as much as possible, but prompt and help as much as needed.

1. Say, "Remember that the first letter in **POW** is **P** - PICK my IDEA." Refer student to their self-statements for creativity or thinking free. Help the student get an idea.

2. Say, "The second letter in **POW** is **O** - ORGANIZE my NOTES. Remind the student to use the story parts reminder "trick" to help. Encourage them to say, "I will use this page to make my notes and organize my notes." Review – "What should your goal be?" "You want to write a good

story - a good story has all seven parts, makes sense, is fun for you to write and for others to read." After the student has generated notes for all the story parts, say – "Remember to look back at your notes and see if you can add more detail or description" - help them actually do this. Make sure all the parts are filled in on the notes sheet. Identify at least 2 things the student did really well.

3. The last letter in **POW** is **W** - WRITE and SAY MORE. Encourage and remind the student to start by saying "What is it I have to do here? I have to write a good story - a good story has all 7 parts and makes sense. I can write my story and think of more good ideas or million dollar words as I write." Help student as much as they need to do this, but try to let them do as much as they can alone. Encourage them to use other self-statements of their choice while they write. If student do not finish writing today, they can continue at the next lesson.

____VI. Graph Story Parts

Have the student shade in the graph to equal the number of story parts they included – have the student determine- does the story have all 7 parts - then fill in graph. If the student misses a part, talk about how to revise the story and set a goal for next time.

____ VII. Lesson Wrap-Up

E. Have the student put their work and charts in their folder.

F. Remind the student that they will fill in the transfer chart again next time.

G. Remind student of the POW + WWW, What=2, How=2 test again next time.

*****Repeat this lesson if the student appears to have difficulty with any of the story parts, with taking notes on the graphic organizer, or is having difficulty transferring notes to the actual story writing.**

IN STORY? Graphic Organizer Without Pictures

_____**W**HEN _____

_____**W**HO _____

_____**W**HERE _____

_____**WHAT** WANT _____

_____**WHAT** HAPPENED(1,2,3) _____

_____

_____

_____**HOW** FEEL _____

_____**HOW** END _____

Writing Academy: POW + WWW with Transfer
Lesson 6

Instructor: _____ Date: _____
Student(s): _____

Purpose: Review POW & Story Parts Reminder, Wean off Graphic Organizer

Objectives: Review and practice POW, story parts, and story part reminder; reinforce transfer, individual collaborative practice; wean off graphic organizer

Materials Needed: Transfer Sheet, Self-Instructions Sheet, Rocket Graphing Sheet, story picture prompts (alien; boy with door), pencil, scratch paper, student folder.

____ I. Introduce tape player

____ II. Check tape player

____ III. Test POW and WWW, What = 2, How =2

Test to see if the student remembers **POW** and the **story parts reminder.**

A. Test to see if the student remembers **POW + WWW, What=2, How=2**. Do it out loud to save time. It is essential that the student memorize these. If student has trouble, practice using rapid fire cue cards. Tell the student you will test them on it each day to make sure they have it.

____ IV. Transfer

Review the meaning of transfer briefly. Ask the student to orally report back one time they used or could have used all or parts of POW and/or WWW in other classes or for other kinds of writing tasks. If necessary, brainstorm together again some classes or other writing tasks they could use both POW and WWW for, being sure to note that we should use POW with WWW whenever we use WWW. Other writing tasks could be: book reports, letters to friends, reports on special topics, writing for a school newsletter, writing about something that happened to you or a special event, and so on. Briefly remind the student that for some tasks, liking writing a report, all parts of the WWW trick might not be right to use – so what could we do? (Change WWW to fit the kind of report we need to write).

____ V. Wean off Graphic Organizer

Explain to the student that they won't usually have a story parts reminder page with them when they have to write stories, so they can make their own notes on blank paper. Show them how to write down the reminder at the top of the page: WWW What =2 How =2. Have them make a space for each story part on their notes page.

____ VI. Individual Collaborative Writing

Give student a blank piece of paper and ask them to take out their self-statements list. Put out the picture prompt (aliens). This time let the student lead as much as possible, but prompt and help as much as needed. This time the student will make notes on blank paper ~ no graphic organizer! Go through the following processes but let the student do as much as possible with prompting.

1. Say, “Remember that the first letter in **POW** is **P** - PICK my IDEA.” Refer student to their self-statements for creativity or thinking free. Help the student get an idea.

2. Say, “The second letter in **POW** is **O** - ORGANIZE my NOTES. Remind the student to use the story parts reminder “trick” to help. Encourage them to say, “I will use this page to make my notes and organize my notes.” Review – “What should your goal be?” “You want to write a good story - a good story has all seven parts, makes sense, is fun for you to write and for others to read.” After the student has generated notes for all the story parts, say – “Remember to look back at your notes and see if you can add more detail or description” - help them actually do this. Make sure all the parts are filled in on the notes sheet. Identify at least 2 things the student did really well.

3. The last letter in **POW** is **W** - WRITE and SAY MORE. Encourage and remind the student to start by saying “What is it I have to do here? I have to write a good story - a good story has all 7 parts and makes sense. I can write my story and think of more good ideas or million dollar words as I write.” Help student as much as they need to do this, but try to let them do as much as they can alone. Encourage them to use other self-statements of their choice while they write. If the student does not finish writing today, they can continue at the next lesson.

____VII. Graph Story Parts

Have the student shade in the graph to equal the number of story parts they included – have the student determine- does the story have all 7 parts - then fill in graph. If the student misses a part, talk about how to revise the story and set a goal for next time.

____ VIII. Lesson Wrap-Up

E. Have the student put their work and charts in their folder.

F. Remind the student that they will fill in the transfer chart again next time.

G. Remind student of the POW + WWW, What=2, How=2 test again next time.

H. Tell students you have done a great job, next time we will take a practice test.

*****Repeat this lesson until student can write a story independently. Select from remaining pictures.**

Writing Academy: POW + WWW Practice Test
Lesson 7

Instructor: _____ Date: _____
Student(s): _____

Purpose: Post Testing Practice and Preparation

Objectives: Review and practice POW, story parts, and story part reminder; write independently; practice post-testing conditions

Materials Needed: story picture prompt (select from remaining pictures), pencil, scratch paper, lined paper, Rocket Graphing Sheet, student folder.

____ I. Introduce tape player

____ II. Check tape player

____ III. Introduce Practice Test

Tell your student that now you will practice taking the test for writing a story so that when we do it again, it were much easier. Give the student a picture and two blank pieces of paper, one for notes and one for story writing.

____ IV. Practice Test

Tell them, ok, now lets pretend it is a test day. What do you do first? **THEY MUST WRITE OUT THE W-W-W, WHAT=2, HOW=2 ON ONE PIECE OF BLANK PAPER - PROMPT THEM TO DO SO IF THEY ARE UNSURE, HELP ONLY AS NEEDED.** Once this is written out, say, "Good", this is what you need to do first every time we do a test for story writing. If student wants to write out POW, explain that he/she does not need to do this, they can just remember POW in their head - when they make notes for the story and then writes the story, they are doing POW!!

Wait and see if the student goes on. If not, ask the student what they need to do next. Prompt and help only as necessary - what they need to do is make notes for each part. When they are done, remind them they can think of more ideas as they write, if they want to. Prompt for out loud self-statements only when you think they are needed. At this point, it is ok if they aren't using much or any out loud speech.

Wait and see if the student goes on. If not, ask the student what they need to do next. Prompt the student to write the story as needed, letting them do it on their own as much as possible. Same on out loud statements, prompt only if needed.

Wait and see if the student goes on. If not, ask the student what they need to do next. At this point, they should read their story, see if they have all the parts, be sure it makes sense, and see if there are any changes they would like to make. You can remind them to also see if they can use any million dollar words if this seems appropriate and not too much for them.

____ V. Graph Story Parts

Now, go over the story with the student, counting the parts, and go ahead and graph this story on their rockets. Compliment them on good work!

____ VI. Lesson Wrap-Up

Tell student “You have done a great job learning the W-W-W strategy, and now you can write stories by remembering the mnemonic, organizing your notes on blank paper, and writing a story that is fun to write and fun for others to read and makes sense. The next time I ask you to write a story for me, I won't be able to help you. This were our test to see if you remember what you have learned. I will ask you to write about three more stories for me. I will make copies of your story that I can keep, and then I will give you back all of your stories and work, your rockets, and a certificate that shows you have learned the trick for writing a good story. Thank you so much for doing such great work

APPENDIX C

Procedural Fidelity Form for Lesson 1A

STUDENT/LEG: ☐ Female / Leg 1 ☐ Male / Leg 1
☐ Female / Leg 2 ☐ Male / Leg 2
☐ Female / Leg 3 ☐ Male / Leg 3

Lesson (1a, 1b, 2, 3, 4, 5, 6, 7):_____

Date of Lesson: _____

Writing Academy Instructor (James, Jessica, Phyllis): _____

Person conducting Fidelity: _____

Tape / School Observation: _____

First day conducting lesson? _____

Total time in instruction: _____

Treatment Fidelity

_____ Tutor makes introductions

_____ Introduces POW (trick for making writing more powerful)

_____ Discusses good stories (what is a good story? why important? what's in them?)

_____ Introduce WWW (flashcards)

_____ Find WWW in the Albert Story

_____ Introduce What = 2 (flashcards)

_____ Introduce How = 2 (flashcards)

_____ Find What = 2 and How = 2 in the Albert Story

_____ Introduce million dollar words and find one in Albert Story

_____ Practice memorizing WWW, What = 2, and How = 2 (flashcards)

_____ Lesson Wrap-Up ("test" and clean up)

Teaching Quality – Rating Scale

DATE: _____

LESSON: _____

STUDENT: _____

1. Was XX well prepared for SRSD instruction when I arrived? _____
2. Was XX positive and enthusiastic with the student? _____
3. Did XX seem to have rapport with the student? _____
4. Was XX able to redirect off-task behavior so that they did not interfere with instruction? _____
5. Did XX carefully monitor and support the student, actively involving them in instruction? _____
6. Was the lesson appropriate paced? _____
7. Did XX appear motivated and enthusiastic about SRSD instruction? _____
8. Did XX end the lesson well and set goals for the next lesson? _____

Observations:

_____

_____

_____

_____

_____

_____

_____

_____

REFERENCES

*References marked with an asterisk indicate studies in which effect sizes were calculated and presented in Chapter Two.

Adkins, M. H. (2005*). Self-regulated strategy development and generalization instruction: Effects on story writing among second and third grade students with emotional and behavioral disorders*. Unpublished doctoral dissertation, University of Maryland, College Park, MD.

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht: Kluwer.

Allen, A., & Chavkin, N. F. (2004). New evidence that tutoring with community volunteers can help middle school students improve their academic achievement. *The School Community Journal, 14*, 7-18.

*Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality, 13*, 195-208.

*Baker, S., Gersting, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly, 35*, 494-519.

Baker, S. K., Ketterlin-Geller, L. R., Chard, D. J., Apichatabutra, C., & Doabler, C. (2009). Teaching writing to at-risk students: The quality of evidence of self-regulated strategy development. *Exceptional Children, 75*, 303-318.

Bell, K. M. (2001). *Partnerships in literacy tutoring: Using community volunteers to provide one-to-one tutoring to struggling readers in third, fourth, and fifth grade.* Unpublished doctoral dissertation. University of Pittsburgh.

Bereiter, C, & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

*Burns, M. K., Senesac, B. V., & Symington, T. (2004). The effectiveness of the HOSTS program in improving the reading achievement of children at-risk for reading failure. *Reading Research and Instruction, 43,* 87-104.

Cohen, P. A., Kulik, J. A., & Kulik, C-L, C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19,* 237-248.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100,* 907-919.

Danoff, B., Harris, K. R., Graham, S. (1993). Incorporating strategy instruction within the writing process in the regular classroom: Effects on the writing of students with and without learning disabilities. *Journal of Reading Behavior, 25,* 295-322.

Edl, H. M. (2007). *Examining the impact of video self-modeling on the reading fluency of upper elementary and middle school students with significant reading disabilities*. Unpublished doctoral dissertation. Indiana University.

Erion, J. (2006). Parent Tutoring: A meta-analysis. *Education and Treatment of Children, 29,* 79-106.

Erlbaum, B., Vaughn, S., Hughes, M. T., Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of intervention research. *Journal of Educational Psychology, 92*, 605-619.

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). Educational Research: An introduction (8th Ed.). Boston: Allyn and Bacon.

Gersten, R., Fuchs, L., Compton, D., Coyne, M., Greenwood, C., Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research. *Exceptional Children, 71,* 149-164.

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology, 98*, 732-749.

Good, R. H., & Kaminski, R. A. (Eds.). (2007). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement. Accessed online at https://dibels.uoregon.edu/measures/files/2_benchmark_6th_ed.pdf on July 10, 2009.

Graham, S. (2006a). Strategy instruction and the teaching of writing: A meta-analysis. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 187-207). New York: Guilford.

Graham, S. (2006b). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology* (pp. 457-478). Mahwah, NJ: Erlbaum.

Graham, S., & Harris, K. R. (1989). A component analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology, 81,* 353-361.

Graham, S., Harris, K. R., & Mason, L. H. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*, 207-241.

Graham, S., & Perin, D. (2007a). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445-476.

Graham, S., & Perin, D. (2007b). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading, 11*, 313-335.

Hammill, D., & Larsen, S. (1996). *Test of written language (TOWL-3).* Circle Pines, MN: AGS.

Harris, K. R., Graham, S., & Mason, L. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*, 295-340.

Harris, K. R., Graham, S., & Mason, L. (2003). Self-regulated strategy development in the classroom: Part of a balanced approach to writing instruction for students with disabilities. *Focus on Exceptional Children, 35*, 1-16.

Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Erlbaum.

Hillocks, G. (1986). *Research on written composition: new directions for teaching.* Urbana, IL: National Council of Teachers of English.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. *Exceptional Children, 71,* 149-164.

Juel, C. (1996). What makes literacy tutoring effective. *Reading Research Quarterly, 31,* 268-289.

Kennedy, C. H. (2005). Single-case designs for educational research. Boston: Allyn & Bacon.

Kiuhara, S. A., (2009). The effects of self-regulated strategy development training on the persuasive essay writing of high school students with disabilities. Unpublished doctoral dissertation. University of Utah.

Lane, K. L., Harris, K. R., Graham, S., Weisenbach, J. L., Brindle, M., & Morphy, P. (2008). The effects of self-regulated strategy development on the writing performance of second-grade students with behavioral and writing difficulties. *Journal of Special Education, 41,* 234-253.

Lane, K. L., Wolery, M., Reichow, B., & Rogers, L. (2007). Describing baseline conditions: Suggestions for study reports. *Journal of Behavioral Education, 16,* 224-234.

Lienemann, T. O., Graham, S., Leader-Janssen, B., & Reid, R. (2006). Improving the writing performance of struggling writers in second grade. *The Journal of Special Education, 40,* 66-78.

MacArthur, C., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing, and dictation. *Journal of Special Education, 21*, 22-42.

McCutchen, D. (1988). "Functional automaticity" in children's writing: A problem of meta-cognitive control. *Written Communication, 5*, 306-324.

*McGrady, D. S. (1983). *The effect of programmed tutoring upon the reading comprehension of fourth-grade students enrolled in a chapter 1 reading program*. Unpublished doctoral dissertation. Purdue University, IN.

McLukin, D. L. (2006). *An analysis of the expectations, training, and perceived support of the volunteers at an after-school community-based reading tutorial program*. Unpublished doctoral dissertation. The University of Michigan.

*McPeak, S. H. (2000). *The effects of trained adult volunteers in Vygotskian collaboration with elementary students on cognitive processing and reading achievement*. Unpublished doctoral dissertation. Seattle Pacific University, Seattle, WA.

Meier, J. D., & Invernizzi, M. (2001). Book buddies in the Bronx: Testing a model for American Reads. *Journal of Education for Students Placed at Risk, 6*, 319-333.

*Meyer, B. J. F., Middlemiss, W., Theodorou, E., Brezinski, K. L., McDougall, J., & Barlett, B. J. (2002). Effects of structure strategy instruction delivered to fifth-grade children using the internet with and without the aid of older adult tutors. *Journal of Educational Psychology*, 94, 486-519.

*Morris, D., Shaw, B., & Perney, J. (1990). Helping low readers in grades 2 and 3: An after-school volunteer tutoring program. *The Elementary School Journal, 91,* 133-150.

National Assessment of Educational Progress (NAEP, 2009). *Chronology of National Assessment of Educational Progress Assessments from 1969 to 2007*. Retrieved July 15, 2009, from http://nces.ed.gov/nationsreportcard/about/assesshistory.asp

National Commission on Writing (2003, April). *The neglected R: The need for a writing revolution*. Retrieved August 13, 2008, from www.collegeboard.com

National Commission on Writing (2004, September). *Writing: A ticket to work… or a ticket out: A survey of business leaders*. Retrieved August 1, 2008 from www.collegeboard.com

Neuman, S. B., & Roskos, K. (1993). Access to print for children of poverty: Differential effects of adult mediation and literacy-enriched play settings on environmental and functional print tasks. *American Educational Research Journal, 30,* 95-122.

*Nielson, B. B. (1991). *Effects of parent and volunteer tutoring on reading achievement of third grade at-risk students*. Unpublished doctoral dissertation. Brigham Young University, Provo, UT.

Olinghouse, N. G. (2008). Student- and instruction-level predictors of narrative writing in third-grade students. *Reading and Writing, 21*, 3-26.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220-244.

Persky, H.R., Daane, M.C. & Jin, Y. (2003). *The nation's report card: Writing 2002*. (NCES 2003–529). U.S. Department of Education.

*Ramey, M. (1991). *Compensatory education sustained gains from Spring 1988 to Spring 1990*. Report No. 91-1. Eric document 338774.

*Ramey, M. (1990). *Compensatory education sustained gains from Spring 1985 to Spring 1987*. Report No. 90-4. Eric document 330-764.

Reid, R., Luschen, K., & Lienemann, T. O. (2009). *Utilizing paraeducators to help improve story writing of struggling second-grade writers.* Unpublished manuscript, Department of Special Education, University of Nebraska, Lincoln, Nebraska.

*Rimm-Kaufman, S. E., Kagan, J., & Byers, H. (1999). The effectiveness of adult volunteer tutoring on reading among "at risk" first grade children. *Reading Research & Instruction, 38,* 143-152.

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79,* 3-38.

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100,* 879-906.

Saddler, B., Moran, S., Graham, S., & Harris, K. R. (2004). Preventing writing difficulties: The effects of planning and drafting strategy instruction on the writing performance of struggling writers. *Exceptionality, 12,* 13-17.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The Nation's Report Card: Writing 2007* (NCES 2008–468). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

*Sankaranrayanan, A. (1998). *Cognitive profiles of young children: A comparison of good and poor readers*. Unpublished doctoral dissertation. Harvard University, Cambridge, MA.

Scardamalia, M., & Bereiter, C. (1986). *Written composition*. In M. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 778-803), New York: MacMillan.

Schneider, M. (April 2008). *National assessment of educational progress. The Nation's report card: Writing 2007*. Retrieved on August 15, 2008 from http://nces.ed.gov/whatsnew/commissioner/remarks2008/4_3_2008.asp.

Scruggs, T. E., Mastropieri, M. A. (2001). How to summarize single-participant research: Ideas and applications. *Exceptionality, 9,* 227-244.

Serve America Act (2009). Senate bill S.3487.) Accessed online on June 2, 2009 at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_public_laws&docid=f:publ013.111.pdf

Smetana, L. (2005). Collaborative storybook reading: Bringing parents and at-risk kindergarten students together. *Reading Horizons, 45,* 283-320.

Tennessee Department of Education (2008). Report Card 2008. Web site: http:state.tn.us/education/reportcard/

Topping, K. J., & Hill, S. V. (1995). University and college students as tutors for schoolchildren: A typology and review of evaluation research. In S. Goodlad (Ed.), *Students as tutors and mentors* (pp. 13-31). London, PA: Kogan Page.

Tracy, B., Reid, R., & Graham, S. (2009). Teaching young students strategies for planning and drafting stories: The impact of self-regulated strategy development. *Journal of Educational Research, 102,* 323-329.

*Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S., K., & O'Connor, R. E. (1997a). Community-based early reading intervention for at-risk first graders. *Learning Disabilities Research and Practice, 12*, 29-39.

*Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S., K., & O'Connor, R. E. (1997b). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. *Learning Disability Quarterly, 20*, 126-139.

*Vadasy, P. F., Jenkins, J. R., & Pool, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. *Journal of Learning Disabilities, 33*, 579-590.

*Vadasy, P. F., Sanders, E. A., Peyton, J. A., & Jenkins, J. R. (2002). Timing and intensity of tutoring: A closer look at the conditions for effective early literacy tutoring. *Learning Disabilities Research and Practice, 17*, 227-241.

*Wallach, M. A., & Wallach, L. (1976). *Teaching all children to read*. The University of Chicago Press.

Wasik, B. A. (1998). Volunteer tutoring programs in reading: A review. *Reading Research Quarterly, 33*, 266-292.

Wasik, B. A., Bond, M. A., & Hindman, A. (2002). *After-school tutoring and children at risk.* In R. Garner, &. Zhao, and M. Gillingham (Eds.) Hanging Out. Bergin & Garvey, Westport, CT.

Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 179-200.

*Weiss, J. A., Thurlow, M. L., Christenson, S. L., & Ysseldyke, J. E. (1989, March). *Paired reading with adult volunteer tutors as a reading intervention for students with reading difficulties*. Paper presented at American Educational Research Association, San Francisco. Eric document #305-606.

White House, Office of the Press Secretary. (1997). President Clinton's America Reads Work-Study Challenge. Washington, DC: Author.

*Wilfong, L. G. (2006). *Combining the power of poetry, repeated readings, and community volunteers for literacy intervention: The poetry academy*. Unpublished doctoral dissertation. Kent State University, Kent, OH.

Woo, D. G. (2005). America Reads: The effects of a federal work-study tutoring program on literacy achievement and attitudes of teachers, tutors, and children. Unpublished doctoral dissertation. Rutgers University, New Brunswick, NJ.