The Effects of Instruction in Peer-Revision on the Persuasive Writing of Students with Emotional and Behavioral Disabilities

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Sara J. Mills
Master of Teaching
University of Virginia, 1999
Bachelor of Science
Northwestern University, 1995

Chairperson: Margo A. Mastropieri, University Professor
College of Education and Human Development

Fall Semester 2012
George Mason University
Fairfax, VA
This research was supported in part by grants: Grant No. R324A0070199-07 from the U.S. Department of Education, Institute of Educational Sciences, and from the Office of Special Education Programs, U.S. Department of Education Grant No. H325D070008. This paper does not necessarily reflect the position or policies of that funding agency, and no official endorsement should thereby be inferred.
DEDICATION

To my family, who fills my life with love and makes all things possible –
my husband Richard,
my children Emily and Daniel,
my parents Burt and Carol, and
my sisters Mary and Kristin.

Thank you.
ACKNOWLEDGEMENTS

This dissertation is the culmination of many years of work, which would not have been possible without the support of many people. First, I would like to thank my chairperson, Dr. Margo Mastropieri, for her tireless efforts to teach me the ins and outs of educational research and for providing me priceless experience as a member of her research team. Her persistence and thoughtful feedback kept me moving forward through all phases of my doctoral study. Thank you to Dr. Tom Scruggs and Dr. Kelley Regan, who have enriched my doctoral experience with their knowledge, expertise, and support. I truly appreciate the time and dedication the members of my dissertation committee have so generously given to help me conduct this research.

Additionally, thank you Margo and Tom for funding my studies through the Doctoral Leadership Grant in Special Education. Thank you to the Graduate School of Education at George Mason University, which provided me with fellowships and dissertation funding. These financial supports allowed me to pursue my studies.

Thank you to the teachers, students, and administrators who participated in this research study. You welcomed me with open arms and supported me with your time, flexibility, and follow through. I enjoyed working with you and am most thankful for all you did to make this study a success. Thank you to my colleagues Stephanie, Nancy, Mary, Mary, Catherine, Pat, and Yojanna, who coded behavior, watched videos, and scored dependent measures. I could not have done this without you! Thank you to my fabulous editor, Theresa, for her intelligence, humor, and patience.

It has truly taken a village for me to become Dr. Mills, and to my village I am deeply grateful. The unconditional support I have received from family and friends over the last five years has been humbling. Thank you to my parents, Burt and Carol, and my in-laws, Ernest and Elizabeth, who offered me encouragement at every turn, and who have travelled across the country countless times to help my family. Thank you to my sister Mary for talking with me about my research and helping me get to the finish line. Thank you to my sister Kristin for the late-night pep talks and valuable advice. Thank you to Donisha and Ms. Frances, who have lovingly cared for my children so that I could pursue
my studies. Thank you to the many neighbors and friends who have helped out in whatever way they were asked, whenever they were asked, especially Clair, Diana, Veronika, and Carrie. Thank you to my PhD buddies, Yojanna, Dani, Nancy, Stephanie, and Clara. Your friendship has meant the world to me. I could not have made it through this program without you by my side.

In closing, I would like to thank my family. You have been my rock. Thank you to my biggest cheerleaders, Emily and Daniel. Your encouraging words, hugs, love notes, and kind gestures kept me going, even when I wanted to give up. Thank you for always reminding me about what is most important. My biggest thank you goes to my husband, Rich. From the first time I said that I was thinking about enrolling in a PhD program, you supported me without reservation. You stepped up to the plate time and time again to accommodate my erratic schedule and the demands of my program. You consoled me during difficult times and shared in my joys. I am eternally grateful.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
</tr>
<tr>
<td>List of Figures</td>
</tr>
<tr>
<td>Abstract</td>
</tr>
</tbody>
</table>

1. Introduction ..................................................................................................................11
   Writing Instruction ........................................................................................................3
   Characteristics of Students with Emotional and Behavioral Disorders (EBD) ..................6
   Behavioral Characteristics of Students with EBD ..........................................................7
   Academic Characteristics of Students with EBD ...........................................................8
   Writing for Students with EBD .....................................................................................9
   Peer Learning for Students with EBD ............................................................................11
   The Present Study ........................................................................................................11
   Purpose and Research Questions ..................................................................................12
   Definition of Terms ......................................................................................................13
   Summary .......................................................................................................................21

2. Literature Review ........................................................................................................23
   Characteristics of Students with EBD ...........................................................................23
   Functional Characteristics of Students with EBD .........................................................27
   Academic Characteristics of Students with EBD ..........................................................28
   Academic Interventions for Students with EBD ............................................................33
   Peer-mediated Instruction for Students with EBD ..........................................................35
   Writing Research ...........................................................................................................39
   Meta-analyses of Writing Research ...............................................................................40
   Strategy Instruction .......................................................................................................48
   Revision Research ..........................................................................................................54
   Quality Single-Subject Research Design .........................................................................75
   Summary .......................................................................................................................76
3. Methods .................................................................................................................................79
   Setting .................................................................................................................................79
   Participants ..........................................................................................................................80
   Dependent Measures ..........................................................................................................90
   Materials .............................................................................................................................94
   Procedures ..........................................................................................................................101
   Fidelity of Treatment .........................................................................................................115
   Reliability and Scoring .......................................................................................................116
   Design ..................................................................................................................................121
   Summary ............................................................................................................................124

4. Results ..................................................................................................................................125
   Fidelity of Treatment .........................................................................................................125
   Data Analysis Methods ......................................................................................................126
   Persuasive essays ...............................................................................................................129
   Revisions .............................................................................................................................200
   Self-efficacy for persuasive writing scale .........................................................................205
   Writing fluency ....................................................................................................................205
   Peer-conferencing student performance ..........................................................................205
   Strategy knowledge questioning .........................................................................................210
   Social validity ......................................................................................................................213
   Summary ................................................................................................................................218

5. Discussion ............................................................................................................................221
   Research Questions ............................................................................................................221
   Study Quality .......................................................................................................................235
   Limitations ..........................................................................................................................237
   Implications for Practice .....................................................................................................239
   Directions for Future Research .........................................................................................240
   Summary ..............................................................................................................................241

Appendices ..............................................................................................................................243
   A. Sample Persuasive Essay Writing Prompts ...............................................................243
   B. Essay Scoring Conventions ...........................................................................................244
   C. Holistic Quality Scoring Rubric .....................................................................................246
   D. Persuasive Writing Self-efficacy Scale ..........................................................................248
   E. Woodcock Johnson III Writing Fluency Subtest Sample Questions .........................250
   F. On-task Behavior Coding Sheet ....................................................................................251
G. Social Validity Interview ................................................................. 252
H. POW+TREE Mnemonic ................................................................. 253
I. Graphic Organizer ........................................................................ 254
J. Self-statement Sheet ..................................................................... 255
K. Self-monitoring Graph ................................................................ 256
L. Transition Word List .................................................................... 257
M. Good Writing Handout ............................................................... 258
N. Revision Strategy Guide .............................................................. 259
O. Peer-revision Checklist ............................................................... 260
P. Conferencing Feedback Guide ..................................................... 261
Q. Strategy Knowledge Questionnaire ............................................. 262
R. Sample Fidelity of Treatment Checklist ........................................ 263
S. Peer-revision Steps Completed Checklist ..................................... 264

List of References ............................................................................. 265

Curriculum Vitae ................................................................................ 276
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Student characteristics</td>
<td>84</td>
</tr>
<tr>
<td>Table 2. Peer-conferencing steps</td>
<td>112</td>
</tr>
<tr>
<td>Table 3. Fidelity of treatment percentages by group</td>
<td>126</td>
</tr>
<tr>
<td>Table 4. Mean essay scores for all groups</td>
<td>131</td>
</tr>
<tr>
<td>Table 5. Mean essay scores by student</td>
<td>132</td>
</tr>
<tr>
<td>Table 6. Percent of nonoverlapping data across phases by student</td>
<td>149</td>
</tr>
<tr>
<td>Table 7. Mean number of revisions by type</td>
<td>201</td>
</tr>
<tr>
<td>Table 8. Percent of on-task behavior by type of instruction</td>
<td>208</td>
</tr>
<tr>
<td>Table 9. Difference scores across SRSD studies of adolescents with EBD</td>
<td>224</td>
</tr>
<tr>
<td>Table 10. Mechanical errors difference scores across revision studies</td>
<td>227</td>
</tr>
<tr>
<td>Table 11. Holistic quality difference scores across revision studies</td>
<td>228</td>
</tr>
<tr>
<td>Table 12. Revision difference scores across studies</td>
<td>231</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>Number of essay parts by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing</td>
<td>135</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>Holistic quality by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing</td>
<td>136</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Essay length by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing</td>
<td>137</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>Mechanical errors per 100 words by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing</td>
<td>147</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>Type of revision by student during baseline, SRSD posttesting, revision posttesting, and maintenance testing</td>
<td>204</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>Group mean scores on the strategy knowledge questionnaire</td>
<td>212</td>
</tr>
</tbody>
</table>
ABSTRACT

THE EFFECTS OF INSTRUCTION IN PEER-REVISION ON THE PERSUASIVE WRITING OF STUDENTS WITH EMOTIONAL AND BEHAVIORAL DISABILITIES

Sara Mills, Ph. D.

George Mason University, 2012

Dissertation Chair: Dr. Margo A. Mastropieri

A growing body of research supports the use of self-regulated strategy development (SRSD) to improve the writing skills of students with emotional and behavioral disorders (EBD). This single subject multiple baseline, multiple probe design study extends this research base by investigating the effects of peer-revision instruction on the persuasive essay writing of 10 eighth grade students (7 boys, 3 girls) with EBD. The study included: (a) baseline testing, (b) SRSD persuasive writing instruction, (c) SRSD posttesting, (d) revision instruction, (e) revision posttesting, and (f) six- to eight-week maintenance testing. Student performance on persuasive essay writing, on-task behavior, self-efficacy, writing fluency, and a social validity interview were collected.
Writing instruction, which followed the six stages of SRSD, was provided over eight to nine, 50-minute sessions for an average of 412.5 minutes. Revision instruction was provided over 7-10, 50-minute sessions for an average of 450 minutes of instruction. The revision strategy required students to give their partners feedback about four aspects of their essay: (a) What did the author do well? (Compliments were given at the beginning and end of the peer conference.) (b) Does it have all the parts [of a good persuasive essay]? (c) Is it clear? and (d) Is it persuasive?

Results showed that students’ writing improved on essay measures of content, quality, and length following the first instructional phase. Percent of nonoverlapping data (PND) with baseline were uniformly high. Revision instruction resulted in a decrease in mechanical errors in students’ essays, but did not further improve the content, quality, or length of students’ essays. However, PND with baseline were maintained. Additionally, students made more revisions following revision instruction, although these gains were not maintained over time. During all phases of the study, student primarily made surface-level revisions. Students were able to maintain gains in the content, quality, length, and mechanics of their essays over time. They were also able to work effectively together during peer-conferences, as measured by high rates of on-task behavior and their ability to complete nearly all components of the peer-revision conferences. Students did not show improvement in self-efficacy or writing fluency as a result of this study. Lastly, students reported that they found both the SRSD and revision strategies useful and enjoyable. Limitations, educational implications, and directions for future research are discussed.
1. INTRODUCTION

From the inception of America’s public schools in the 19th century, boys and girls across the country have been tasked with learning the three Rs – reading, writing, and ‘rithmetic. However, over the last several decades, writing has received considerably less attention than reading and arithmetic in schools. At the same time, technological advances have made writing a daily requirement for many jobs, and a regular routine in the lives of individuals. Whether it is composing an email, writing a report, or blogging effective writers know how to organize their thoughts, compose clear sentences, and use proper spelling, grammar, and mechanics to convey ideas. Writing provides individuals with an outlet to express ideas and emotions, organize their thoughts, and persuade others to a way of thinking. Too often, however, schools are failing to adequately teach students critical writing skills that will allow them to be successful in their future employment and personal lives.

The 2007 National Assessment of Educational Progress (NAEP) results indicated that the majority of students completing high school (58%) were at the basic level. The basic level indicates “partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at a given grade” (U. S. Department of Education, 2008, p. 6). Nearly 1 in 5 students (18%) performed even lower than this minimal level. While there was an increase in the number of students at the basic level between 2002 and 2007,
there was not an increase in the number of students at the proficient or advanced levels (U.S. Department of Education, 2008). In 2007, only 24% of twelfth graders were rated as proficient or advanced, which means they had mastered the writing knowledge and skills necessary for completing high school (U.S. Department of Education).

It comes as no surprise, then, that many students are not adequately prepared to meet the writing demands of college or work after graduating from high school. A survey conducted by Achieve, Inc. (2005) found that 35% of college students reported that there were gaps between their high school preparation and the writing expectations in higher education. College instructors saw a larger gap between preparation and expectations, reporting that 50% of their students are not prepared for college level writing. In 2000, 14% of all college freshmen were required to take a remedial writing course (Parsad & Lewis, 2003).

The problem continues once students leave school, impacting their ability to obtain and maintain jobs. The National Commission on Writing (2004) reported that writing is a “threshold skill” for those seeking salaried employment; meaning those who cannot communicate clearly in writing are not hired for salaried positions. In their national survey of business leaders, they found that business leaders report written communication skills also play a role in promotion decisions. Furthermore, businesses spend more than $3.1 billion each year to remediate their employees writing deficits. These findings were corroborated by the Achieve, Inc. (2005) survey, which found that 38% of both recent high school graduates and their employers felt high school graduates
were “inadequately prepared for the quality of writing that is expected” in the workplace (p. 6).

The gap between the writing demands of the “real world” and the lack of preparation students receive in school is of particular concern for students with disabilities. In 2007, NAEP testing on the differences in skill level between students with and without disabilities was substantial. Only 54% of eighth-grade students with disabilities were at, or above, the basic level, while 91% of their non-disabled peers were at, or above, the basic level (U.S. Department of Education, 2008). Moreover, only 6% of students with disabilities were at the proficient level; meaning that they had the knowledge and skills necessary to proficiently complete grade level work. In contrast, 33% of their non-disabled peers were at, or above, the proficient level (U.S. Department of Education). Given these extreme disparities, there is a particular need for improved writing instruction for students with disabilities.

Writing Instruction

Writing is a complex task, requiring the integration of basic writing skills (e.g., spelling, grammar, punctuation) with higher order thinking (e.g., organization, reasoning). In the past, writing instruction in schools often focused on isolated skills, such as spelling or grammar. In the 1980’s, the focus of writing instruction shifted to encompass the entire writing process. Hayes and Flowers (1980) put forth a model of writing that consisted of three stages: (a) planning, (b) translating, and (c) revising. Subsequently, the workshop approach (Graves, 1985) to writing instruction became the norm in American schools. In this approach, students take part in the entire writing
process to produce written work—planning, drafting, revising, and eventually, publishing their writing. Popular programs, such as Writer’s Workshop (Calkins, 1986), begin with students as young as kindergarten and continue through the elementary school years. Writing at the secondary level focuses primarily on demonstrating content knowledge through writing.

Research on effective writing instruction has similarly shifted to incorporate a process writing approach. In the field of special education, most research has focused on teaching students a strategy for writing. That is, students are taught to engage in all steps of the writing process, using a specific strategy to guide what they do at each step. For example, Cognitive Strategy Instruction in Writing (Englert, 2009) uses the acronym POWER to teach students the steps of the writing process—Plan, Organize, Write, Edit, and Revise. Through dialogue, the teacher thinks aloud and models the writing process. Students are also given “think sheets” to scaffold their thinking through each step, from planning to revising.

**Self-regulated strategy development (SRSD)** uses a similar approach, but incorporates self-regulation skills in the context of writing instruction, including goal setting, self-instruction, self-monitoring, and self-evaluation. SRSD also uses mnemonics to remind students of the essential components of writing for different genres. For instance, SRSD story writing follows the mnemonic POW+WWW, What = 2, How = 2:

Pick my idea, Organize my notes, Write and say more
Who are the main characters? When does the story take place? Where does the story take place?

What do the main characters want to do? What happens when the main characters do it? How does the story end? How do the main characters feel?

SRSD persuasive writing, on the other hand, follows the POW+TREE mnemonic: Pick my idea, Organize my notes, Write and say more + Topic, Reasons, Explanations, and Ending.

While most writing research in special education has focused on strategy instruction, little research has included instruction in the revision stage of the writing process. Mills (2011) found that only 10 studies have been conducted that teach revision to students with learning disabilities (LD), and only one study has been conducted teaching revision to students with emotional and behavioral disorders (EBD). Given this limited number of studies, there is no conclusive research base that identifies the essential components of effective revision instruction. In six of the studies, revision instruction was taught in the context of the entire writing process. That is, the writing strategy included instruction in planning, drafting, and revising essays. In the remaining five studies, revision was taught as a stand-alone strategy; students did not receive any instruction in planning and drafting. Although results varied somewhat based on the type of study conducted (e.g., quasi-experimental vs. single-subject studies), strategies that incorporated revision within the writing process generally obtained higher, more consistent effect sizes (ES = 1.07, PND = 80%) than did revision-only strategies (ES = 0.40, PND = 91.5%; Mills, 2011). One characteristic that consistently distinguished
more effective revision strategies from less effective strategies was the use of peer-learning for revision. *Peer-mediated instruction* studies obtained large effect sizes (ES = 1.15, PND = 100%), while the effect sizes for individually implemented revision strategies were in the moderate range (ES = 0.47, PND = 84.33%; Mills, 2011).

While writing is a challenge for many students, students with EBD struggle with writing tasks more than do most of their peers. They often perform below grade level, or years behind their peers in writing (Nelson, Benner, Lane, & Smith, 2004; Reid, Gonzalez, Nordness, Trout, & Epstein, 2004; Trout, Nordness, Pierce, & Epstein, 2003). Writing performance for students with EBD is similar to peers with LD and *Attention Deficit/Hyperactivity Disorder (ADHD)*; Trout et al.). These deficits are particularly pronounced for students with *externalizing behaviors*, as opposed to *internalizing behaviors*. Adding to their difficulty, a disproportionate number of students with EBD also exhibit *language disorders*, which impact their *expressive, receptive, and pragmatic language skills*. As a result, they have difficulty expressing themselves both verbally and in writing.

**Characteristics of Students with Emotional and Behavioral Disorders (EBD)**

Students with EBD have a number of characteristics that distinguish them from the general population. To begin, students with EBD are predominantly male—80% of students with EBD in elementary school, and 76% of secondary students with EBD are male (Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005). Additionally, African Americans are overrepresented, making up a quarter of all students with EBD, while comprising only 17% of the total school-age population (Wagner et al.). Furthermore, a
large proportion of students with EBD are living in poverty (33.2% in elementary, 29.8% in secondary; Wagner et al.). The vast majority of students with EBD are served in regular school buildings, although only 25% of students with EBD spend the majority of their day in general education classrooms (Bradley, Doolittle, & Bartolotta, 2008).

Students with EBD have a high rate of comorbidity, meaning they have additional disabilities in conjunction with EBD (e.g., a student may have EBD and LD). In particular, students with EBD have a high occurrence of language disorders that affect their receptive, expressive and pragmatic language skills. Seventy-one percent of students with EBD have language deficits (Benner, Nelson, & Epstein, 2002). When looking at public school settings only, the proportion of students with both EBD and language deficits is even higher; 9 out of 10 students with EBD served in public school settings have language deficits (Benner et al.).

Behavioral Characteristics of Students with EBD

Students with EBD, by definition, display serious behavioral difficulties that interfere with their learning. Their behaviors are classified as either externalizing or internalizing. Externalizing behaviors refer to “acting out” behaviors such as noncompliance, yelling, and physical aggression. Internalizing behaviors, on the other hand, are directed inward and include behaviors such as anxiety and depression. Individuals with EBD can also display a combination of both externalizing and internalizing behaviors.
Academic Characteristics of Students with EBD

In accordance with the Individuals with Disabilities Education Act (IDEA), to qualify for special education services, a student’s disability must have an adverse impact on his or her educational functioning. Students with EBD perform below their same-age peers in all academic areas (Nelson et al., 2004; Trout et al., 2003). Approximately 4 out of 5 students with EBD perform below age-level norms in reading, writing, and mathematics (Nelson et al.; Reid et al., 2004). Their greatest academic deficits are in math and spelling (Reid et al.). Students who exhibit externalizing behaviors are more likely to have academic deficits than are students with internalizing behaviors (Nelson et al.). Little is known about differences in academic functioning for girls and students of color because this information is not regularly reported in the research on academic characteristics of students with EBD (Trout et al.). What is clear from research is that students with EBD have extensive academic deficits that require intensive academic interventions.

For over 40 years, those in the field of special education have highlighted the need for more useful information and effective strategies to better serve students with EBD (Bradley, Henderson, & Monfore, 2004). Most educational research has focused on behavioral interventions for these students (Wehby, Lane, & Falk, 2003), with the assumption that improvement in student behaviors will lead to improvement in academic performance. Unfortunately, the research suggests this assumption is not accurate. When compared to peers with other mild disabilities, such as learning disabilities (LD), students with EBD show a distinct lack of progress over time (Anderson, Kutash, &
Duchnowski, 2001). Therefore, research-based academic interventions targeted to the needs of students with EBD are sorely needed.

**Writing for Students with EBD**

Written expression is one academic area where students with EBD perform lower than most of their peers (Nelson et al., 2004). While there is over 30 years of research on writing interventions for students with LD, from early elementary grades through college, to date there are fewer than a dozen studies published in peer-reviewed journals that focus on teaching writing to students with EBD (Cuenca-Sanchez, Mastropieri, Scruggs, & Kidd, 2012; Lane et al., 2008; Little et al., 2010; Mason, Harris, & Graham, 2002; Mason, Kubina, Valasa, & Mong Cramer, 2010; Mason & Shriner, 2008; Mastropieri et al., 2009; 2010; Morris Kindzierski, 2009; Regan, Mastropieri, & Scruggs, 2005). Therefore, a great need exists to develop a research base to identify effective instructional strategies to teach writing to students with EBD.

The majority of the intervention research on writing for students with EBD uses strategy instruction to teach the planning and drafting portions of the writing process (Mason & Shriner, 2008; Mastropieri et al., 2009). More specifically, eight of the most recent studies (Cuenca-Sanchez et al., 2012; Lane et al., 2008; Little et al., 2010; Mason et al., 2002; 2010; Mason & Shriner, 2008; Mastropieri et al., 2009; 2010) used the self-regulated strategy development (SRSD) approach for teaching persuasive or story writing to students with EBD in elementary and middle schools. The first of these (Mason et al., 2002) is a descriptive study of a third grade girl with behavioral and learning difficulties
who learned the SRSD strategy for story writing. Following instruction, the student had increased the length of her stories and included more story parts.

In two studies (Lane et al., 2008; Little et al., 2010), SRSD was taught to second grade students at risk for EBD within the context of a schoolwide positive behavioral support system. In both studies, students improved the length, quality, and number of text structure elements in their writing.

Two additional studies (Mason et al., 2010; Mastropieri et al., 2009) used SRSD to improve the writing fluency of middle school students with EBD. In both studies, the SRSD strategy for persuasive writing was taught to students. Students were then instructed how to apply the strategy to write a complete essay within 10 minutes. Results from Mastropieri et al. (2009) indicated that students with EBD, who were performing below grade level in writing, improved the length, quality, and number of text structure elements following instruction. Participants in the Mason et al. (2010) study were all performing on-grade level in writing prior to receiving instruction in writing fluency. As a result of instruction, participants demonstrated more consistency in their writing, which improved the overall quality of their written products.

The other studies that used SRSD to teach writing to students with EBD (Cuenca-Sanchez et al., 2012; Mason & Shriner, 2008; Mastropieri et al., 2010) reported similarly positive results following the strategy instruction. Mason and Shriner (2008) taught the strategy to second through fifth grade students with EBD, while Cuenca-Sanchez et al. (2012) and Mastropieri et al. (2010) focused on eighth-graders with more severe EBD. All three studies reported improvements in text length, quality, number of transition
words, and number of text elements following instruction. Additionally, Mastropieri et al. noted improvements in students’ writing fluency. Students in the Mastropieri et al. study maintained their gains over a 12-week time period. Taken together, these studies indicate that SRSD is a promising intervention for teaching writing to elementary and middle school students with EBD.

Peer Learning for Students with EBD

While improving academic outcomes for students with EBD is a real and pressing concern, instruction for students with EBD must also take into account the social and behavioral learning needs of these students. Students with EBD can benefit from explicit instruction in social skills. Integrating instruction in social skills into academic instruction is a potentially powerful way to meet the social and behavioral needs of students with EBD, while still maximizing the amount of time spent on academic instruction. Peer-tutoring, for example, has been shown to increase the academic achievement for students with EBD. It also improves the behavioral performance of adolescents with EBD (Spencer, 2006; Spencer, Simpson, & Oatis, 2009). If peer-learning opportunities are not well-structured, however, students with EBD perform more poorly both academically and behaviorally (Sutherland, Wehby & Gunter, 2000).

The Present Study

As these studies show, students with EBD can improve their writing performance with direct strategy instruction in writing. However, current studies have been limited to the planning and drafting stages of writing. Including revision instruction, as well as peer-learning components can have positive effects of student learning.
Instruction in persuasive writing has many benefits for students with EBD. First, their writing skills improve as a result of such instruction (Mason et al., 2010; Mastropieri et al., 2009). Based on research conducted with students with LD, it is reasonable to expect that additional instruction in revision will further improve the mechanical aspects of writing for students with EBD (Graham & MacArthur, 1988; MacArthur, Schwartz, & Graham, 1991; and Stoddard & MacArthur, 1993). Combining strategic writing instruction with instruction in revision means that students will be more likely to meet grade level expectations in writing, and perform better in school. Second, explicit strategic writing instruction teaches students with EBD to organize their thoughts and express their ideas in a coherent, meaningful way. Third, by focusing on persuasive writing, students with EBD learn to take the perspective of others and effectively self-advocate for their needs and wants (Cuenca-Sanchez, 2012).

Structured and scaffolded opportunities to work with peers also have the potential to improve the academic and social performance of students with EBD. Peer-mediated instruction increases students’ time on task, increases opportunities for all students to respond, and incorporates immediate feedback throughout the task (Spencer, 2006). Peer-tutoring arrangements, in particular, have a positive impact on the academic outcomes of students with EBD, especially when students have the opportunity to act as the tutor in such arrangements (Ryan, Reid, & Epstein, 2004).

**Purpose and Research Questions**

The purpose of this study is to (a) replicate and extend the research on SRSD writing instruction for students with EBD by including explicit, strategic instruction in
revising; and (b) to extend the research base on effective peer learning for individuals with EBD by teaching students to work with partners to revise their essays. Research questions include:

- Does persuasive writing improve (as measured by length, content, quality, and mechanics) after instruction in SRSD and after instruction in revision?
- What types of revisions do students make using the peer-revision process? What impact do these revisions have on the length, content, quality, and mechanics of the essay?
- Does student self-efficacy for writing improve following SRSD and revision instruction? Are changes in self-efficacy maintained over time?
- Does students’ writing fluency improve as a result of instruction in the SRSD and revision strategies?
- Are students with EBD able to accurately conduct all steps of the peer-revision strategy? Are students with EBD socially able to work together as measured by time on task (as compared to time on task during teacher-led instruction)?
- Do students find the SRSD and revision strategies easy to implement, useful, and enjoyable?

Definition of Terms

**Attention Deficit/Hyperactivity Disorder (ADHD)**

ADHD is diagnosed based on the following criteria, set forth in the Diagnostic and Statistical Manual, 4th Edition (DSM-IV; American Psychiatric Association, 2000):

A. Either (1) or (2):
(1) six (or more) of the following symptoms of *inattention* have been persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level:

**Inattention**

(a) often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities
(b) often has difficulty sustaining attention in tasks or play activities
(c) often does not seem to listen when spoken to directly
(d) often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (not due to oppositional behavior or failure to understand instructions)
(e) often has difficulty organizing activities
(f) often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork or homework)
(g) often loses things necessary for tasks or activities (e.g. toys, school assignments, pencils, books, or tools)
(h) is often easily distracted by extraneous stimuli
(i) is often forgetful in daily activities

(2) six (or more) of the following symptoms of *hyperactivity-impulsivity* have persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level:

**Hyperactivity**

(a) often fidgets with hands or feet or squirms in seat when sitting still is expected
(b) often gets up from seat when remaining in seat is expected
(c) often excessively runs about or climbs when and where it is not appropriate
   (adolescents or adults may feel very restless)
(d) often has trouble playing or doing leisure activities quietly
(e) is often "on the go" or often acts as if "driven by a motor"
(f) often talks excessively

Impulsivity
(a) often blurts out answers before questions have been completed
(b) often has difficulty awaiting turn
(c) often interrupts or intrudes on others (e.g., butts into conversations or games)

B. Some hyperactive-impulsive or inattentive symptoms that caused impairment were
   present before age 7 years.

C. Some impairment from the symptoms is present in two or more settings (e.g. at school
   [or work] and at home).

D. There must be clear evidence of clinically significant impairment in social, academic,
   or occupational functioning.

E. The symptoms do not occur exclusively during the course of a Pervasive
   Developmental Disorder, Schizophrenia, or other Psychotic Disorder and are not
   better accounted for by another mental disorder (e.g. Mood Disorder, Anxiety
   Disorder, Dissociative Disorder, or a Personality Disorder).

Based on these criteria, three types of ADHD are identified:

Attention-Deficit/Hyperactivity Disorder, Combined Type: if both criteria A1 and
A2 are met for the past 6 months

Attention-Deficit/Hyperactivity Disorder, Predominantly Inattentive Type: if criterion A1 is met but criterion A2 is not met for the past six months

Attention-Deficit/Hyperactivity Disorder, Predominantly Hyperactive-Impulsive Type: if Criterion A2 is met but Criterion A1 is not met for the past six months

Comorbidity

Comorbidity refers to two or more conditions occurring within one individual (Hallahan, Kauffman, & Pullen, 2008). For example, an individual may have both LD and EBD, or EBD and ADHD.

Cooperative Learning

Cooperative learning is one type of peer-mediated instruction. It utilizes group structures for students to learn information and perform tasks. As originally intended, there are five critical features of cooperative learning: (a) positive interdependence, meaning students have to work together to accomplish a task; (b) individual accountability; (c) promotive interaction, meaning that students meet face-to-face to promote each other’s work (Sutherland, Wehby, Gunter, 2000); (d) group processing; and (e) development of small group social skills (Sutherland et al., 2000).

Emotional and Behavioral Disorders

Emotional and behavioral disorders (EBD) refers to those disorders meeting the criteria for emotional disturbance according to guidelines in IDEA 2004 (see definition below), or meeting criteria outlined in the DSM-IV for various emotional or behavioral disorders.
Emotional Disturbance

IDEA 2004 defines “emotional disturbance” as

... a condition exhibiting one or more of the following characteristics over
a long period of time and to a marked degree that adversely affects a child’s
educational performance:

(A) An inability to learn that cannot be explained by intellectual, sensory,
or health factors.

(B) An inability to build or maintain satisfactory interpersonal
relationships with peers and teachers.

(C) Inappropriate types of behavior or feelings under normal
circumstances.

(D) A general pervasive mood of unhappiness or depression.

(E) A tendency to develop physical symptoms or fears associated with
personal or school problems.

(ii) Emotional disturbance includes schizophrenia. The term does not
apply to children who are socially maladjusted, unless it is determined that they
have an emotional disturbance under paragraph (c) (4) (i) of this section. (34
C.F.R. § 300.8 (c) (4), 2006, p. 46756)

Expressive Language Deficits

Expressive language deficits refer to disabilities affecting an individual’s spoken
or written language.
**Externalizing Behaviors**

Externalizing behaviors are observable behaviors directed at others (Hallahan, Kauffman, & Pullen, 2008). These “acting out” behaviors include behaviors such as physical aggression, yelling, or defiance.

**Individuals with Disabilities Education Improvement Act (IDEA 2004)**

IDEA 2004 is the Federal law governing special education. It details how states and local education agencies provide services to students with disabilities, from birth to age 22. Part B applies specifically to children ages 3-22, while Part C applies to infants and toddlers, from birth to age 3.

**Intellectual Disabilities**

Intellectual disability is a disability characterized by significant limitations both in intellectual functioning and in adaptive behavior, which covers many everyday social and practical skills. This disability originates before the age of 18 (American Association on Intellectual and Developmental Disabilities, 2011).

**Internalizing Behaviors**

Internalizing behaviors are negative behaviors that are directed toward the self, such as depression, anxiety, or withdrawal (Hallahan et al., 2008).

**Learning Disabilities (LD)**

IDEA 2004 defines a specific learning disability as:

…..a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or
to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia.

Disorders not included. Specific learning disability does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage. (34 C.F.R. § 300.8 (c) (10))

**Metacognitive Processes**

Metacognitive processes are those processes by which individuals think about their own thinking and learning. For example, individuals use metacognitive processes to self-regulate their own learning—setting goals for their learning, monitoring their progress toward those goals, and evaluating their performance on learning tasks.

**Mild Disabilities**

Mild disabilities are those disabling conditions that are occur more frequently in the general population and require less intensive interventions for support. Examples of mild disabilities include LD, EBD, ADHD, and mild intellectual disabilities.

**No Child Left Behind Act of 2001 (NCLB)**

The Federal law, formally known as the Elementary and Secondary Education Act regulates what states and local education agencies must do to receive Federal education money. One of the most debated features of NCLB is its requirement that all students participate in high-stakes testing in reading and mathematics, beginning in third grade (U.S. Department of Education, 2004).
Peer-mediated Instruction

Peer-mediated instruction is student-directed learning methods in which students implement teacher-designed instruction to teach their peers. It includes approaches such as peer tutoring and cooperative learning.

Peer Tutoring

Peer tutoring is a peer-mediated instructional approach in which one student acts as a tutor, and another student acts as the tutee to learn information or practice skills. In same-age peer tutoring, peers with similar skill levels work together, taking turns as tutor and tutee. In cross-age peer tutoring, an older, more skilled student works with a younger, less-skilled student and the tutor and tutee roles are fixed (Mastropieri & Scruggs, 2010).

Pragmatic Language Deficits

While individuals with pragmatic language deficits understand the meaning of words, they have significant difficulty understanding how language is used in social situations. For example, an individual with pragmatic language deficits may not understand voice inflection cues, or the role of turn-taking in conversations.

Receptive Language Deficits

Receptive language deficits refer to disabilities affecting an individual’s ability to understand spoken or written language.
**Self-regulated Strategy Development**

Self-regulated strategy development (SRSD) is one research-based approach to writing instruction that combines strategy instruction and explicit instruction in self-regulation techniques to improve students’ writing.

**Strategy Instruction**

Strategy instruction refers to instruction that teaches students strategies that can be applied in various situations to accomplish given tasks. In writing, strategy instruction refers to instruction that teaches students strategies to plan, draft, and/or revise. Instruction proceeds from teacher modeling to guided practice to independent practice.

**Writing Process**

The writing process, as conceptualized by Hayes and Flowers (1980) consists of three stages—planning, translating (or drafting), and revising. Writers move in and out of these stages as they write; it is not conceptualized as a linear process.

**Summary**

Writing is a critical skill that individuals use on a daily basis. Unfortunately, many students leave high school unprepared to meet the writing demands of college and employers. This lack of preparation is particularly notable for students with disabilities, including students with EBD. Students with EBD struggle with writing. Moreover, they have behavioral, social, and academic characteristics that negatively impact their classroom performance.

Writing strategy instruction, such as SRSD, has been identified by research as an effective academic intervention for students with LD and, more recently, for students
with EBD. However, little is known about effective revision instruction for students with
disabilities. Peer learning has also been found to be an effective academic intervention
for students with EBD. The present study combines SRSD writing instruction with a
peer-revision strategy to improve the persuasive writing of students with EBD.
2. LITERATURE REVIEW

This chapter begins with a review of the literature on the characteristics of students with emotional and behavioral disorders (EBD), focusing particularly on academic characteristics. Then, research on academic interventions for students with EBD is reviewed, including a summary of the use of peer-mediated instruction with students with EBD. Finally, an in-depth review of writing intervention research for students with mild disabilities is provided, with an emphasis on Self-regulated Strategy Development (SRSD) for writing with students with EBD and revision instruction.

While the Individuals with Disabilities Education Improvement Act (IDEA 2004) uses the term “emotional disturbance” (ED), this research study uses the term emotional and behavioral disorders (EBD). EBD is used by many practitioners in the field because it more accurately defines the range of difficulties that children served under the ED label experience (Hallahan, Kauffman, & Pullen, 2008). When discussing IDEA 2004 and the number of students identified for special education service, the term, “students with ED” will be used. In all other instances, the term “students with EBD” will be used.

Characteristics of Students with EBD

IDEA 2004 defines “emotional disturbance” as
... a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child’s educational performance:

(A) An inability to learn that cannot be explained by intellectual, sensory, or health factors.

(B) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers.

(C) Inappropriate types of behavior or feelings under normal circumstances.

(D) A general pervasive mood of unhappiness or depression.

(E) A tendency to develop physical symptoms or fears associated with personal or school problems.

(ii) Emotional disturbance includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance under paragraph (c) (4) (i) of this section. (34 C.F.R. § 300.8 (c) (4), 2006, p. 46756)

The 29th Annual Report to Congress on the Implementation of IDEA reports that there are 471,306 students ages 6-21 who have a primary special education label of ED (U.S. Department of Education, 2007), which represents 7.81% of the total special education population, but only 0.90% of the total number of students enrolled in public schools. Of these, 131,254 are students ages 6-11, which is 4.81% of the total number of 6-11 year-old students with disabilities. For students aged 12-17, 310,842 are labeled as students with ED, which represents 10.43% of the total special education population at
this age group. The remaining 29,210 students with EBD are 18-21 years old. The
difference in prevalence rates from elementary-age to secondary-age students is largely a
result of younger students with a primary label of speech-language impairment being
recategorized as they get older (Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005).
Nearly 80% of students receiving services as students with ED are boys and a
disproportionately large number of these boys are African American, while a
disproportionately small number are Hispanic (Wagner et al., 2005). Further, a large
proportion of parents indicate that their children with EBD have comorbid conditions
such as ADHD (about two-thirds) and LD (about one-fourth; Wagner et al.).

A study of two national longitudinal data sets, the Special Education Elementary
Longitudinal Study (SEELS) and the National Longitudinal Transition Study–2 (NLTS-2),
by Wagner and colleagues (2005) found that students with EBD are disproportionately
at risk on a number of indicators when compared to their peers with other types of
disabilities and those without disabilities. Specifically, one third of students with EBD
live in single-parent homes. One fifth of these students live in homes where the head of
household is not a high school graduate and is not employed. Almost half (45%) live
with another person with a disability. Taken together, these statistics point to the severe
economic stress that families of students with EBD often face.

The same study (Wagner et al., 2005) found that students with EBD are more
likely than their peers, with other and without, disabilities to have a number of negative
school experiences. For example, while students with EBD start to demonstrate school-
related problems at about the same time as other students with disabilities, on average,
they receive services a year later. Specifically, students with EBD and students with other disabilities first start having difficulty at 4.6 and 4.4 years of age, respectively. Students with other disabilities begin to receive special education services by 6.7 years of age, while students with EBD do begin to receive services until 7.8 years of age (Wagner et al., 2005). Given the effectiveness of early intervention, this year is critical for improving students’ long-term outcomes. Students with EBD are retained at rates similar to their peers with other disabilities, but more than their peers without disabilities. Additionally, students with EBD have higher rates of school mobility than their peers, and are four times more likely to be suspended or expelled than their peers with and without disabilities. The parents of students with EBD report that it is difficult for them to get services for their children. Additionally, these parents report that they spend more time than parents of students with other disabilities working with their children on homework.

Postschool outcomes for students with EBD are equally dismal. Half of all students with EBD drop out of high school, the highest rate for any disability category (Wagner et al., 2005). Only 1 out of 5 adolescents with EBD pursues some sort of education after high school (Bradley, Doolittle, & Bartolotta, 2008). Unemployment rates for young adults with EBD during the first few years after high school range from 42% to 70%, which is higher than the unemployment rates for their peers with other disabilities (Bradley et al., 2008; Carter & Wehby, 2003). Of the young adults with EBD who have jobs one year after high school, 35% do not have a job two years after high school (Carter & Wehby, 2003). Finally, youth with EBD have high rates of involvement with the justice system, including a startling arrest rate of 58% (Bradley et al., 2008).
Functional Characteristics of Students with EBD

The behaviors of students with EBD are typically classified into three categories: *externalizing, internalizing*, or *comorbid*. Students with externalizing behaviors are seen as acting out against others, while students with internalizing behaviors tend to suffer from disorders like anxiety and depression (Hallahan, Kauffman, & Pullen, 2008). A student with EBD may also have two or more conditions at the same time, which is known as comorbidity. The majority of students with EBD exhibit externalizing behaviors. A study of the types of behavior problems exhibited by public school students with EBD found that twice as many students demonstrated externalizing behaviors as internalizing behaviors (Nelson, Babyak, Gonzalez, & Benner, 2003). Furthermore, elementary-age students with EBD had more complex and severe problems than secondary students with the disability. Results also suggested that girls had to exhibit more severe behaviors than boys to be identified as and provided services for students with EBD (Nelson et al., 2003).

A large proportion of students with EBD also have language disorders (Benner, Nelson, & Epstein, 2002; Nelson, Benner, Neill, & Stage, 2006; Wagner et al., 2005). Those with *receptive language deficits* have more antisocial behaviors, while those with *expressive language deficits* tend to be more withdrawn. Additionally, individuals with receptive language deficits are more likely to also have reading deficits (Benner et al., 2002). A review of the literature on language disorders in students with EBD (Benner et al., 2002) found that 71% of children identified as having EBD had particular language deficits. Of these, 71% had *pragmatic language deficits*, 64% had expressive language
deficits, and 56% had receptive language deficits. Primary studies that focused on students identified with language deficits found that 57% of these children also had EBD. The comorbidity of the two disorders was stable over time.

Parent surveys provide further evidence of the co-occurrence of language deficits and EBD (Wagner et al., 2005). About one third of parents reported that their children with EBD have trouble carrying on conversations. Additionally, 44.2% of parents of elementary-aged students with EBD and 29.4% of parents of secondary-aged students with EBD reported that their children had trouble understanding what others say. These language deficits have a large impact on learning as research has shown that students are expected to learn through listening at least 60% of the time in elementary school and 90% of the time in high school (Benner et al., 2002). Therefore, interventions that develop language skills will benefit students with EBD.

**Academic Characteristics of Students with EBD**

According to the federal special education legislation, IDEA, to qualify for special education services as a student with ED, emotional or behavioral problems must adversely affect the student’s ability to learn and achieve in school. Several studies have shown that students with EBD perform at lower levels than their same-age peers in reading, writing, and mathematics (Nelson, Benner, Lane, & Smith, 2004; Reid, Gonzalez, Nordness, Trout, & Epstein, 2004; Trout, Nordness, Pierce, & Epstein, 2003). Trout and colleagues (2003) reviewed the research on the academic status of students with EBD. Ninety-one percent of the 65 studies included in the review found these students to be “academically deficient,” performing below grade level or years behind
their nondisabled peers in reading, writing, and mathematics. Similarly, a meta-analysis by Reid and colleagues (2004) found that students with EBD have moderate to large deficits in academics (ES = -0.64). Students with EBD performed poorly across all academic areas, revealing the largest gaps between them and their same-age nondisabled peers in math and spelling. This was true for both older (12 years and older) and younger (less than 12 years) students, indicating that these academic deficits were not remediated over time.

Nelson et al. (2004) found similar results in their study investigating the academic differences between students with EBD in grades K-12 and the same-age normative sample. The study included 155 students with EBD (126 boys, 29 girls), and included measures of behavior, i.e. Child Behavior Checklist: Teacher Report Form (Achenbach, 1991) and performance on a standardized academic achievement test, via Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). Data were also gathered from student records on ethnicity, the hours of special education services students received each day, and IQ scores. Results indicated large academic deficits for students with EBD in grades K-12, with an effect size of 0.94. An effect size of this magnitude reveals that 83% of the students with EBD performed below the mean of the norm group across subtests and clusters, which included reading, writing, and mathematics. No significant differences in academic achievement were found between boys and girls. As in other studies, there was little difference between the performance of younger students and older students, the exception being mathematics in which older students performed more poorly than younger students. Finally, students with
externalizing behaviors were more likely to have academic achievement deficits than were students with internalizing behaviors (Nelson et al., 2004).

While research has documented the academic deficits of students with EBD, many questions remain. Few studies of the academic performance of students with EBD provide information about the functioning of girls as compared to boys (Reid et al., 2004; Trout et al., 2003). Only a small number of studies include information about participants’ ethnicity or socio-economic status, limiting understanding about how these important variables relate to academic functioning (Reid et al., 2004; Trout et al., 2003). Additionally, nearly all available studies were conducted in self-contained settings (Trout et al., 2003). While most students with EBD are served in self-contained settings, one third of these students spend most of their day in general education settings. Strategies that may be implemented in self-contained settings may not be appropriate for more general populations. Therefore, research on academic strategies that benefit students with EBD within general education classrooms are greatly needed.

Academic characteristics of students with EBD compared to students with LD. Understanding how students with EBD perform compared to their peers with other mild disabilities can also be useful for determining how students can best be served and which interventions might be appropriate for students with different disabilities. Unfortunately, the results of comparison studies have been highly variable. When looking across all such studies, Trout et al. (2003) found students with EBD to perform most similarly to students with attention deficit/hyperactivity disorder (ADHD). Additionally, they performed similarly to their peers with learning disabilities (LD) in
math and written expression. Finally, they outperformed their peers with *intellectual disabilities (ID)* in all academic areas.

One difficulty in interpreting the results of studies comparing students with EBD to other students with mild disabilities (most notably, students with LD), is that the studies typically use cross-age samples, capturing individual students at one point in time. As such, they provide little understanding of how the academic performance of the two groups varies over time. An important study by Anderson, Kutash, and Duchnowski (2001) investigated the academic progress of students with EBD and students with LD over time. Participants (79% male) included 42 students with EBD and 61 students with LD. The majority of participants (66%) were Caucasian; 34% were African American. Standardized test scores in reading and mathematics were collected when the students were in kindergarten or first grade, and again when they were in fifth or sixth grade as part of the special education eligibility and reevaluation process. Data on the predictor variables of attendance, behavioral discipline referrals, retention in early grades, school mobility, and type of special education setting were also collected for each student.

Results showed that students with EBD missed significantly more school, received significantly more referrals for behavioral issues, and received full-time special education services more often than did students with LD. However, the two groups did not differ significantly with regard to rates of early retention or school mobility. Both groups increased their scores in mathematics over time, but the differences between the groups remained consistent. In reading, scores for students with EBD were higher than scores for students with LD, at testing time one, and did not increase over time. Reading
scores for students with LD, on the other hand, improved over time. Finally, a step-wise multiple regression revealed three significant relationships between the predictor variables and achievement in reading and mathematics:

1. For students with EBD, lower reading scores at testing time two were associated with retention in kindergarten or first grade.

2. For students with LD, higher mathematics scores at testing time two were associated with more part-time special education services and not being retained in kindergarten or first grade.

3. For students with LD, higher reading scores at testing time two were associated with more part-time special education services (Anderson et al., 2001).  

These results suggest several important conclusions. First, retention was associated with lower academic performance, over time, for students with EBD and students with LD. Second, students with LD received more part-time special education services and showed more academic improvement than did students with EBD. Finally, three of the five predictor variables did not have any significant relationship to academic performance over time, including attendance, behavioral referrals, and school mobility. All in all, emotional and behavior disorders appear to have a more adverse impact on academic achievement, over time, than do learning disabilities.
Academic Interventions for Students with EBD

Why, then, do students with EBD have such significant academic deficits? Wehby, Lane and Falk (2003) propose several reasons. One possible reason is that the behaviors of students with EBD interfere with high-quality instruction. Another explanation may be that the behaviors of students with EBD influence teacher behaviors so that teachers provide less instruction. A third reason may be that most of the preparation for teachers working with students with EBD focuses on behavior management rather than on academic instruction. Lastly, limited research on academic interventions for students with EBD may impact the necessary level of training new teachers receive in this area. Whatever the reason(s), there is clearly a need for more focus on high-quality academic instruction for students—in the classroom, in teacher preparation programs, and in research institutions.

The lack of research on effective academic instruction for students with EBD is a serious concern. With the passage of *No Child Left Behind* in 2001, teachers are required to use evidence-based practices in the classroom. Without evidence of effective instructional practices for students with EBD, teachers are left without guidance about what works for this population of students. Passage of the initial federal special education legislation occurred in 1975: Yet, a descriptive review by Mooney, Epstein, Reid, and Nelson (2003) found only 55 articles in peer-reviewed journals from 1975 to 2002 that included students with EBD and had at least one dependent variable that was an academic outcome.
Analysis of the characteristics of these intervention studies revealed several important limitations in the literature on academic interventions for students with EBD (Mooney et al., 2003). First, the number of participants with EBD in intervention research has been steadily declining since the early 1980s. In fact, only 358 participants with EBD were identified over the 28-year period included in this review. As in studies of the characteristics of students with EBD, intervention studies rarely focuses on females, and provide inconsistent information about ethnicity, socio-economic status, and IQ—all of which impact the effectiveness of instruction in the classroom.

As in other types of studies, little intervention research focused on students with EBD is done in the general education classroom. The vast majority of research conducted with students with EBD occurs in special education classrooms or alternative settings. Most of the intervention studies used a single-subject design and employed curricular measures as dependent variables. Additionally, most research focused on teacher-mediated instruction, rather than child-mediated or peer-mediated instruction. The primary academic focus was on reading or mathematics, with very little research conducted in other subject areas. By and large, basic skills were addressed, rather than higher-level skills such as reading comprehension and mathematics reasoning. From a research design point of view, interrater agreement is usually reported, but social validity and treatment fidelity are not often reported (Mooney et al., 2003).

The small number of students with EBD included in academic intervention studies, coupled with the inconsistent reporting of such studies makes it difficult to determine which instructional practices are effective for this population of students.
Therefore, a great need exists to study academic interventions for students with EBD. Most past research in this area has focused on basic skills in reading and mathematics (Mooney et al., 2003). Very little is known about effective writing instruction for these students. Of particular interest to the current study is research on peer-mediated instruction for students with EBD and writing instruction. These two areas of research are reviewed below.

Peer-mediated Instruction for Students with EBD

Peer-mediated instruction for students with disabilities has been researched for over 30 years (Spencer, 2006; Spencer, Simpson, & Oatis, 2009). It has been used to teach reading, mathematics, writing, science, social studies, and behavior management (Ryan, Reid, & Epstein, 2004; Spencer, 2006; Spencer et al., 2009). Peer-mediated instruction has proven beneficial to a wide range of students across a variety of settings because it increases students’ time on task, provides every student with more opportunities to respond, and incorporates immediate feedback throughout the task (Spencer, 2006). These are all strategies that have been shown to improve academic achievement.

Peer-mediated instruction can take a number of forms, such as cooperative learning or peer tutoring. While some research has been conducted using cooperative learning strategies with students with EBD, the results of such studies have not supported the use of that approach for this population of students (Sutherland, Wehby, & Gunter, 2000). Peer tutoring, on the other hand, has been shown to improve the academic skills
of students with EBD. In some instances, peer tutoring has also had a positive impact on students’ classroom behavior (Ryan et al., 2004; Spencer, 2006; Spencer et al., 2009).

Peer tutoring can be implemented in several ways (Ryan et al., 2004). In cross-age, peer-tutoring formats, students act as tutors to other students of different ages (e.g., older students with EBD tutoring younger students struggling in math). Same-age peer tutoring involves students of the same age working together (e.g., one strong reader acts as the tutor to a weaker reader in the same class to complete a reading task). Many studies (Ryan et al., 2004) have used a reciprocal or reverse-role tutoring strategy where each participant takes on the role of tutor and tutee during the intervention (e.g., partners take turns reviewing vocabulary terms with each other).

Four research syntheses of peer-mediated instruction have been conducted over the last 10 years. The first (Sutherland et al., 2000) focused on the use of cooperative learning with students with EBD. Only eight articles were included in the review, representing 59 students with EBD, and 177 students identified as having either LD or EBD (numbers of students by disability area were not broken down in some primary studies). The authors concluded that the results of these studies did not support the use of cooperative learning with students with EBD. Poor study designs and reporting of results in primary studies made it unclear whether cooperative learning provided any benefits to the academic achievement, on-task behavior, or social interactions of students with EBD. Overall, the authors emphasized the need to train students in the group work and communication skills needed to successfully participate in cooperative learning, something the primary studies did not attempt to do.
A review by Ryan et al., (2004) focused on the impact of peer-mediated instruction, in general, on the academic outcomes of students with EBD. The review included 14 studies of cross-age, same-age, or classwide peer tutoring or cooperative learning. Most studies targeted either reading/language arts or mathematics. An overall effect size of 1.875 was reported across studies. Same-age peer tutoring arrangements had the strongest effects (ES = 1.92), followed by cross-age peer tutoring (ES = 1.12). While most studies did not report the data necessary to calculate effect sizes for tutors and tutees, those that did report such data showed that combining the roles of tutor and tutee (ES = 2.12, for four studies) and acting as the tutor (ES = 2.02, for two studies) had the largest effects on academic achievement. Acting as the tutee had only a moderate impact on achievement (ES = 0.63, for one study). Peer-mediated instruction was particularly effective for improving the academic achievement for adolescents (ES = 2.55), and benefited children as well (ES = 0.83). Overall, peer-mediated instruction proved to be a highly effective intervention for improving student achievement in all content areas.

Two literature reviews conducted by Spencer (2006) and colleagues (Spencer et al., 2009) also found peer tutoring to be an effective intervention for students with EBD. The first of these reviews (Spencer, 2006) included 38 studies conducted between 1972 and 2002 that included students with EBD. Most of these studies (a) incorporated cross-age tutoring designs, and (b) focused on reading. Spencer reported an overall ES of 0.72, based on data available from 13 of the 38 studies. (The other studies did not provide the data necessary to calculate effect sizes.) Studies were analyzed by grade level. At the
elementary level, students made gains in basic reading skills through the use of peer tutoring, but did not improve on measures of reading comprehension. At the middle school level, results were mixed. At the high school level, students made both academic and behavioral gains.

Spencer et al. (2009) updated this review to include nine new studies using children with EBD as tutors or tutees that were conducted between 2001 and 2007. All of these studies used same-age peer tutors. Five of the studies focused on reading, three focused on math, and one focused on behavior management. In the five studies with elementary-age students with EBD, students consistently showed gains in basic reading skills, but gains on behavioral measures were inconsistent. In two middle school studies, students showed both academic gains and improved behaviors in the peer-tutoring conditions. Lastly, two high school studies showed moderate gains in academics and improved behaviors. Across all studies, students reported enjoying the peer-tutoring interventions more than their typical instruction.

While most peer-mediated instruction for students with EBD has focused on drill and practice of basic skills, one study has used peer-mediated instruction in conjunction with strategy instruction to improve student learning in social studies (Spencer, Scruggs, & Mastropieri, 2003). This study included 30 middle school (seventh and eighth grade) students in a special school for students with EBD. A cross-over design was used in which all students participated in both the experimental and control conditions. In the experimental condition, students worked with same-age peers on a paragraph summarization strategy to read the grade-level textbook. Students took turns reading and
summarizing, with one student acting as the tutor and the other acting as the tutee. Then the roles reversed. In the control condition, students used guided notes and whole-class read aloud to cover the same textbook material. The study lasted four weeks, covering two units of instruction for two weeks each. Each of the six participating classes was placed in the experimental condition for one unit and in the control condition for the other.

Results showed that students obtained higher scores on content tests and on-task behavior measures during the tutoring condition. Effect sizes ranged from 0.40 to 0.89, favoring the tutoring condition in all cases. There were no differences in performance across grade levels or between boys and girls. Students reported that they liked working with their peers, although they did not particularly care for the summarization strategy. This study is important in that it demonstrated the effectiveness of peer-mediated instruction for teaching learning strategies to students with EBD (Spencer et al., 2003).

Writing Research

In 1980, Hayes and Flowers wrote their highly influential paper, “Identifying the Organization of the Writing Process.” In it, they discussed writing as a three-stage process of planning, translating (or drafting), and revising. In their conceptualization of the writing process, writers move in and out of the different stages as they write: Writing is not a linear activity. Metacognitive processes govern each stage of the writing process. What the writer thinks, and how the writer guides his thinking throughout the writing process, are important components of the entire process.
There are important differences in the ways novice and expert writers approach the writing task (Scardamalia & Bereiter, 1986). Unlike expert writers, for whom writing is a problem-solving process, novice writers engage primarily in knowledge telling. That is, for novices, writing is a matter of telling what they know about a topic. Their thinking is more linear than that of expert writers. Scardamalia and Bereiter (1986) suggest that, with guidance that provides “executive structure” (p. 792) during the writing process, novice writers can produce more mature text. For example, reprocessing, which refers to further processing of text that has already been written (both while writing and after writing), can be improved by providing novice writers with the executive procedures to revise their work.

The works of Hayes and Flowers (1980) and Scardamalia and Bereiter (1986) have guided writing research in the field of special education for nearly three decades. Instructional approaches tend to emphasize aspects of the writing process, such as planning and translating. Additionally, instruction in writing strategies (Mason & Graham, 2008; Rogers & Graham, 2008) has emerged as the most effective method for teaching students with mild disabilities to write.

**Meta-analyses of Writing Research**

Meta-analyses are a useful tool for summarizing a body of research on a given topic. Effect sizes (ES) are calculated to determine the relative effectiveness of different variables under study. For group experimental and quasi-experimental studies, the most common effect size reported is Cohen’s $d$. Standard guidelines for determining the
relative effects of interventions based on the \(d\) statistic (Lipsey & Wilson, 2001) are as follows:

- \(\text{ES} = 0.20\) are considered small effects.
- \(\text{ES} = 0.50\) are considered moderate effects.
- \(\text{ES} > 0.80\) are considered large effects.

For single-subject research (Scruggs, Mastropieri, & Casto, 1987), the most commonly reported effect size is percent of non-overlapping data (PND). Interventions with

- \(\text{PND} < 50\%\) are considered ineffective.
- \(\text{PND}\) ranging from 50-70\% are considered to have small effects.
- \(\text{PND}\) from 70-90\% are considered moderate effects.
- \(\text{PND} > 90\%\) are considered large effects.

**Meta-analyses of writing research for students without disabilities.** The Hayes and Flower (1980) conceptualization of the writing process changed writing research and instruction. As a result of their work, the teaching of writing has shifted from a focus on isolated aspects of writing, such as grammar instruction, to the writing process. The research on writing also shifted to focus on writing processes. Since 1980, seven meta-analyses of writing research for students without disabilities have been conducted (Bangert-Drowns, 1993; Bangert-Drowns, Hurley, & Wilkinson, 2004; Goldberg, Russell, & Cook, 2002; Graham & Hebert, 2010; Graham & Perin, 2007a, 2007b; Hillocks, 1986). Five of these focus on learning-to-write (Bangert-Drowns, 1993;
Goldberg et al., 2002; Graham & Perin, 2007a, 2007b; Hillocks, 1986), and two focus on writing-to-learn (Bangert-Drowns et al., 2004; Graham & Hebert, 2010).

The first of the learning-to-write meta-analyses (Hillocks, 1986) examined research on writing for students from third grade through college, looking both at the mode of presentation (e.g., role of the teacher, types of activities, specificity of objectives) and at the specific focus of the writing instruction (e.g., sentence-combining, peer feedback). Of the modes of presentation compared, the environmental presentation mode had the largest effect on student learning (ES = 0.44). In this mode, students were cooperatively engaged in the writing process to meet specific writing goals. The specific writing intervention with the largest effect size was inquiry activities, which involve developing ideas and content for writing by analyzing data (ES = 0.56). Student evaluation of writing (ES = 0.36), and sentence-combining instruction (ES = 0.35) also obtained moderate effect sizes.

Two meta-analyses were conducted to examine the impact of word processing on writing skills. First, Bangert-Drowns (1993) examined experimental and quasi-experimental studies in which two groups of students received identical writing instruction, but one group was allowed to use word processors to write. A small, positive effect (ES = 0.27) was found for the use of word processing for writing. Interestingly, the only study feature that had a statistically significant relationship with effect size was writing ability. That is, weak writers benefited more than average or strong writers from the use of word processors (ES = 0.49). Additionally, weak writers showed less variance in writing ability after instruction that included word processing, and the duration of this
instruction did not impact quality of writing. Taken together, these findings led Bangert-
Drowns to conclude that word processing for writing had a motivational impact on weak
writers.

A second meta-analysis of research on the use of technology for writing was
conducted by Goldberg et al. (2002). This meta-analysis focused on primary research
studies that were conducted with students in kindergarten through 12th grade, from 1992
to 2002. Findings indicated a moderate effect from use of word processing on both the
quantity (ES = 0.50) and quality (ES = 0.41) of students’ writing. There were no
significant differences in effect size found based on writing ability.

Two recent meta-analysis conducted by Graham and Perin (2007a, 2007b)
focused on writing interventions for adolescents (grades 4 through 12). The first
(Graham & Perin, 2007a) included only experimental and quasi-experimental studies,
while the second (Graham & Perin, 2007b) added single-subject and qualitative studies to
the previously reported findings. In each, the authors evaluated the effectiveness of
various instructional approaches, providing both weighted and unweighted mean effect
sizes. (Weighted mean effects are reported here.) First, the process writing approach
was moderately effective (ES = 0.32), although results were quite variable. Professional
development improved the effectiveness of the process writing approach (ES = 0.46) and
reduced variability of results. The effectiveness of explicit teaching approaches varied
based on the content of instruction. Most relevant to the current study, SRSD strategy
was very effective (ES = 1.14). The effectiveness of methods for scaffolding students’
writing also varied greatly. Peer assistance was one effective scaffolding strategy (ES =
0.75). As in previous meta-analyses, word processing had a moderate effect (ES = .55; PND = 77%, for writing quality measures) on students’ writing.

Finally, two general education meta-analyses have been conducted for writing-to-learn strategies. Bangert-Drowns et al. (2004) found a small, positive effect (ES = 0.26) of writing-to-learn on academic achievement. Using metacognitive prompts and increasing the length of the intervention improved effectiveness, while longer writing assignments decreased effectiveness. Additionally, smaller effects were found for students in middle school (sixth through eighth grades).

Recently, Graham and Hebert (2010) analyzed research on writing to improve reading. They found that having students write about what they read improves reading, particularly writing personal responses to reading (ES = 0.77). Also noteworthy, teaching spelling skills has a large effect on improving reading skills (ES = 0.68 to 0.79). Lastly, simply increasing the amount of time students write has a small effect on improving reading skills (ES = 0.30).

**Meta-analyses of writing research with students with learning disabilities.** No meta-analyses of writing research for students with EBD have been conducted to date. However, findings from five meta-analyses of writing instruction for students with LD (Gersten & Baker, 2001; Graham, 2006; Graham & Harris, 2003; Mason & Graham, 2008; Rogers & Graham, 2008) can provide insight into effective writing interventions for students with mild disabilities. Students with LD make up the largest group of students in special education. As such, much of the research on instructional strategies in the field of special education focuses on this population. As Rock, Fessler, and Church
(1997) pointed out, the academic characteristics of students with EBD are similar to those of students with LD. For example, both students with LD and students with EBD demonstrate deficits in executive functioning, inattentiveness, and hyperactivity. In the absence of research on students with EBD, information on students with LD is a useful starting place for identifying effective strategies that may work with students with EBD.

The findings of meta-analyses of writing research for students with LD have been relatively consistent and generally mirror the findings of meta-analyses of research involving students without disabilities. In their meta-analysis of research on expressive writing for students with LD, Gersten and Baker (2001) reviewed 13 studies focused on students in first through ninth grades. They found an overall strong, positive effect for writing instruction for students with LD (ES = 0.81). Writing interventions had a large impact on students’ inclusion of text structural elements in their writing (ES = 1.11). Instruction had a moderate effect on students’ metacognitive understanding of the writing process (ES = 0.64), and self-efficacy for writing (ES = 0.61). Instruction had little effect on students’ attitudes toward writing tasks (ES = 0.40). The authors concluded that research supports (a) instruction in the writing process, (b) awareness of text structures, and (c) providing feedback to students about their writing.

Two meta-analyses reviewed several programs of writing research for students with LD (Mason & Graham, 2008; Rogers & Graham, 2008). Both reviews found strategy instruction to be the most effective method for teaching writing, as measured by effect sizes. For example, Rogers and Graham (2008) conducted a meta-analysis of single-subject studies in writing for students in grades 1 through 12. They analyzed
primary studies that focused on strategy instruction for planning and drafting, editing, and paragraph construction, as well as those with a focus on teaching grammar, goal setting, word processing, reinforcement, pre-writing activities, instruction in sentence construction, and self-monitoring. The instructional strategy that included the most single-subject research studies was strategy instruction for planning and drafting (25 of the 88 studies in the meta-analysis). Percent of non-overlapping data (PND) effect sizes were uniformly high for this intervention: PND = 96%, for structural elements included in essays; PND = 85%, for generalization to other writing genres; PND = 91%, for productivity; and PND = 99%, for quality. Similarly, strategy instruction for editing found a PND = 84%, for errors corrected; and strategy instruction for paragraph construction had a PND = 97%, for structural elements included.

In contrast to the high effect sizes found for strategy instruction, other writing approaches were less effective (Rogers & Graham, 2008). For example, for instruction in goal setting for productivity, the effect size across seven studies was moderate (PND = 79%). Similarly, the effect size for word processing was moderate (PND = 70%). For instruction in pre-writing activities, the effect size was even lower (PND = 52%, for quality). The effect size for self-monitoring instruction was moderate (PND = 51%). Only the effect sizes for reinforcement (PND = 96%), sentence construction (PND = 86%) and teaching grammar (PND = 83%) were as high as strategy instruction.

Mason and Graham (2008) also analyzed writing research across programs of instruction, this time focusing on adolescents with LD in grades 4 through 12. Both group and single-subject studies were included in the analysis. Results of group studies
are reported as effect sizes (ES), while the results of single-subject studies are reported as PND. Again, the instructional programs that focused on strategic instruction outperformed other writing interventions. The strategic instruction model had effect sizes in the effective range (PND = 83% to 100%, ES = 1.69), as did cognitive strategy instruction for writing (ES = 0.93), interactive dialogues (ES = 2.51), and self-regulated strategy development (PND = 92% to 100%, ES = 0.72 to 1.32). Other approaches to writing instruction obtained more moderate effects overall, including goal setting (ES = 0.76) and using computers (ES = 0.79).

In 2006, Graham conducted a meta-analysis focused exclusively on strategy instruction. He found the overall effect size for strategy instruction to be in the very high range (ES = 1.15). Results were high across measures of quality, text structure elements, length, and revision. The findings held for (a) type of student—whether LD, at risk, average achieving, or high achieving; (b) across grades 1 through 12; (c) genres (e.g., persuasive writing, story writing); and (d) strategy taught. Students who were taught writing through strategy instruction were also able to generalize their skills to other writing tasks, and maintained their skill gains over time. Of the types of strategy instruction analyzed, SRSD had higher effect sizes in group experimental studies than did other approaches, but this was not the case in single-subject studies. Graham and Harris (2003) conducted a meta-analysis of SRSD research, the results of which are included in the discussion that follows.

The results of meta-analyses of writing instruction for students with LD clearly point to strategy instruction as the most powerful method for improving students’ skills.
Several programs of research over the last 30 years use strategy instruction to improve the writing skills of struggling learners. These programs of research are briefly described below.

**Strategy Instruction**

Strategy instruction focuses on improving students’ written products by teaching them the processes involved in writing, and scaffolding their learning along the way. The progression of strategy instruction moves from teacher modeling, to guided practice, to independent practice. Several strategy instruction models have been developed for use with writers who have LD. For example, the *interactive dialogue approach* developed by Bernice Wong utilizes verbal teacher modeling, as well as teacher-student and student-student conferencing throughout the writing process, to promote improved writing skills for students with LD (Wong, Butler, Ficzere, & Kuperis, 1996). Englert (2009) developed the *Cognitive Strategy Instruction in Writing* approach to teach writing to students with LD. This approach focuses on teaching students text structures, and employs teacher modeling and scaffolds, such as graphic organizers. As with the interactive dialogue approach, research on Cognitive Strategy Instruction in Writing found that student learning was greatly enhanced by student-student collaboration during the writing process.

Schumaker and Deshler (2009) took a different approach to strategy instruction in writing. Instead of focusing on the writing process as a whole (planning, translating, and revising), the researchers developed discrete strategies for teaching skills, such as sentence construction, paragraph writing, and editing. As with other strategy instruction
interventions, the **Strategic Instruction Model (SIM)** writing strategies follow a sequence of developing background knowledge, teacher modeling, guided practice, and independent practice to support student learning.

Self-regulated Strategy Development (SRSD) is another instructional approach to writing that has over 25 years of research supporting its effectiveness with struggling writers (Graham & Harris, 2003). Developed by Graham and Harris (2009), SRSD combines strategy instruction with instruction in self-regulation. It has been shown to be effective for writers of all ability levels, from elementary to secondary grades. All writers, with varying degrees of success, are able to maintain their gains over time and generalize their skills to other genres. It is effective for both narrative and expository texts (Graham & Harris, 2003). A recent meta-analysis by Baker, Chard, Ketterlin-Geller, Apichatabutra, and Doabler (2009) found that SRSD met the criteria for an evidence-based practice (Gersten et al., 2005; Horner et al., 2005) and recommended its use for learners with special needs.

**Self-regulated strategy development.** There are six stages of SRSD instruction (Graham & Harris, 2009). The first stage, **Develop Background Knowledge**, addresses pre-skills needed to learn and apply the writing strategy and self-regulation procedures. In the second stage, **Discuss It**, students are introduced to the writing strategy and make a commitment to learn it. In **Model It**, the third stage, the teacher models the strategy while thinking aloud and students develop self-statements that they can use for writing. During the fourth stage, **Memorize It**, students memorize the strategy’s mnemonic and their self-statements. **Support It**, the fifth stage, provides guided practice for students to implement
the strategy and to perform the self-regulatory skills of goal setting, self-instruction, and self-monitoring. Finally, *Independent Performance* is the sixth stage of the SRSD model. At this stage, students use the writing strategy and self-regulation skills to write independently.

De La Paz (2007) reviewed studies that contrasted elements of strategy instruction to determine which components were most important for improving students’ writing skills. In her review of 12 studies, the author determined that there was a modest gain in writing performance when self-regulation components were added to strategy instruction, as SRSD does. Additionally, peer support appeared to aid transfer and generalization of writing skills. Supporting students with the mechanical demands of writing was also beneficial. Finally, inconclusive results were found across studies that sought to increase motivation to improve writing performance.

**Self-regulated strategy development and students with EBD.** While there is a wealth of research establishing SRSD as an effective, research-based instructional strategy for students with LD (Baker et al., 2009), less is known about its effectiveness with students with EBD. A number of recently published studies have begun to point to the effectiveness of SRSD in improving the writing skills of students with EBD (Lane et al., 2008; Little et al., 2010; Mason, Kubina, Valasa, & Mong Cramer, 2010; Mason & Shriner, 2008; Mastropieri et al., 2009; 2010).

Mason and Shriner (2008) studied the effects of instruction in the SRSD persuasive writing strategy, POW+TREE, on the writing ability of sixth second through fifth grade students with EBD. POW+TREE stands for: Pick your idea, Organize your
notes, and Write and say more; plus Topic, Reasons, Ending, and Examine (i.e., the parts of a persuasive essay). After 11 to 13 individual, 30-minute, instructional sessions, five of the six students increased the number of persuasive essay parts, the total number of words, and the number of transition words included in their essays. Additionally, they improved the overall quality of their writing. Maintenance and generalization scores varied across students, with the authors concluding that these variations were due to individual student behaviors rather than skill levels.

Lane et al. (2008) conducted a study assessing the effectiveness of SRSD, in combination with schoolwide positive behavioral supports, for improving the writing skills of six second-grade students at risk for emotional and behavioral problems who were also below average writers. The POW+WWW, What = 2, How = 2 strategy was used and represents:

- **POW** = Pick my idea, Organize my notes, Write and say more.
- **+WWW** = Who is the main character? When does the story take place? Where does the story take place?
- **What** = 2 What does the main character do or want to do? What happens when the main character tries to do it; what happens with the other characters?
- **How** = 2 How does the story end? How does the main character feel; how do the other characters feel?

After 10 to 15, 30-minute instructional sessions, all students’ essays improved in terms of story elements included, length, and quality. All students maintained gains over
baseline at maintenance testing. Furthermore, students and teachers both rated the goals, procedures, and outcomes of the SRSD instruction favorably.

In a follow-up study, Little et al. (2010) taught 13 second-grade students with behavioral and writing problems the POW+TREE strategy for persuasive writing within the context of a schoolwide positive behavioral support system. In this study, students were grouped according to whether they exhibited internalizing or externalizing behaviors. Additionally, three students in the study had below-average IQ scores. Again, results showed that students increased in the number of text structural elements, number of words, and overall quality from baseline to postinstruction. They were able to maintain their gains over baseline at maintenance testing. There were no differences in performance for students with internalizing and externalizing behaviors.

Mastropieri et al. (2010) also used the POW+TREE strategy to teach students with EBD to write persuasive essays, this time focusing on middle school students. In this design experiment, 10 eighth-graders in a public day school were taught to write persuasive essays over 55, 30-minute instructional sessions. Throughout the course of the study, the researchers modified SRSD materials and instructional procedures that had been used in previous studies with students with LD, to meet the needs of students with severe EBD. As a result of this intensive instruction, students increased the length, number of essay parts, number of transition words, and overall quality of their persuasive essays. Additionally, their fluency scores improved. Students were able to maintain gains over baseline, 12 weeks after posttesting.
Lastly, two studies have been conducted using SRSD to teach middle school students with EBD to write fluently. First, in a follow-up to the previous study, Mastropieri et al. (2009) taught 12 eighth-graders with severe EBD the POW+TREE strategy for persuasive writing. In the second instructional phase of the study, students were taught to apply the POW+TREE strategy to write one-paragraph essays in 10 minutes. Students improved the length, number of paragraphs, number of essay parts, number of transition words, and holistic quality of their essays, following instruction in SRSD. After fluency instruction, students’ scores decreased somewhat, but were still significantly higher than their baseline scores. They were able to maintain above-baseline scores at maintenance testing.

In their study of the effects of fluency instruction on the writing of students with EBD, Mason et al (2010) used POW+TREE to teach five middle-school students a quick-write strategy. After a relatively short intervention period of six to seven 30-minute sessions, all students showed improvement in their ability to write persuasive essays within 10 minutes. Quality scores improved for all students. However, the total number of essay parts did not increase significantly, and the total number of words actually decreased for four of the five students. The researchers noted that all students were on-grade level writers when the intervention began, but they were highly inconsistent in their performance. Following the quick-write instruction, their performance stabilized markedly. This consistency in writing performance improved the overall quality of their writing.
In spite of the promising results of research on teaching SRSD to students with EBD, studies conducted to date have only addressed two aspects of the writing process, as articulated by Hayes and Flowers (1980)—planning and translating. Only one study has been conducted teaching revising skills to students with emotional and behavioral disorders (Morris Kindzierski, 2009), and it did not specifically teach students to write essays using strategy instruction. Nearly all of the studies that have been conducted on revision focus on students with LD, and most were conducted more than 15 year ago.

Revision Research

Revision means making any changes at any point in the writing process. It involves identifying discrepancies between intended and instantiated text, deciding what could or should be changed in the text and how to make desired changes, and operating, that is, making the desired changes. Changes may or may not affect meaning of the text, and they may be major or minor. Also, changes may be made in the writer's mind (a) before being instantiated in written text, (b) at the time text is first written, and/or (c) after text is first written. (Fitzgerald, 1987, p. 484)

In the only review of research on revision to date, Fitzgerald (1987) provided an overview on what is known about the revision skills of beginning writers and more advanced writers. Beginning writers, in contrast to more advanced writers, do not revise much. Younger students do very little revising without support, whether it comes from a teacher or a peer. This is true for many older writers as well. Revising behaviors tend to change as writers become more competent—more skilled writers make more revisions, and do more revising while writing rather than waiting until they have completed a first
draft. For all writers—beginning and skilled—most revisions are surface-level changes. However, more skilled writers make a larger proportion of meaning-change revisions than beginning writers.

The research on interventions for revision was limited at the time of Fitzgerald’s (1987) review. Results of the available studies indicated that revision did improve the quality of writing for more skilled writers, but did not necessarily improve the quality of writing for less skilled writers. On the whole, procedural support, direct instruction, and feedback from others tended to produce positive results. Simply giving directions to revise did not produce consistent positive results.

A study of 10th grade writers supports the importance of teaching revision skills to students who are less skilled writers (Kieft, Rijlaarsdam, & Galbraith, 2007). In this study, 113 high school students were assigned to either a planning strategy condition or a revising strategy condition. For more competent writers, participation in the planning strategy condition resulted in larger gains in writing quality. Less skilled writers, on the other hand, benefitted more from instruction in the revising strategy.

Two basic approaches have been used in the research on revision instruction for students with disabilities: (a) independent revision strategies (Graham & MacArthur, 1988; Reynolds, Hill, Swassing, & Ward, 1988) and (b) peer-revision strategies (MacArthur, Schwartz, & Graham, 1991; Stoddard & MacArthur, 1993; Wong et al., 1996). One recent study (Morris Kindzierski, 2009) examined which method— independent or peer-revision—produced more meaningful revisions for students with EBD. In an alternating treatment, single subject study, eight students with EBD wrote
both narrative and descriptive essays, and then revised the essays using an independent
checklist or a peer-revision checklist, both of which contained identical questions. No
instruction was given on how to complete the revision checklists. Results indicated that
students made fewer conventional errors, used fewer words, and had more repetition of
words when revising independently. In the peer revision condition, students used more
personal references, wrote longer drafts, and their writing was more organized.

**Independent revision strategies.** A number of different strategies have been
used to teach students to revise independently. Scardamalia and Bereiter (1986)
suggested that teaching executive procedures for revising could benefit struggling
writers. In a series of two studies, Graham and colleagues (Graham, 1997; Graham,
MacArthur, & Schwartz, 1995) attempted to test this theory by teaching two different
strategies for revising to students with LD who were also struggling writers. In one
study (Graham, 1997), fifth and sixth grade students (six boys, six girls) were taught the
compare, diagnose, operate (CDO) strategy developed by Scardamalia and Bereiter
(1983) to provide an executive processing framework for revising. The study consisted
of two conditions, with students writing and revising one essay in each condition. The
first condition was the typical revision condition where students were simply told to
revise the essay they had written. The second condition was the CDO condition in which
students were taught the strategy and applied it to revise an essay they had written. The
CDO strategy required students to read each sentence of the essay and evaluate the
sentence as to whether it met their goals or not. Students were given seven note cards
with evaluative statements on them, and chose which card fit for that sentence. Then, if
the evaluation indicated that the sentence needed to be revised, students chose from five options (also written on note cards) for revising the sentence (e.g., add information). The final step, operate, required the student to make the change that was needed. Students performed this CDO procedure on each sentence in the essay. The teacher first modeled the strategy for the students, then they applied the CDO revision strategy to their own essay.

Students’ essays were analyzed to provide information about the revisions that were made. Revisions were scored along several dimensions: (a) level (surface, word, phrase, or t-unit); (b) type (addition, deletion, substitution-rearrangement); (c) impact on meaning (meaning-changing or meaning-preserving); and (d) quality (Did the revision make the paper better or worse, or was there no change in quality?). The total number of revisions per 100 words was also recorded. Additionally, essays were scored for length, as measured by total number of words, and overall quality. Two quality measures were used. First, a holistic quality score was given. Second, a score of +1, 0, or -1 was given to indicate whether the revisions improved the paper from first draft to second draft (+1 indicated a higher quality second draft, -1 indicated a higher quality first draft, and 0 indicated no change in quality). Lastly, student interviews about the strategy were also conducted.

Results of the student interviews indicated that 10 of the 12 students thought that the CDO procedure made revising easier, and that it made their papers better. Eleven of 12 participants recommended that the strategy be taught to other students. On the essay scoring measures, no differences were seen between conditions on length or quality.
Furthermore, there were no differences or changes in quality from first to second draft across conditions. Some differences were seen in types of revisions undertaken in each condition. The number of surface-level revisions made was not impacted by revision condition, so only the non-surface revisions were analyzed. The CDO procedure increased the number of meaning-preserving revisions, but did not impact the number of meaning-changing revisions. When using the CDO strategy, students’ non-surface revisions were rated as higher in quality and lower in quality, but there was no difference in the number of revisions rated as no change in quality across conditions. Most revisions were substitutions (40%), additions (33%), and deletions (27%). In analyzing the evaluative statements students chose during the CDO procedure, almost all of the evaluative statements were based on the writer’s personal reactions to the text. Only four revisions were selected based on the reader’s perceptions of text. Nearly half of the revisions that students made to their papers during the CDO conditions were not prompted by the CDO procedure.

While students in the CDO condition did not outperform students in the typical revision condition on every measure, students with LD did appear to benefit from the assistance with executive procedures during the writing process. However, the lack of impact on text quality suggested that a different type of instruction was also needed for struggling writers to effectively revise their work. To that end, Graham, MacArthur, and Schwartz (1995) investigated the impact of goal setting on the revision behaviors of struggling writers with LD. The study included 67 students in grades four through six who were randomly assigned to one of three revision goal conditions. The first group
was given a general goal to revise their paper to make it better. The second group was given a specific goal to add three additional things to their essays to make them better. The third group was given the same goal of adding three things to their essays to make them better, and was given a procedure for carrying out the task. The procedure required students to write on a separate sheet of paper at least five things they could add to their story, and then put a check mark next to the three they liked best.

On the first day, students wrote a story in response to a given prompt. Two to four days later, students revised their essays according to the revision goal they had been assigned (i.e., general goal, add information goal, or add information goal plus procedural facilitation). Essays were scored for holistic quality, length, and revisions. All changes between drafts were counted as revisions, and were scored along four dimensions described in the previous study (Graham, 1997). Results showed that there were differences between the general revision goal condition and the two add information goal conditions. However, there was no difference between the add information goal condition and the add information goal with procedural support condition. That is, the procedural facilitation strategy did not additionally benefit students who received a goal to add three new things to their essays.

More specifically, students in the two information-adding conditions outperformed the students in the general goal condition on quality and revision measures. There were no differences between groups in the length of their essays, when the length of the first draft was used as the covariate. On revision measures, across all groups, students made an average of 23 revisions per 100 words. Unlike the Graham (1997)
study, in this study there were no differences across groups for meaning-preserving revisions, but there was a significant difference across groups on meaning-change revisions. Therefore, subsequent analyses are reported only for meaning-change revisions. Students in the two information-adding conditions had more meaning-change revisions than students in the general goal condition. For students in the general goal condition, most revisions involved surface-level mechanics (72%). For students in the adding-information conditions, only 50% of their revisions focused on mechanics. Most meaning-change revisions were additions, with students in the two information-adding conditions making more additions than students in the general goal condition. Seventy-five percent of revisions were rated as making the text better, with no significant differences across groups when measured by parametric measures. Using nonparametric measures, however, students in the two information-adding conditions had more revisions that made the text better than students in the general goal condition. All in all, giving students a specific goal for the amount of information to add to a story when revising, improved their ability to revise. The addition of a procedure to facilitate that goal did not make a significant difference for the students with LD who participated in this study.

Taken together with the results of the previous study, Graham and his colleagues concluded that simply providing executive procedures to students with LD was not enough to help them become more proficient revisers. Instead, they needed self-regulation strategies, such a goal setting, along with executive procedures, to learn how to revise more effectively.
A number of other strategies have been used to teach students how to revise their papers independently, without the aid of a teacher or peer. In one example, Fitzgerald and Markham (1987) taught 30 average sixth-grade writers specific types of revisions (e.g., additions, deletions). The three-day instructional cycle began with the teacher modeling the particular revision strategy that was the focus of instruction that week. The next day, students practiced revising a prepared text with peers, using the targeted revision strategy. Students also independently wrote a brief story on Day 2 and had a conference with the teacher about possible revisions that could be made to the written piece. On the third day of instruction, students independently revised their story.

In this posttest-only experimental design, students’ revisions were evaluated at four stages of the writing process: (a) in process during the first draft (by comparing what was written to pre-writing plans); (b) new revisions marked on the first draft; (c) revisions between marks on the first draft and the final draft (any revisions in the final draft that were not previously marked on the first draft); and (d) in-process revisions on the final draft (indicated by crossed out or inserted words, etc.). Results indicated that students in the experimental group identified more possible revisions per 100 words at the four stages of the writing process than students in the control group, although there was no difference in the specificity of the intended revisions between the control and experimental groups. At posttesting, students in the experimental group made twice as many meaning-change revisions as surface-change revisions, while the control group showed no difference in type of revision. Additionally, the students in the experimental group made more revisions per 100 words than those in the control group, with most
revisions made during the third stage of the writing process. Finally, the quality of the essays improved from stage one to stage four for students in the experimental group, while the quality was unchanged from first to final draft for students in the control condition. These results suggest that students who received instruction in types of revision increased the number and quality of revisions, and these revisions improved the overall quality of their essays.

In another study, Reynolds, Hill, Swassing, and Ward (1988) conducted a quasi-experimental study of 54 sixth-, seventh-, and eighth-grade students to determine the relative effectiveness of two types of revision strategies—one for content revisions and one for mechanical revisions—and the effect order of instruction had on the types of revisions students made in their essays. Both experimental and control groups received the same instruction in planning and drafting. Then, the first experimental group received instruction in the evaluative and declarative phrases strategy for content revision, and received instruction in COPS for mechanical revisions. The second experimental group received instruction in the reverse order: COPS first, evaluative and declarative phrases second. The control group did not receive instruction in revision.

After each phase of revision instruction, students revised two paragraph-long essays they had written during baseline testing. Posttest scores were analyzed for content and mechanics using Deiderich’s (1974) analytic scale. The scale rated each of the following eight factors on a 5-point scale: areas of general merit, including (a) ideas, organization, wording, and flavor; and (b) mechanics, including usage, punctuation, spelling, and handwriting. Both experimental groups increased scores on measures of
mechanical revisions, but did not increase scores on content revisions. The control group
did not increase its scores on content or mechanical revisions.

Graham and MacArthur (1988) also taught students with LD to use an
independent revision strategy. In this multiple baseline study, three fifth- and sixth-
graders with LD were taught a six-step revision strategy:

1. Read your essay.
2. Find the sentence that tells what you believe. Is it clear?
3. Add two reasons why you believe it.
4. SCAN each sentence.
   
   S = Does it make sense? C = Is it connected to my beliefs?
   
   A = Can I add more? N = Note errors.

(5) Make changes on the computer.

(6) Re-read the essay and make final changes.

Students were given essay probes at baseline and postinstruction. Both the first
and final drafts of each essay were scored according to a number of factors. Revisions
(any changes between drafts) were scored according to four factors: (a) the total number
of revisions; (b) the level of change of each revision (surface change or non-surface
change); (c) all non-surface revisions were coded by type (addition, deletion, substitution,
rearrangement); and (d) all revisions were coded as meaning-preserving or meaning-
changing. In this and subsequent studies by MacArthur and colleagues (MacArthur et al.,
1991; Stoddard & MacArthur, 1993), surface-change revisions included “changes in
spelling, capitalization, morphological changes for tense of number, changes in format,
and abbreviations” (p. 139). All other changes were considered non-surface level revisions. In addition to scoring of revisions, each essay was also scored for total number of words written; total number of spelling, capitalization and punctuation errors; overall quality, and overall change in quality from draft one to draft two; and the purposes of the revisions. The purposes of the revisions matched the six stages of the revision strategy and were scored to measure the degree to which each student implemented the revision strategy. Finally, students were given a self-efficacy for writing scale.

Results showed that students increased the total number of revisions following instruction in the revision strategy. While more revisions at baseline were surface-level revisions, following instruction, two of the three students changed to having more non-surface-level revisions. By far, the majority of revisions were additions (84%). This was true at both baseline and posttesting. The number of revisions that changed the meaning of text increased as a result of instruction. Students also showed increases on other essay measures. The total number of words written on final drafts increased for all participants. Two participants also increased the total number of words on the first draft of their essays. The number of spelling, capitalization, and punctuation errors did not decrease between first and second drafts at baseline or following instruction. However, two of the three students had fewer errors on the first drafts of their essays postinstruction. On the quality measure, there was no change from first to second draft at baseline testing, but at posttesting, second drafts were better than first drafts. Additionally, all students showed increased scores on the self-efficacy for writing measure following training in the
revision strategy. Finally, analysis of the purpose of revision indicated that students did
follow the steps of the revision strategy.

One recent study attempted to teach an individual revision strategy within the
context of the writing process (Saddler & Assaro, 2007). This multiple baseline study
included six second-grade students (three boys, three girls) with LD who were also
struggling writers. The students were taught the self-regulated strategy development
approach to story writing (POW+WWW What = 2, How = 2), as well as a self-
questioning technique for revision. To teach the revision strategy, the teacher did a think
_aloud_ as she revised a story. She asked herself questions to identify problems, model
corrections, and reread the revised text. The method was based on Wong et al. (1996)
think-aloud model, which is described later in this chapter.

Students wrote three to five baseline stories in response to picture prompts, and
three posttest stories. Essays were scored along five dimensions: (a) number of story
elements (e.g., characters, emotions); (b) amount of planning time; (c) quality of first
draft, which was a holistic score; (d) quality of the revised draft, scored from -2 to +2
when compared to the first draft; and (e) revisions. Revisions were coded by type
(capitalization, punctuation, spelling, addition, deletion, substitution) and level (surface,
morphological, or non-surface). Results showed that all students increased the number of
story elements and the quality of their first drafts. Planning time increased for all
students, although it was still very limited. The number of revisions for each student
increased from zero at baseline to three or four at posttest. Most revisions were surface-
level additions or word changes, but they were generally rated as improving the quality of the text. Overall, the revision strategy had a modest impact on student revision.

**Peer-revision strategies.** Most research studies teaching revision to students with LD utilized peer revision. Some studies measure the outcomes of instruction in revision, specifically (MacArthur et al., 1991; Stoddard & MacArthur, 1993), while others look at overall writing outcomes (De La Paz & Graham, 2002; Englert et al., 1991). Englert and colleagues (1991), for example, included peer revision within the Cognitive Strategy Instruction in Writing approach, focusing on overall writing performance, rather than revisions, in particular. The Englert study included 183 fourth- and fifth-grade students, 55 of whom were identified as students with LD. The experimental and control groups included students identified as LD, students who were low achieving in writing, and students who were high-achieving writers. The intervention included explicit instruction in the entire writing process, following the mnemonic POWER: Plan, Organize, Write, Edit/editor, and Revise. A series of think sheets guided students through each step of the writing process. The editor think sheet, for example, prompted students to: (a) place stars (*) next to the parts of the text that they liked and question marks (?) next to the parts that were confusing; (b) rate whether text structure features were present; and (c) complete a guide for making a revision plan. The revision think sheet, then, required students to reflect on their editing plans by listing the suggestions generated during peer-editing conferences and deciding which ones to implement. Students consulted with peers at several points in the writing process, including planning and editing.
Students were tested both pre- and postintervention on four measures. First, a test of metacognitive knowledge for writing was given. Next, students composed two essays in response to given prompts: one explanation essay, and one compare-contrast. Then, students were given a near-transfer measure that asked them to write an essay on the topic of their choice. Finally, a far-transfer task of reading comprehension was given. The three essays (explanation, compare-contrast, and self-selected) were scored on four dimensions: (a) inclusion of text structure elements, (b) holistic quality, (c) productivity (i.e., the number of ideas contained in the paper), and (d) reader sensitivity.

Results indicated that experimental students made gains on the metacognitive test while control students did not. Additionally, experimental students showed an increased use of text structure elements, greater sensitivity to audience, and better holistic quality of their essays following instruction. High-achieving students scored significantly higher than the low-achieving and LD groups, but there was no difference in the performance of the low-achieving and LD groups. Effects across groups were consistent, except in the area of audience sensitivity. The higher achieving group showed higher levels of audience sensitivity. The low-achieving and LD groups were higher than control students on this measure; control students’ scores on the audience sensitivity measure actually decreased over time. On the near-transfer task of writing an essay to a self-selected topic, all groups of experimental students performed well, but control students’ performance decreased. One of the most exciting conclusions from the study was that, while students with LD showed big differences from low-achieving and high-achieving students at pretesting, they did not demonstrate significant differences from their peers at
posttesting. This indicates that the writing intervention eliminated the differences between students with LD and their nondisabled peers.

De La Paz and Graham (2002) also focused on overall writing outcomes in a study that sought to teach SRSD within a comprehensive writing program and included instruction in planning, translating, and revising. In this quasi-experimental study, the PLAN and WRITE strategy was used to teach typically achieving middle school students how to write essays on statewide high-stakes tests. The strategy included steps for (a) evaluating the essay prompt, (b) planning, (c) goal setting, (d) writing a thesis statement, (e) including transition words, and (f) using high-level vocabulary. The revision portion of the program involved both peer revision and individual teacher conferences. Peer-revision conferences followed a checklist focused on the following elements: (a) ideas and development; (b) organization, unity, and coherence; (c) vocabulary; and (d) sentence structure, grammar, and usage. Teacher-student conferences included discussion about any difficulties the student was having and writing goals.

A total of 58 seventh- and eighth-graders (20 boys, 38 girls) participated in the study. Instruction was provided to 10 intact classes by five classroom teachers. A sample of students from these classrooms was selected for inclusion in the study. A stratified random sampling procedure was used to ensure that students sampled from each classroom were representative of the students taught by that teacher. PLAN and WRITE instruction was conducted four days a week for six weeks. Students in the control group received the same amount of writing instruction using the same essay prompts and the same number of essays. The “traditional” instruction the control group received included
instruction in the five-paragraph essay structure, organization and brainstorming, grammar, and spelling.

Pretesting, posttesting, and one-month maintenance testing involved having students write essays given prompts similar to those used on statewide high-stakes tests. Essays were scored along four dimensions. First, planning was scored on a five-point scale where one indicated that the student did not plan and five represented a complete plan. Plans were assessed for the degree of elaboration and accuracy. Second, essays were scored for length, as measured by total number of words. Third, vocabulary was assessed by counting the number of words included in the essay that had seven or more letters. Fourth, overall quality was rated using an eight-point scale focused on the quality of ideas, structure of text, word choice, sentence structure, and mechanics. Teachers were also interviewed about the effectiveness of the intervention, recommendations, and other comments.

Results showed that students who received strategy instruction outperformed the control group on all essay measures at posttest and were able to maintain those gains at maintenance testing. While both the strategy group and the control group’s planning scores improved from pre- to posttesting, the strategy group’s scores improved more (posttest ES = 1.17, maintenance ES = 1.04), with most students achieving a score of four or five on their plans. Students in the strategy group also had longer essays than students in the control group (posttest ES = 0.82, maintenance ES = 1.07), scored higher on vocabulary (posttest ES = 1.13, maintenance ES = 0.94), and outperformed the control group on overall quality (posttest ES = 1.71, maintenance ES = 0.74). These results
indicate that SRSD is a writing strategy that benefits all students, not just those with learning disabilities.

A number of studies investigated the effects of revision instruction, in particular, on the revising abilities of students with LD. MacArthur and colleagues (1991) for example, followed their study of an independent revision strategy (Graham and MacArthur, 1988), with two more studies of revision instruction for students with LD (MacArthur et al., 1991; Stoddard & MacArthur, 1993), this time using a peer-revision approach. The first of these (MacArthur et al., 1991), a quasi-experimental study, included 40 fourth-, fifth-, and sixth-graders with learning disabilities. The student-editor strategy required peers to meet together twice during the revision process. During the first meeting, student partners focused on substantive revisions. To do this, the student-editor listened while the author read his essay aloud, and then told the author what the paper was about and what she liked about it. Then, the student-editor reread the paper to herself and made notes about the revision questions: (a) Is there anything that is not clear? and (b) Could more details or information be added? The student-editor then discussed the suggestions with the author. After one student partner acted as the editor, the partners switched roles.

Following the first peer-revision meeting, both partners revised their essays independently, based on the feedback from their partner. The second partner meeting focused on mechanical revisions. During this meeting, each partner discussed the changes he or she had made to the first draft and edited their partner’s paper for complete
sentences, capitalization, punctuation, and spelling. Lastly, each partner made final revisions independently.

In this study, dependent measures included pre- and postinstruction writing samples, a pre- and postinstruction metacognitive interview that asked about the criteria of good writing and the writing process, and a strategy use checklist to determine whether students had followed all steps of the peer-editing strategy. Each writing sample was given over three days. On the first day, students wrote a first draft in response to a given prompt. On the second day, students made revisions to their first draft using pencil and paper, and then made changes on the computer. On the third day, students met with a partner and gave suggestions, then made individual revisions on the computer. All three drafts were scored on a variety of measures: (a) total number of words; (b) total number of spelling, punctuation, and capitalization errors; (c) overall quality; (d) change in quality from first to final draft; and (e) revisions, including total number of revisions, level of revisions (i.e., surface or non-surface level), impact on meaning (i.e., meaning changing or non-meaning changing), and quality of revisions.

Results of the study indicated that all students in the experimental condition used the majority of the steps in the peer-revision strategy. As a result, students in the experimental group made more revisions, postinstruction, than students in the control group, even though experimental students made fewer revisions at baseline testing. The experimental group showed no difference between draft one and draft two (i.e., when revising independently) on non-surface level revision, but they had significantly more non-surface revisions on draft three after revising with a peer. Students made more
surface-level revisions on draft two. The experimental group outperformed the control
group on the quality measure (ES = 1.33). Furthermore, they showed a significant
decrease in the number of spelling and capitalization errors, but no difference in
punctuation errors following instruction. The metacognitive interview yielded few
differences between groups.

Stoddard and MacArthur (1993) used a similar format for teaching revision to six
seventh- and eighth-grade students with LD using a single-subject, multiple probe across
pairs design. In this study, students wrote personal narratives. The revision strategy had
instructional procedures similar to the six stages of self-regulated strategy instruction
(SRSD). As in MacArthur et al. (1991), the first peer-revision conference focused on
content-level revisions. First, the revision partner listened as the author read the essay
aloud, and then the partner said what he liked best. Next, the partner reread the essay to
himself and made notes in response to the revision questions. The revision questions
included

1. Parts – Does it have a good beginning, middle, and end?
2. Order – Does it follow a logical sequence?
3. Details – Where could more details be added?
4. Clarity – Is there any part that is hard to understand?

After discussing these issues with the author, the author revised her own essay
independently. The second peer conference, then, focused on mechanical errors.

Similar to the previous study, essays were scored for number of words; proportion
of spelling, capitalization, and punctuation errors; and the number and type of revisions.
All revisions were coded as being surface or non-surface level changes, as meaning preserving or meaning changing, and the quality of the change (better, no change, or worse). Additionally, two quality measures were used: overall quality of the final draft, and change in quality from first draft to final draft.

Results indicated that students increased non-surface level revisions from baseline to posttesting and maintained those gains (PND = 96%). Additionally, the quality of these non-surface revisions increased from first to final drafts (PND = 100%). Three of the six students increased the number of surface-level revisions from baseline to posttesting (PND = 58%), but the quality of these revisions did not improve. The overall quality of the essays increased substantially following instruction, and the quality of postinstruction essays improved from first to final drafts, indicating that students’ revisions helped make their writing better. Furthermore, the length of students’ essays increased from first to final drafts following instruction, while they did not increase following revision during the baseline phase. Spelling errors decreased following instruction, while capitalization and punctuation errors were unchanged. Finally, on a measure generalizing to paper-pencil writing, students made more revisions when handwriting than when typing.

A final study using peer revision with students with LD was conducted by Wong et al. (1996) using the interactive dialogue approach to write opinion essays. This quasi-experimental study included 28 eighth- and ninth-grade students classified as LD or low achieving in writing. In this approach, peers met during both the planning and the revision stages of the writing process. The focus of these conferences was on the
cogency of the author’s argument and the clarity of his writing. During the peer-revision conference, the role of the student-critic was to identify ambiguities in the partner’s writing, and ask the author for clarification. Each student also participated in a conference with the teacher to discuss any ambiguities and the cogency of the author’s arguments. Finally, the COPS strategy was used to correct mechanical errors.

Pre- and posttest essays were scored for clarity of writing and cogency of the writer’s argument. Two essays were given during pretesting, one essay was given during posttesting, and one essay was done during one-week maintenance testing. Three additional measures were given at pretest and maintenance: (a) an attitude toward writing scale, (b) a self-efficacy for writing scale, and (c) a metacognition for writing scale. Results indicated that the students in the experimental condition showed significant gains from pre- to posttest on both clarity (ES = 2.17) and cogency measures (ES = 2.74), and those gains were maintained at one-week maintenance testing. No pre-post differences were found on the attitude toward writing (ES = 0.12) and metacognition (ES = 0.61) questionnaires. There was a significant difference in self-efficacy for writing (ES = 0.70) for the interactive dialogue group. Students in the experimental condition outperformed control students on both clarity (ES = 2.55) and cogency (ES = 2.52). No pre-post differences were found on the attitude and metacognition measures for students in the experimental group. However, they did show a significant difference from pre- to posttest on the self-efficacy measure. At posttest, the interactive dialogue group outperformed the control group on both the self-efficacy for writing (ES = 0.73) and metacognition in writing (ES = 0.87) questionnaires, although the authors note that the
metacognition finding must be interpreted with caution because there was no difference in pre- and posttest metacognition scores for the experimental group. There were no differences between groups on the attitude toward writing questionnaire (ES = 0.43).

Quality Single-Subject Research Design

Horner et al. (2005) articulated a number of indicators of high quality single subject design. First, participants and setting must be thoroughly described to allow for replication by future researchers. Participant characteristics, as well as the process for selecting participants, must be described in detail. Additionally, the setting must be described with enough detail to allow for replication.

The second quality indicator refers to the dependent variable. That is, the dependent variable must be described in sufficient detail, operationally defined, and scored in a quantifiable way. In addition, the dependent variable must be measured repeatedly over time, and information about interrater reliability must be provided.

The third indicator of high quality single subject research design relates to the independent variable. The independent variable must be operationally defined. It also must be systematically applied by the researcher, and measures of fidelity of implementation must be provided.

The fourth indicator describes procedures for establishing a stable baseline of participant performance. Baseline must include repeated measures, and the pattern established at baseline must be sufficient to predict future performance without application of the independent variable. As with other high-quality research criteria, the baseline phase must be described sufficiently to allow for replication.
The fifth indicator requires demonstrations of experimental control during each phase of the study. To achieve this, three applications of the effect must be demonstrated at three different points in time. The design of the study must control for threats to internal validity.

The sixth indicator for high quality single subject research is related to external validity. That is, the experimental effect must be replicated across participants, settings, or materials. The seventh and final indicator is social validity. The dependent variable in the study must be socially important, and the change in the dependent variable as a result of the intervention must be significant. Additionally, the independent variable must be practical and cost-effective to implement. Ideally, the independent variable will be applied over time, in typical settings, by typical intervention agents.

Summary

Students with EBD demonstrate significant academic delays in reading, writing, and mathematics when compared to their peers without disabilities (Nelson et al., 2004; Reid et al., 2004; Trout et al., 2003). Language disorders, which disproportionately affect students with EBD, contribute to students’ poor performance and necessitate instruction focused on receptive and expressive language skills (Benner et al., 2002; Nelson et al., 2006; Wagner et al., 2005). Despite the great academic needs of students with EBD, little research has focused on effective academic interventions for this population of students (Mooney et al., 2003). Peer-mediated instruction is one intervention with some support indicating that it is an effective intervention for
improving the academic outcomes and, in some cases, classroom behaviors of students with EBD (Ryan et al., 2004; Spencer, 2006; Spencer et al., 2009).

Research on writing over the last 30 years has pointed to strategy instruction as the most effective method for improving the writing of students with mild disabilities (Mason & Graham, 2008; Rogers & Graham, 2008). While most of this research has focused on students with LD, the same strategies have shown promising results for students with EBD (Lane et al., 2008; Mason & Shriner, 2007). However, few of these studies have also included instruction in revision. For weak writers, instruction in revision is necessary for improving their writing outcomes (Kieft et al., 2007). The available research on revision suggests that instruction that begins with teacher modeling and progresses through guided practice and independent practice is effective for teaching students revising skills (Mills, 2011). Additionally, including feedback from peers during the revision process can improve students’ writing (Mills, 2011).

Developing instructional strategies to meet the academic needs of students with EBD, while at the same time addressing the behavioral issues that negatively impact their performance in the classroom, is critical for improving outcomes for students with EBD. Writing is one area where students with EBD demonstrate particular weakness and, if unaddressed, can have a lasting impact beyond school. Only a handful of research studies conducted to date, have focused specifically on the writing skills of students with EBD (e.g., Mason et al., 2010; Mastroianni et al., 2009; 2010). Only one of these studies has addressed instruction in revision—a key component of the writing process (Morris Kindzierski, 2009). Research on other students with mild disabilities has shown that
working with peers during the revision process greatly increases the effectiveness of instruction (Mills, 2011). Because peer-mediated instruction can have positive impacts on the academic success and behavior of students with EBD, instruction in peer-revision may be one way to meet both the academic and behavioral needs of these students.

Clearly, the need is great for broad and multiple investigations into effective, evidence-based practices that work for students with EBD. Further exploration of the strategies which have been investigated and found effective for other populations of students is a starting place for building this body of research and impacting outcomes for students with EBD.
3. METHODS

A single-subject multiple baseline, multiple probe across participants design (Kennedy, 2005) was used to determine the additive effects of instruction in peer revision on the persuasive writing of 10 middle-school students with emotional and behavioral disorders (EBD). The study consisted of six phases: (a) baseline testing phase, (b) self-regulated strategy development (SRSD) phase, (c) SRSD posttesting phase, (d) revision instructional phase, (e) revision posttesting, and (f) maintenance testing phase. This chapter describes the methods used in the current study, including setting, participants, dependent measures, materials, procedures, and scoring procedures. Additionally, the single-subject multiple baseline design is explained.

Setting

This study took place in a large suburban middle school along the East Coast. The school serves 1055 seventh and eighth grade students (51.5% male, 48.5% female). Students at the school are predominantly White (62.8%), followed by Asian or Pacific Islander (23.0%), other (5.8%), Hispanic (5.4%), and Black (2.9%). Ten percent of the students in the school are identified as limited English proficient; 8.4% receive free and reduced price meals.

Special education serves 13.1% of the school population. The special education program at the school serves students with a variety of mild disabilities in both team-
taught and self-contained classes. The school also utilizes a schoolwide positive behavioral intervention and support (PBIS) system, although it is not consistently implemented in the self-contained classes serving students with EBD. Rather, each teacher of a self-contained class for students with EBD selects his or her own classroom rules, and follows a structured series of consequences for infractions. This approach is taken because school staff feel that the students served in the self-contained program do not “buy into” the schoolwide positive behavior system (Special Education Teacher, personal communication, September 23, 2010).

Participants

The study began with 12 participants, but one male participant was dropped from the study after missing three weeks of school due to a death in the family. One female participant was dropped from the study after being transferred to another educational setting. Subsequently, 10 eighth-grade students participated in this study. Seven of these students (70%) were male and three (30%) were female. The average age of participants was 13 years 7 months (range = 13.1 to 14.3, SD = 5.34). Seven participants (70%) were identified as White, one participant (10%) was Hispanic, and two (20%) were multiracial. One student was identified as a student with limited English proficiency. Most participants ($n = 8$, 80%) were found eligible for special education services due to emotional disturbance (ED), with five of these (62.5%) having comorbid disabilities. Four students had both ED and a learning disability (LD), and one student had ED and other health impairment (OHI) due to a medical diagnosis of attention deficit/hyperactivity disorder (ADHD). Two participants (20%) received ED services
under the special education category of OHI. One of these students had both OHI and autism; the other received services provided by ED specialists, autism specialists, LD specialists, and physical disabilities (PD) specialists under his OHI designation.

Selection criteria. Two criteria were used to select participants for this study. First, students had to be receiving special education support to address identified emotional issues, in accordance with their individualized education program or IEP. Additionally, students had to be enrolled in one of two eighth-grade, self-contained, language arts classes designed to remediate deficits in reading and writing. It was during this class period that the writing intervention was implemented by the researcher.

Performance on standardized tests. Information about students’ ability and achievement on standardized tests was retrieved from their special education eligibility paperwork. Current ability test scores (i.e., tests administered within the last two years) were not available for all students. Of those with current scores ($n = 6, 60\%$), five students had taken the Wechsler Intelligence Scale for Children ($4^{th}$ ed., WISC IV; Wechsler, 2003) earning an average full scale IQ score of 99.6 (range = 79 - 120, $SD = 18.25$). One student took the Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003), earning a Composite Intelligence Index score of 109. Five students (50%) had current scores on standardized writing tests. Two students completed the Woodcock-Johnson III Tests of Achievement (W-J III; Woodcock, McGrew, & Mather, 2001). One student received a Broad Written Language standard score of 95; the other received a standard score of 65 on the same scale. Two students completed the Test of Written Language-3 (TOWL-3; Hammill & Larsen, 1996). One student received a
Written Language Quotient standard score of 76; the other received a standard score of 99. One student completed the Kaufman Test of Educational Achievement (2nd ed., KTEA-II; Kaufman & Kaufman, 2004), earning a Written Language Composite score of 89. Taken together, these scores indicate a diverse group in terms of intellectual ability, with students performing in the below average, low average, average, high average, and superior ranges. In terms of writing achievement, students performed in the low, low average, and average ranges.

Performance on state testing. Students in this school district participate in statewide high-stakes assessments at regular intervals throughout their public education. Relevant to this study, students were required to take the state assessment in writing when in the fifth grade. Nine of the 10 participants in this study participated in these tests. One student did not take the test because of an exemption for students with limited English proficiency. The test consists of two parts. The first part is multiple-choice, including 32 questions that address planning, composing, revising, and editing a variety of types of writing. On this portion of the test, participants in this study received an average score of 411.56 (range = 322 - 487, $SD = 62.09$). A score of 400 is considered passing, and a score of 500 is considered proficient.

On the second part of the test, students are presented with a prompt and asked to write an essay response. Students are scored on both their composition and written expression skills, and on usage and mechanics. A score of 30 or more on either measure indicates an area of strength for the student; a score below 30 indicates an area of need. On the written expression measure, participants scored an average of 30.22 (range = 22 -
38, $SD = 5.95$). On the mechanics measure, participants’ mean score was 31.78 (range = 22 - 41, $SD = 6.69$). Overall, five students passed the fifth grade, statewide writing test (Tim, Ally, Carl, Brian, and Navid), and four students did not (John, Andy, Mary Kate, and Clare). Mario did not take the test.

Table 1 provides an overview of participant characteristics. Brief descriptions of each student follows.
<table>
<thead>
<tr>
<th>Student</th>
<th>Gender</th>
<th>Ethnicity</th>
<th>Age</th>
<th>Eligibility</th>
<th>Behavioral Goals</th>
<th>Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ally</td>
<td>Female</td>
<td>White</td>
<td>13.1</td>
<td>ED</td>
<td>Use coping skills to react to frustration. Turn assignments in on time. Use assertive language to ask for what she needs and take responsibility for actions.</td>
<td>WISC IV (12/05): FSIQ = 122 W-J III (12/05): BWL = 96</td>
</tr>
<tr>
<td>Andy</td>
<td>Male</td>
<td>White</td>
<td>13.2</td>
<td>ED</td>
<td>Make on-topic, appropriate comments in class. Be prepared for class with work and materials. Complete assignments on time.</td>
<td>WISC IV (10/07): FSIQ = 92</td>
</tr>
<tr>
<td>Carl</td>
<td>Male</td>
<td>White</td>
<td>14.3</td>
<td>LD, ED</td>
<td>Demonstrate active listening skills. Use study skills for assignment completion. Self-advocate when stressed or anxious. Follow classroom procedures and complete assignments.</td>
<td>WISC IV (10/09): FSIQ = 82 TOWL-3 (10/09): WLQ = 76</td>
</tr>
<tr>
<td>Clare</td>
<td>Female</td>
<td>White</td>
<td>13.5</td>
<td>OHI, Autism</td>
<td>Begin task and maintain attention throughout whole task.</td>
<td>RIAS (5/10): CIX = 109 KTEA-II (5/10): WLC = 89</td>
</tr>
<tr>
<td>John</td>
<td>Male</td>
<td>Multi</td>
<td>13.10</td>
<td>ED, LD</td>
<td>Age-appropriate conversations with peers and adults. Complete assignments on time.</td>
<td>WISC IV (12/04): FSIQ = 98 W-J III (11/04) BWL = 82</td>
</tr>
<tr>
<td>Mario</td>
<td>Male</td>
<td>Hispanic</td>
<td>14.2</td>
<td>LD, ED</td>
<td>Begin and complete classroom assignments. Respect peers and adults in conversation.</td>
<td>WISC IV (1/08): FSIQ = 105 W-J III (1/08): BWL = 81</td>
</tr>
<tr>
<td>Mary</td>
<td>Female</td>
<td>White</td>
<td>13.9</td>
<td>LD</td>
<td>Attempt classroom assignments independently. Immediately respond to adult directions.</td>
<td>WISC IV (5/09): FSIQ = 79 W-J III (5/09): BWL = 65</td>
</tr>
<tr>
<td>Tim</td>
<td>Male</td>
<td>White</td>
<td>13.2</td>
<td>OHI</td>
<td>Complete assignments on time. Transition in school environment with appropriate social skills.</td>
<td>WISC IV (12/08): FSIQ = 120</td>
</tr>
</tbody>
</table>

Ally was a 13 year 1 month old, White female. She was identified as a student with ED due to both internalizing (e.g., depression), and externalizing problems (e.g., hyperactivity), as rated on the Behavioral Assessment Scale for Children (2nd ed., BASC-2; Reynolds & Kamphaus, 2004), administered when she was eight years old. Ally’s IEP goals included using coping skills to react to frustration, turning in assignments on time, and self-advocacy. According to her report cards, Ally’s writing was considered on grade level, and she did not have any writing goals on her IEP. On good days, Ally happily participated during group instruction and enjoyed engaging with her peers. On bad days, Ally either refused to work, putting her head down on the desk and becoming unresponsive to teacher attempts to communicate, or she lashed out verbally, threatening to injure herself. Ally had a history of self-mutilating behavior.

Andy was a 13 year 2 month old, White male. He was identified as a student with ED. Testing completed when Andy was nine years old indicated that his teachers and parents rated him as at-risk, or clinically significant, on a range of items on the BASC-2 (Reynolds & Kamphaus, 2004) including internalizing behaviors (particularly depression), adaptive skills, social skills, functional communication, and externalizing behaviors. Andy’s IEP goals focused on making on-topic comments in class, being prepared for class, and completing classroom and homework assignments. He had an IEP goal for using the writing process (prewriting, writing, and editing) to write multi-paragraph essays. In class, Andy was well-behaved and got along well with his teacher and peers.
Brian was a 13 year 7 month old, White male. He was identified as a student with ED and OHI (due to a medical diagnosis of ADHD). Brian displayed internalizing behaviors. BASC-2 testing conducted when Brian was nine years old found that parent, student, and teacher reports were all consistent with a diagnosis of depression. On the TOWL-3 (Hammill & Larsen, 1996) administered when Brian was 12 years old, his Written Language Quotient was in the average range (SS = 99). However, his in-school writing performance was significantly lower as indicated by the limited amount of writing he produced, which consisted of only a few sentences. Although Brian was an avid reader, he was an extremely reluctant writer. He often refused to complete writing tasks, with his refusal sometimes escalating into physical defiance (e.g., turning over desks). His IEP goals focused on task avoidance and coping with anxiety and frustration. Brian also had an IEP goal for writing five-paragraph essays using assistive technology.

Carl was a 14 year 3 month old, White male. He was identified as a student with both LD and ED. BASC-2 (Reynolds & Kamphaus, 2004) testing completed when Carl was 13, indicated that he was in the clinically significant range for school problems and adaptive skills. His Written Language Quotient score on the TOWL-3 (Hammill & Larsen, 1996) was in the low range (SS = 76). Carl’s IEP goals focused on attending (i.e., active listening, organizing, and monitoring task completion); self-advocacy, when anxious or stressed; and following classroom procedures. He also had an IEP goal focused on writing multi-paragraph essays, including prewriting, drafting, revising, and editing. In class, Carl was polite, actively engaged in lessons, and supportive of his peers.
Clare was a 13 year 5 month old, White female. She was identified as a student with autism and OHI. BASC-2 (Reynolds & Kamphaus, 2004) testing conducted in 2010 indicated that on the behavioral symptoms checklist she was in the clinically significant range for internalizing problems (i.e., anxiety). On the adaptive skills composite she was at-risk for externalizing problems (i.e., hyperactivity). Achievement tests showed Clare performing in the low average range (SS = 89) on the Written Language Composite of the KTEA-II (Kaufman & Kaufman, 2004). Clare’s IEP included goals for maintaining attention and completing multi-paragraph essays. In class, Clare tried to please the teacher and be a good student. She often made comments when the teacher or other students were talking, and seemed anxious about her inability to control these outbursts. She worked quickly on all tasks, and did the minimum amount required.

John was a 13 year 10 month old, multiracial male. His areas of special education eligibility included ED and LD. Testing conducted when John was 10 years old using the BASC-2 (Reynolds & Kamphaus, 2004) indicated that he was at-risk for internalizing problems. His IEP goals focused on both social skills (i.e., having age-appropriate conversations with peers and adults) and study skills (i.e., completing assignments on time). In class, John was an active participant in teacher-led activities. When working independently, or with peers, he often tried to finish his work quickly, sometimes appearing to do the minimum amount necessary to meet the requirements of the task. John also frequently told stories about out-of-school activities that appeared exaggerated and aimed at eliciting a response from the teacher.
Mario was a 14 year 2 month old, Hispanic male. He was identified as a student with LD and ED, as well as having limited English proficiency. On the BASC-2 (Reynolds & Kamphaus, 2004), which was administered when he was 12 years old, Mario scored in the clinically significant range for externalizing problems (i.e., hyperactivity, aggression, and conduct), and at-risk for internalizing problems (i.e., anxiety, depression). He scored in the below average range (SS = 81) on the Broad Written Language Composite of the W-J III (Woodcock, McGrew, & Mather, 2001), administered when he was 10 years old. Mario’s IEP goals addressed task avoidance and respectful interactions with peers and adults. Additionally, he had an IEP goal for writing multi-paragraph essays, following the steps of the writing process. In class, Mario would, at times, push the limits of acceptable behavior, and regularly was sent to work at a carrel in the hallway, rather than sit next to the other students. He had several friends in the class, but disagreements quickly escalated to the point where adult intervention was required.

Mary Kate was a 13 year 9 month old, White female. She was identified as having ED and LD. Additionally, Mary Kate had cerebral palsy, and ADHD. BASC-2 (Reynolds & Kamphaus, 2004) testing, completed when she was 12, indicated that Mary Kate was in the clinically significant range on the behavioral symptoms index and school problems composite, and at-risk for both internalizing (i.e., depression, withdrawal) and externalizing problems (i.e., hyperactivity, aggression). Achievement testing conducted at the same time showed Mary Kate earning a score in the very low range (SS = 65) on
the Broad Written Language Cluster of the W-J III (Woodcock, McGrew, & Mather, 2001).

Mary Kate’s IEP included goals for task avoidance and reactions to anxiety and frustration, as well as an IEP goal for writing multi-paragraph essays. In class, Mary Kate was a very reluctant writer, regularly refusing to write. When pressed by the teacher, she would become verbally aggressive. Additionally, Mary Kate had difficulty maintaining her attention during class. Some of her behaviors were similar to those of students with autism (e.g., difficulty with transitions, compulsive interests, difficulty forming relationships with peers), and she received instruction in a self-contained class for students with autism for part of the school day.

Navid was a 14 year 2 month old, multiracial male. He was identified as a student with ED. On the BASC-2 (Reynolds & Kamphaus, 2004), administered when Navid was 13, he was rated as at-risk for attention problems and school problems. Navid’s self-ratings did not show him having any at-risk or clinically significant scores on the measure. In spite of this, Navid struggled to get along with peers and deal with frustration in school. His IEP goals included accepting help from the teacher and responding appropriately to redirection. He did not have a writing goal on his IEP.

Navid also had a notable number of absences from school and was regularly tardy. On other days, he would be so preoccupied by events that had happened at home or at school that he would be unable to complete his work.

Tim was a 13 year 2 month old, White male. His special education eligibility was other health impairment (OHI). Under this label he received a number of services,
including services provided by specialists in ED, autism, counseling, LD, and PD. Tim exhibited internalizing behaviors. He was under the care of a doctor for ADHD and an anxiety disorder. His IEP goals included study skills (i.e., completing assignments, transitioning in the school environment) and written expression (i.e., writing multi-paragraph essays). While his IEP indicated that Tim could use voice recognition software, there were problems with his laptop that precluded the use of the software during the course of this study. Instead, he typed his essays using a word processing program. In class, Tim had a great deal of difficulty maintaining his focus. He participated minimally in whole class instruction and often failed to complete in-class assignments.

Dependent Measures

Dependent measures were selected to assess students’ writing skills, self-efficacy for persuasive writing, and writing fluency. Additional data was gathered to assess students’ on-task behavior during both teacher-led and student-led instruction, as well as during independent work time. Finally, interviews with students were conducted to obtain information about the usefulness and effectiveness of the writing and revision strategies.

Writing skills. Students wrote persuasive essays in response to a writing prompt: (a) during the baseline phase, (b) following each instructional phase, and (c) six weeks after instruction was completed. Students were given a choice of two prompts and asked to write a good persuasive essay in response to one of them (see Appendix A). A list of prompts used in previous SRSD studies (Mastropieri et al., 2009; 2010) was given to the
students’ special education teacher for review. The special education teacher provided
feedback about the level of difficulty of each prompt and the level of interest students
were likely to have for each prompt. One potential prompt was dropped from the list
because students had written an essay on the same topic during the previous academic
year. After the teacher reviewed the list of prompts, the prompts to be used for baseline,
post-instructional phases, and maintenance measures were selected. All prompts were
judged to be comparable in difficulty with the potential to also generate a high level of
interest from participants. First and final drafts of student essays were reviewed and
rated in accordance with scoring conventions and a Holistic Quality Scoring Rubric (see
Appendix B-C). The scoring conventions were adapted from the conventions used to
score essays in the Mastropieri et al. (2009; 2010) studies, and the Holistic Quality
Scoring Rubric is the same rubric used in these studies.

Persuasive Writing Self-efficacy Scale. Students’ self-efficacy for writing was
measured using a researcher-designed Persuasive Writing Self-efficacy Scale (see
Appendix D). To create the scale, the researcher followed Bandura’s guidelines for
creating self-efficacy scales (Bandura, 2006). For example, Bandura recommended using
the language “can” rather than “will,” providing a response scale from 0% confident to
100% confident, and linking the scale to a specific task. This 13-item Persuasive Writing
Self-efficacy Scale asked students to rate their confidence in performing different parts of
the writing task such as planning, drafting, and revising. Because writing self-efficacy is
very task-specific, students were asked to fill out the self-efficacy scale in reference to
one of the writing prompts as shown in Appendix A. That is, students were presented
with two essay prompt choices and selected the prompt they wished to write about. After selecting a prompt, students completed the self-efficacy scale before they began planning and drafting their essays. Using a Likert-type scale ranging from 1 (0% confident) to 5 (100% confident) students were asked to respond to a series of questions such as, “How sure are you that you can write an essay that contains at least 10 persuasive essay parts?” and “How sure are you that you can make a plan before writing your essay?” The self-efficacy scale was administered a total of four times—before students wrote their first essays: at (a) baseline, (b) post-SRSD instruction, (c) post-revision instruction, and (d) maintenance.

**Writing fluency.** Writing fluency refers to the ability to quickly and accurately write sentences that express complete thoughts. Although writing fluency was not explicitly taught during this study, fluency was assessed to determine whether it increased simply as a result of intensive instruction and practice in writing. The W-J III writing fluency subtest (Woodcock, McGrew, & Mather, 2001) was used to measure writing fluency at baseline testing and following the revision instructional phase. The W-J III is a standardized test of academic achievement that includes subtests in reading, writing, and mathematics, along with other academic and skill areas. The writing fluency subtest presents students with a picture and three words. They must write one sentence that includes the three words and tells about the picture. The format of the test can be seen in Appendix E, which contains the practice items for this test. Responses can include any additional words needed to make a complete sentence, but the stimulus words
cannot be modified in any way (e.g., cannot write “singing” if the stimulus word is “sing”). Students must complete as many sentences as possible within seven minutes.

On-task behavior. Time sampling and coding of students’ on-task behavior occurred during 52.2% of all instructional sessions, SRSD instruction (48.5%), and peer-revision instruction (55.5%). Selection of instructional sessions for assessing on-task behavior was purposeful, in order to provide a balance across (a) groups, (b) instructional phases, and (c) types of instruction (i.e., teacher-led instruction and student-led instruction, during peer conferences, and independent work time). Videotapes of the selected instructional sessions were used to code behavior for Groups 1 and 2. Live coding by an experienced, trained coder was done for Groups 3 and 4, as at least one student in each group had not provided permission for videotaping.

Student behavior was coded every 30 seconds for 15 minutes. A beep tape was used to indicate when observations should be made. On-task and off-task behavior was recorded on a chart (see Appendix F) using paper and pencil. On-task behavior was defined as students (a) being in the assigned areas of the room, (b) engaging with instructional materials, (c) writing to the given prompt, and (d) asking relevant questions (Mastropieri et al., 2009). During the writing task, students who sometimes appeared to be thinking or looked away from their material for short periods of time (i.e., 10 seconds or less) were still considered to be on-task. Behaviors that deviated from those defined above were coded as off-task. If a student was out of the classroom during a coding interval or could not be seen on the video camera, it was noted as such.
Social validity interviews. Students were interviewed individually after revision posttesting was completed to provide feedback about the usefulness and acceptability of the SRSD writing instruction and peer-revision process. The same social validity interview was used in previous SRSD studies (Mastropieri et al., 2009; 2010). Sample questions included, “How has the POW+TREE strategy helped you become a better writer?” “If you were the teacher, what would you change about the POW+TREE lessons?” and “How did working with a writing partner help you revise?” The entire social validity interview is included in Appendix G. All interviews were audio recorded for review and analysis.

Materials

Materials were designed for the students, the teacher, the observers, and the scorers. Each set of materials are described next and are included in the Appendix.

Student materials. Student materials were developed for each phase of the study. Materials included student folders, which contained all copies of handouts and student products. Specific materials developed for each phase are described below.

Baseline phase. During the baseline testing phase and subsequent instructional and testing phases, students used Neo word processors to draft and revise their essays. The Neo includes a keyboard and small screen that displays four lines of text. After writing, students uploaded their essays into Microsoft Word, using the classroom desktop computer. From the desktop, they were able to print their essays and save their work onto a flash drive provided by the researcher. The Neo word processors include a spell
checker, which students were allowed to use. Students were also allowed to use a class set of dictionaries to look up the spelling of words.

**SRSD instructional phase.** During the SRSD instructional phase, which taught students to plan and draft persuasive essays, students were introduced to several self-regulatory aids to assist them during the writing process. These materials were adapted from Harris, Graham, Mason, & Friedlander (2008) *Powerful Writing Strategies for All Students* and have been used in previous studies (Mastropieri et al., 2009, 2010). Appendix H-L contains copies of the (a) POW+TREE mnemonic, (b) graphic organizer, (c) self-statement sheet, (d) self-monitoring graph, and (e) transition word list, which are described in detail below.

The first self-regulatory aid was a visual mnemonic to help students remember the strategy for persuasive essays, POW+TREE. This handout has a picture of a tree on it, along with a list of the components of POW+TREE: Pick my idea, Organize my notes, Write and say more, Topic, Reasons (3 or more) and counter reasons (1 or more), Explain reasons, and Ending and examine (see Appendix H).

Another self-regulatory aid students used was a graphic organizer for planning their essays. The graphic organizer followed the TREE mnemonic, with spaces for students to write a Topic sentence (the T in TREE); three Reasons (R), along with three Explanations for each reason (E); a counter reason, and an explanation for the counter reason; space to refute the counter reason; and an Ending sentence (the second E in TREE). Spaces were also provided for transition words next to each reason to remind students to include transitions from one idea to the next in their writing (see Appendix I).
A third self-regulatory aid provided to students was a sheet for recording self-statements. The sheet has a picture of a tree on it, along with lines to write down words or phrases students could say to themselves before they started writing, while they wrote, and when they finished writing. The class brainstormed positive self-statements to include on the sheet, and students were encouraged to use these as they worked on their essays (see Appendix J).

The fourth self-regulatory aid used during SRSD instruction was the self-monitoring graph. Students used the self-monitoring graph to record the number of essay parts included in each essay and to set goals for the next essay. Each column on the chart includes a space to record the date and the number of transition words used in the essay. Then, 14 squares are provided for students to color in the number squares corresponding to the number of essay parts they wrote. At the top of each column is a space for students to write down a goal for the number of parts they will include in the next essay (see Appendix K).

Lastly, students received a copy of a transition word list (see Appendix L). The handout was divided into three sections: (a) words used to show a reason, (b) words used to show a counter reason, and (c) words used to show an ending. Each section listed transition words, and provided blank spaces for students to add additional words. Students were taught to use the list as a resource when filling out the graphic organizer.

*Revision instructional phase.* During the peer-revision instructional phase, students had four additional handouts to aid them as they completed the revision process. These materials were developed specifically for this study by the researcher. Appendix
M-P includes copies of the (a) good writing handout, (b) revision strategy guide, (c) peer-
revision checklist, and (d) conferencing feedback guide, which are described in detail
below.

The good writing handout was used to maintain a list of features of good writing,
which students referred to as they evaluated each other’s work. The handout, itself,
contains three prompts: “Good writing (a) looks like…, (b) sounds like…, and (c) makes
me feel….” Each prompt is followed by four blank lines for students to fill in with their
ideas (see Appendix M).

The revision strategy guide led students through providing feedback to their
partners during peer-revision conferences. The guide contains a chart with sections for
providing an opening compliment, evaluating the persuasive essay parts included in the
essay, providing feedback about the clarity of each part and the persuasiveness of the
arguments presented in the essay, and for a closing compliment. The bottom of the
revision strategy guide provides boxes for students to check off when they edit their
partners’ final essays for capitalization, punctuation, and spelling (see Appendix N).

The third handout used during the peer-revision instructional phase was the peer-
revision checklist. The checklist details all of the steps for conferencing, from finding an
appropriate place in the classroom for the conference, to completing the revision strategy
guide, making final edits, and turning in a copy to the teacher. A small box is included
next to each step, and students were instructed to check off the steps as they were
completed. The bottom of the checklist lists the five behavioral expectations for peer-
conferencing: (a) try hard, (b) do your best, (c) be positive, (d) be kind and encouraging to partners, and (e) when in doubt, raise your hand for assistance (see Appendix O).

The final handout was a conferencing feedback guide (see Appendix P). It was used to provide students with possible verbal responses to share with their partners during the peer-revision conference. The guide contains bulleted lists of possible comments corresponding to the sections of the revision strategy guide (i.e., compliment, parts, clarity, persuasiveness). Both positive and negative comments are included, either written as statements or as questions. Sample comments included, “It was very easy to find all of your essay parts.” “I did not see a counter reason. Who would disagree with your opinion?” and, “This might be a strong reason, but it needs more details to support it.”

In addition to the handout materials used during SRSD instruction and peer-revision instruction, digital audio recorders were used to record each peer-revision conference and teacher-student revision discussion. Each pair of students was provided with a recorder to use, and shown how to start and stop recording. Students took turns being responsible for recording their conferences.

Finally, students were administered a strategy knowledge measure (see Appendix Q). The measure contained two questions: What are the parts of a good persuasive essay? and What should you look for when you revise a persuasive essay?

Teacher materials. Teacher materials were developed for use during the SRSD instructional phase and the revision instructional phase.
**SRSD instructional phase.** Teacher materials for the SRSD instructional phase included a set of 11 inch by 18 inch, laminated posters. The posters could be written on with overhead markers and were reused for each instructional group. Posters were made of all SRSD handouts, including (a) POW+TREE mnemonic, (b) graphic organizer, (c) self-statement sheet, (d) self-monitoring graph, and (e) transition word list. In addition, posters were made of the sample essays used during SRSD Lessons 1 and 2.

Additionally, reinforcers were used to encourage on-task behavior and reward task completion. Reinforcers included both tangible items like candy, and non-tangible items like five minutes of library time. Such reinforcers were used primarily during SRSD Lessons 4 and 5 when students were required to write independently.

**Revision instructional phase.** Similar to SRSD instruction, 11 inch by 18 inch, laminated posters were used for instruction during the revision instructional phase. For the revision phase, posters were used for the (a) good writing handout, (b) revision strategy guide, (c) peer-revision checklist, and (d) conferencing feedback guide.

Additionally, to support instruction in the peer-revision process, a researcher-made video was used to model the steps of the peer-revision conference. In this video, the instructor and a colleague were seated next to each other at a table in a conference room. Both had copies of the revision strategy guide and the peer-conferencing checklist, as well as sample essays that they used to model the conference. They followed the checklist to conduct a peer-revision conference for their essays, modeling each step of the process and checking it off as it was completed. A third adult acted as the teacher in the model conference.
Lastly, reinforcers were also used during the peer-revision phase to encourage on-task behavior and reward task completion. As during the SRSD instructional phase, reinforcers included both tangible and non-tangible items. Reinforcers were used primarily during revision Lessons 4 and 5 when students were expected to work more independently.

**Observer materials.** Experienced, trained observers were used to collect on-task behavior data and fidelity of treatment data. The on-task behavior coding sheet can be found in Appendix F. The same on-task behavior coding sheet was used in the Mastropieri et al. (2009, 2010) studies. The fidelity of treatment checklist can be found in Appendix R, and the peer-revision steps completed checklist can be found in Appendix S. The fidelity of treatment checklists for SRSD instruction are the same checklists used by Mastropieri et al. (2010); the fidelity of treatment checklists for revision instruction were developed for this study.

Videotapes of instructional sessions were used for on-task behavior coding and observing fidelity of treatment for Groups 1 and 2. The videotapes for Groups 3 and 4 focused only on the instructor, due to students not providing permissions for taping. Therefore, fidelity of treatment for Groups 3 and 4 was completed based on the videotaped sessions, but live coding was done to capture students’ on-task behavior in Groups 3 and 4. For all on-task behavior coding, a beep-tape was used to prompt observers to record student behavior at 30-second intervals over a 15-minute period. The beep-tape loaded onto an iPaq handheld device and included a beep every 30 seconds.
Scorer materials. Experienced, trained scorers were used to score the first and final drafts of student essays. Each scorer was provided with a typed list of scoring conventions for each item to be scored (e.g., number of essay parts, number of mechanical errors, number of revisions; see Appendix B), along with a copy of the holistic scoring rubric (see Appendix C). Essays were compiled, and both electronic and hard copies were provided to scorers.

Procedures

HSRB approval from George Mason University and the school district were obtained through a large, multi-year writing grant. Parental consent and student assent were obtained before beginning the study. Consent was sought for all students in the targeted classes and all students with signed consent and assent forms participated.

Training data collectors. Experienced research staff members were trained to collect on-task behavior data, to collect fidelity of treatment data, and to score persuasive essays. Data collectors were four graduate research assistants and one research staff member who had collected similar data during previous SRSD studies (Mastropieri et al., 2009; 2010). An additional data collector, a former middle school English teacher, was trained to collect fidelity of treatment data.

Two coders were trained to collect on-task behavior data. The training began with the researcher providing instruction on the scoring conventions for on-task behavior and off-task behavior. Then, behavior coders practiced coding on-task behavior from a videotape of writing instruction from a previous study. After comparing results, the
coders continued to practice coding videotapes of instruction until they were in 100% agreement on the coded behaviors.

Training for fidelity of treatment coders, two of whom were also on-task behavior coders, began with the researcher reviewing all of the lesson plans with the coders. Three of the four fidelity of treatment data collectors had completed similar work in previous studies (Mastropieri et al., 2010). Therefore, the SRSD lessons were familiar to the coders. In-depth explanations of the revision lesson plans were given, including training on how all handouts and materials would be used. Coders were given copies of the detailed lesson plans, along with the fidelity of treatment checklist (Appendix R), to use as a reference.

Finally, five individuals were trained to score students’ persuasive essays. All scorers had experience scoring persuasive essays in previous SRSD studies (Mastropieri et al., 2009; 2010). Training began with a review of the scoring conventions (see Appendix B). Then, scorers practiced scoring a common essay. Scores were compared across scorers, and scoring conventions were clarified. Next, each scorer scored three essays. Then scores on those essays were compared, and conventions were further clarified. The researcher continued to compare essay scores across scorers and clarify scoring conventions until all scorers were in 90 percent agreement across essays.

Grouping students for instruction. The special education teacher was asked to group students who worked well together, for instruction. Students were placed into one of four instructional groups, with two to four students in each group. Group 1 included Tim, Carl, and Andy; Group 2 included Ally and John; Group 3 included Mario, Navid,
and another student who was later dropped from the study; and Group 4 included Clare, Mary Kate, Brian, and another student who was later dropped from the study.

All students began baseline testing at the same time. However, groups entered the instructional phases in a staggered fashion. Group 1 received baseline essay prompts until a stable baseline was established, at which time they entered the SRSD instructional phase. Groups 2, 3 and 4 received additional essay prompts after Group 1 began instruction. After establishing a stable baseline, Group 2 entered the SRSD instructional phase, while Groups 3 and 4 received additional essay prompts. After establishing a stable baseline, Group 3 entered SRSD instruction, and Group 4 received an additional essay prompt. Finally, after establishing a stable baseline, Group 4 entered the SRSD instructional phase.

Instruction took place during two different 50-minute class periods across the school day. The researcher was the instructor for all groups and conducted all testing. Students worked in groups with two to four students in a classroom with the instructor. Students sat at desks while the teacher taught from the front of the room. An instructional assistant was asked to join the class when a student needed an adult to sit with him or her, in order to facilitate the student’s continued work. Such in-class assistance was required for two of the instructional groups (Group 3 and Group 4), with each group requiring the aid of the instructional assistant for approximately 50-75% of the instructional sessions. When student behaviors became extreme, a crisis resource specialist was called to remove the student from class. This occurred twice in Group 2,
three times in Group 4, and three times in Group 3. Students in Group 1 did not require the crisis resource specialist.

**Baseline testing phase.** All students began with baseline testing, including essay prompts (Appendix A), the Persuasive Writing Self-efficacy Scale (Appendix D), and (Appendix E) the W-J III writing fluency subtest (Woodcock et al., 2001). Students received additional essay prompts until a stable baseline was reached. The number of essay parts was the measure used to determine baseline stability. To be considered stable, at least three essays had to be completed, with a range of number of essay parts no greater than two (see essay scoring procedures in Appendix B). Additional essay prompts were given until this requirement was met. Four students (John, Brian, Mary Kate, and Mario) completed one additional essay to meet this stability requirement, while one student (Ally) completed two additional essays to establish stability of baseline. Two of Ally’s five essays were written about self-selected topics, rather than the researcher-generated prompts, and were not scored. The last baseline essay prompt for each group was administered two days before SRSD instruction for that group began, to allow students one day for planning and drafting and one day for revising their essays.

Students composed essays for testing purposes and during instruction using a word processing program. The special education teacher had a class set of Neo word processors (similar to Alpha Smarts), which students used for both drafting and revising. After students were done typing, they uploaded their essays into Microsoft Word to print and save their work. The Neo has a built-in spell checker that students used, although its dictionary was limited. Upon request, students were permitted use of the spell checker in
Microsoft Word, which contains more word choices. Dictionaries were also available for student use during drafting and revising.

Two students asked to write some of their post-test essays using pencil and paper, rather than typing them. These requests were granted. Photocopies of the draft essays were made, and students revised the handwritten copies of their essays. After testing was completed, both the draft and revised versions of the essay were typed up by the researcher for scoring.

Students were given two days to write each essay. On the first day, students received two prompt choices, selected a prompt, and wrote a persuasive essay on the topic. The essays were saved to a flash drive and two copies were printed: one for scoring and one for students to revise the next day. On day two, students were given the hard copies of their essays and asked to mark any revisions. Students made the changes on the computer, saved their work, and printed out a copy for scoring. The first draft and revised essay were then collected for analysis.

Self-regulated strategy development (SRSD) instructional phase. Students in this study learned the SRSD strategy of POW+TREE for writing persuasive essays. Instruction followed the six stages of SRSD instruction, as outlined by Harris et al. (2008) in *Powerful Writing Strategies for All Students*. Groups 2, 3, and 4 completed the lessons during eight, 50-minute instructional sessions, for a total of 400 minutes of SRSD instruction. Group 1 required nine sessions to complete the lessons, for a total of 450 minutes. (Average total time for SRSD instruction = 412.5 minutes.) Students who missed two consecutive sessions, or more than two sessions per week, received make-up
sessions during their daily study skills class. Ally (Group 2) received one make-up instructional session for this reason.

The six stages of instruction were taught during five lessons, with some lessons taking more than one instructional session to complete. The same lesson plans were used as in the Mastropieri et al. (2010) study. During the first lesson, students discussed what makes a good persuasive essay. They were then introduced to the POW+TREE mnemonic and given a visual representation of it. POW stands for Pick my idea, Organize my notes, Write and say more. TREE represents the components of an opinion essay: Topic, Reasons and counter reasons, Explanations, and Ending (see Appendix H).

Next, students practiced finding all of the components of TREE in sample essays and filling in the POW+TREE graphic organizer with this information (see Appendix I). In particular, students were taught to identify the topic sentence, three or more reasons that support the writer’s topic, explanations for these reasons, and the ending. The graphic organizer had spaces for each of these essay parts. Finally, students memorized the mnemonic during the first lesson.

In lesson two, students were introduced to counter reasons and their importance in persuasive essays. Students practiced developing counter reasons for practice essay prompts. Additionally, students were introduced to transition words. Then, they continued to practice finding all of the parts of a persuasive essay and writing the parts on the graphic organizer. During this lesson, counter reasons, explanations for the counter reasons, and refutations of the counter reasons were included in the sample essays.
Lesson three consisted of the teacher modeling how to write a good persuasive essay. Following all of the steps in POW—Pick my idea, Organize my notes, Write and say more—the teacher thought aloud while planning and writing an essay. Afterwards, the teacher guided students to develop self-statements they could use before, during, and after writing to provide positive self-talk throughout the process. The lesson ended with students graphing the number of essay parts included in the new, teacher-generated essay on the self-monitoring graph. (See Appendix J for the self-statement sheet and Appendix K for the self-monitoring graph.)

With the next step in the process, lesson four, the teacher and students drafted an essay together, combining all of the elements they had learned to date. The teacher and students used a graphic organizer to organize notes, used positive self-statements throughout the writing process, graphed their progress, and set goals for the next essay. In lesson 5, the final lesson, students practiced writing an opinion essay independently. The goal was for students to be able to write an essay without using a pre-made graphic organizer (i.e., students drew the graphic organizer for themselves). After completing the essay, students graphed their progress and set goals for the next essay.

Each week during SRSD instruction, students were administered a strategy knowledge measure (see Appendix Q). The measure contained two questions: What are the parts of a good persuasive essay? and What should you look for when you revise a persuasive essay? In this way, the growth in student’s strategy knowledge was demonstrated—first, by improved scores on the parts of a persuasive essay question, and later by improved scores on the revision strategy question.
Additionally, on-task behavior was coded during 48.5% of the SRSD instructional sessions through the use of videotapes (Groups 1 and 2) and live coding (Groups 3 and 4), and included behavior coding during teacher-led instruction and individual work time.

**SRSD posttesting phase.** After students completed the five SRSD lessons, they were given three persuasive essay prompts as a post-test measure. Each prompt contained two essay topics, and students selected one of the topics for their essay. As during baseline testing, students were administered each prompt over two days. On the first day, students were asked to plan and draft their essays. On the second day, students were given the opportunity to revise their essays. Along with the first post-SRSD essay, students were administered the Persuasive Writing Self-efficacy Scale.

One student, Mary Kate, only completed two post-SRSD testing essays. When given the third post-SRSD prompt, Mary Kate first refused to write. The following day, Mary Kate was given a new prompt to write about. Instead of selecting one of the researcher-generated topics, Mary Kate planned and wrote a persuasive essay on a self-selected topic. Because Mary Kate did not respond to the researcher-generated prompt, her essay was not scored.

**Revision instructional phase.** Immediately following SRSD posttesting, students began instruction in the peer-revision strategy. Groups varied in the number of days required to complete revision instruction. Groups 1 and 2 received 10 sessions of revision instruction, for a total of 500 minutes of instruction. Group 3 required nine instructional sessions, or 450 minutes. Group 4 finished revision instruction after 7 sessions, or 350 minutes. On average, students received 450 minutes of instruction over
9 instructional sessions. Students who missed two consecutive sessions, or more than two sessions per week, received make-up sessions during their daily study skills class. Andy (Group 1) received one make-up session and Navid (Group 3) received two make-up sessions due to absences.

Revision strategy instruction was designed to follow procedures similar to SRSD instruction (Harris et al., 2008), beginning with a discussion of the importance of revision, moving on to teacher modeling, guided practice, and independent practice. All lesson plans and materials were designed by the researcher. Lesson one began with a discussion of what revision is and why it is important. Then, the class brainstormed to complete the following statements. Good writing (a) looks like … , (b) sounds like … , and (c) makes one feel … . (see Appendix M.) Each student maintained a copy of this list and added to it throughout the revision lessons. The good writing list was also a resource for students to use when they provided feedback to their peers during conferences, particularly when providing compliments about their partner’s work.

Next, the revision strategy was introduced. The revision strategy prompted students to provide feedback about four aspects of an essay:

1. What did the author do well? (Compliments were given at the beginning and the end of the peer conference.)
2. Does it have all the parts [of a good persuasive essay]?
3. Is it clear?
4. Is it persuasive?
The class then read a sample essay and applied the revision strategy. The revision strategy guide was used to record students’ suggestions for revision (see Appendix N). The strategy guide contained a chart with spaces for students to fill in to provide an opening compliment; check off which parts were present in the essay, and whether each part was clear and persuasive; and a space for a closing compliment. The bottom portion of the guide provided boxes for students to check after they had edited their partner’s paper for capitalization, punctuation, and spelling errors. After identifying revisions that needed to be made, students took turns making the suggested changes on the computer. Finally, students memorized the revision strategy during this lesson.

Lesson two focused on editing. After reviewing the revision strategy and what good writing looks like, the class briefly reviewed the basic rules for capitalization and punctuation. The capitalization rules that were emphasized included (a) capitalize the first word in every sentence, and (b) capitalize proper names. The punctuation rule emphasized was to include an end mark (i.e., period, question mark, or exclamation point) for every sentence. Students also discussed strategies for finding and correcting spelling errors.

Next, students practiced editing a paragraph, and discussed their edits with the class. After that, the class edited the revised essay completed during lesson one, with an emphasis on capitalization, punctuation, and spelling. As in the previous lesson, students took turns making the editing changes on the computer. Finally, students completed guided practice using the entire revision strategy to improve a second essay, and received feedback from the teacher throughout the process.
Lesson three focused on peer conferences. To begin, expectations for peer-revision conferences were introduced (see Appendix O). These expectations included both the logistics and the content of the conferences. Table 2 details these steps.
<table>
<thead>
<tr>
<th>Steps</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Find a place in the classroom to conference where you will not</td>
<td>When working with partners, students spread out to work at desks throughout the room so that they did not distract or interact with other groups.</td>
</tr>
<tr>
<td>distract other students.</td>
<td></td>
</tr>
<tr>
<td>2. Turn on the audio recorder to record your conference.</td>
<td>Each pair was provided with a digital audio recorder to record their conference. Partners took turns being “in charge” of the recorder.</td>
</tr>
<tr>
<td>3. Use quiet voices. Be respectful of other students who are working.</td>
<td>Students were asked to work quietly so that they did not distract other groups.</td>
</tr>
<tr>
<td>4. Take turns reading your essays to each other.</td>
<td>The conference began with each partner reading his or her essay out loud. This was done to ensure that partners knew what the author meant to say (in case of spelling errors), and to provide an opportunity for author review of his or her essay.</td>
</tr>
<tr>
<td>5. Re-read your partner’s essay to yourself and fill out the POW+TREE</td>
<td>Partners switched essays, and completed the revision strategy guide based on their partner’s essay.</td>
</tr>
<tr>
<td>revision sheet.</td>
<td></td>
</tr>
<tr>
<td>6. Conference with your partner about your suggestions.</td>
<td>Students verbally told their partner what they had written on the revision strategy guide. Students were encouraged to use the conferencing feedback guide to provide suggestions.</td>
</tr>
<tr>
<td>7. Revise your own essay on the computer. Save your work. Print out</td>
<td>After the conference, students made changes to their essay based on feedback from their partner. Students were not required to make all of the changes suggested by their partner.</td>
</tr>
<tr>
<td>a copy.</td>
<td></td>
</tr>
<tr>
<td>8. Meet with the teacher to talk about your revisions.</td>
<td>As they finished making revisions, each student met individually with the teacher. He or she read their revised essay out loud to the teacher. Then the teacher asked, “What suggestions did your partner make for revising your essay?” and, “What changes did you make?” The teacher also discussed any additional writing issues with the student, if needed.</td>
</tr>
<tr>
<td>9. When both partners are done revising, meet again to edit your</td>
<td>After making revisions, partners switched essays again to edit. Again, students were to record their interaction using the digital audio recorder.</td>
</tr>
<tr>
<td>partner’s essay for capitalization, punctuation, and spelling.</td>
<td></td>
</tr>
<tr>
<td>Record your conference.</td>
<td></td>
</tr>
<tr>
<td>10. Make your final corrections on the computer. Save your work.</td>
<td>After receiving their partner’s edits, each student made final corrections, printed, and saved their essay. Their final draft, first draft, revision strategy guide, and graphic organizer were turned in to the teacher for evaluation.</td>
</tr>
<tr>
<td>Print out a final copy and turn it in to the teacher.</td>
<td></td>
</tr>
</tbody>
</table>
Teacher conferences in step eight provided an opportunity for students to discuss their revisions with the teacher and explain why they made the changes they did. Teacher-student conferences also provided an opportunity for the teacher to give feedback to students about the clarity of their writing and the cogency of their persuasive arguments (Wong, Butler, Ficzere, & Kuperis, 1996).

During lesson three, time was also spent talking about behavioral expectations for peer conferencing. The five expectations were (a) try hard, (b) be nice, (c) be positive, (d) be kind and encouraging to partners, and (e) when in doubt, raise your hand for assistance. To help students provide feedback to their partners that met these expectations, students were given the conferencing feedback guide (see Appendix P). The guide listed comments, both positive and negative, that could be used to give a peer feedback about his or her essay. For example, the guide included comments like, “You had a reason, but I did not see an explanation,” “You stated your topic clearly,” or “This reason does not match the topic sentence that you wrote.”

Next, students watched a video of a model peer-revision conference. In the video, the instructor and a colleague modeled each step of the peer-revision conference, using the revision strategy guide, peer-revision checklist, and conferencing feedback guide. A third colleague played the role of the teacher, modeling the teacher-student conference in step eight. As students watched the video, they were tasked with checking off the steps of the conference on the peer-revision checklist. They were also asked to review the conferencing feedback guide and identify any feedback from the guide that was used
during the model lesson. Lesson three ended with students reviewing and memorizing the steps of the revision strategy.

Lesson four provided guided practice to student pairs as they completed the peer-revision conference. Each student selected one of the essays he or she wrote during SRSD instruction to use for this conference. Students met in groups of two or three (depending upon the number of students in the instructional group), with the teacher selecting peer-revision partners. All of the peer-conferencing steps were completed as the teacher provided feedback to students. Particular attention was paid to helping students provide their suggestions in positive and constructive ways. Peer conferences were audio recorded to allow for fidelity of treatment scoring. Following guided practice, the teacher led a whole-class discussion about what went well during the peer-revision conferences and what students needed to improve upon during the next conference.

During lesson five, students practiced planning, drafting, and revising persuasive essays, combining both the POW+TREE strategy and the peer-revision strategy to complete the entire writing process. Again, the teacher selected peer-revision partners, choosing different partners than those used in lesson 4, when possible. This was done so that students had an opportunity to receive feedback from different peers throughout the course of instruction. At the end of the lesson, the teacher again led a discussion with students about what went well during the peer conferences and what needed to be improved.

As during the SRSD instructional phase, students were given the strategy knowledge measure each week during the peer-revision instructional phase.
Furthermore, behavior was coded for 55.5% of instructional sessions, including both teacher-led and student-led (i.e., peer conferences) instruction, as well as independent work time.

Revision instruction posttesting phase. Students received comprehensive posttesting following instruction in the peer-revision strategy. To begin, students received a series of three persuasive essay prompts administered under the same conditions as baseline testing and post-SRSD testing. Again, Mary Kate completed only two of the three post-instruction essays. Additionally, students completed the Persuasive Writing Self-efficacy Scale before writing the first post-revision essay, and took the W-J III writing fluency subtest. The researcher also conducted one-on-one social validity interviews with students. Lastly, students were administered a generalization essay prompt asking them to write an essay on a topic covered in their science or social studies curriculum.

Maintenance testing phase. Maintenance testing occurred six weeks (Groups 1 and 2) or eight weeks (Groups 3 and 4) after the final post-instruction writing prompt was given. During maintenance testing, students were given one persuasive essay prompt along with the Persuasive Writing Self-efficacy Scale. Testing was completed over two sessions, allowing students one session for planning and drafting their essay, and one session for revising.

Fidelity of Treatment

To ensure that the instructional program was implemented as intended, instructional sessions were videotaped. For Groups 3 and 4, in which some students did
not provide permission for videotaping, the camera was focused only on the instructor. To assess the fidelity of treatment, four experienced coders viewed all instructional sessions and completed fidelity of treatment checklists (see Appendix R). The checklists included the key components of each lesson.

Reliability and Scoring

**Persuasive essay scoring.** Students were administered persuasive essay writing prompts during the baseline phase (three to four prompts), following SRSD instruction (three prompts), and following revision instruction (three prompts). At the time of post-revision instruction testing, students also completed one generalization essay. The generalization essay prompts reflected content in science and social studies that was covered in the eighth grade curriculum. A final essay prompt was given at maintenance testing. Groups 1 and 2 were administered maintenance testing six weeks after revision instruction posttesting was complete; Groups 3 and 4 were administered maintenance testing eight weeks after revision posttesting was complete. One student, Ally (Group 2), refused to participate in maintenance testing.

For each essay students completed, they were given two essay prompts from which to choose. After selecting an essay, students were given the class period to plan and draft an essay. During the following class period, students were asked to revise their essays. Both the first and final drafts of essays were collected and scored.

First and final drafts of each essay were scored across four dimensions. First, essay length was evaluated by counting total number of words. The word count feature on Microsoft Word was used for this purpose. Second, the total number of essay parts
was tallied. Parts included: topic sentence, reasons, explanations, counter reason(s), explanation for the counter reason, refute, and ending sentence. Each part was given one point, with each reason, explanation, and counter reason receiving a point. For example, if a student had a complete essay, he would receive one point for a topic sentence, three points for three reasons (one point for each reason), three points for three explanations (that is, an explanation for each reason), one point for a counter reason, one point for explaining the counter reason, one point for refuting the counter-reason, and one point for an ending sentence, for a total of 11 points. An essay scored higher than 11 points on the essay parts measure if it included more than three reasons or additional counter reasons.

Third, essays were scored according to a holistic scoring rubric (Appendix C), similar to the rubric used in other SRSD studies (Mastropieri et al., 2009; 2010). This rubric evaluated essays on a 10-point scale, focusing on content and organization. Writing mechanics were not evaluated as part of the holistic rubric. The fourth dimension across which essays were scored was mechanical errors. Errors in punctuation, capitalization, and spelling were tallied and reported as the proportion of errors per 100 words. Punctuation errors were defined as sentences that did not end in a period, exclamation point, or question mark (MacArthur, Schwartz, & Graham, 1991). Capitalization errors were defined as first words of sentences that were not capitalized (MacArthur et al., 1991). Spelling errors included all errors in spelling. Misspellings of people’s names were not counted as errors because many names are routinely spelled in different ways.
The revisions made from first draft to final draft were also tallied and categorized. Specifically, all changes made from the first draft to the final draft were counted as revisions. Next, each revision was classified as either a meaning-change revision or a surface-change revision (MacArthur et al., 1991; Stoddard & MacArthur, 1994). Surface-change revisions referred to (a) revisions of mechanics (i.e., capitalization, punctuation, and spelling), (b) format, or (c) morphological changes (e.g., change in tense or number). All other revisions were classified as meaning-change revisions.

The researcher scored student essays according to the above measures. In addition, four trained, experienced scorers evaluated 41.82% of the essays. They then compared their results to compute interrater agreement. Any discrepancies in scoring were discussed. When consensus was reached on how the essay should be scored, the agreed upon score was used. If no agreement could be reached, the researcher’s scores were used. Interrater agreement for essay scoring was 95.29%.

**Persuasive writing self-efficacy scale scoring.** Students were administered the self-efficacy scale before completing a persuasive essay prompt, at four points during the study with (a) the first baseline essay, (b) the first SRSD posttesting essay, (c) the first revision posttesting essay, and (d) the maintenance essay. The self-efficacy scale contained 13 items, rated on a 1-5 point Likert-type scale, with 1 representing 0% confidence, and 5 representing 100% confidence.

Students’ ratings on each item were summed to determine the total scale score. The total score then was divided by the total number of items on the scale (13) to calculate a composite self-efficacy score for each student. Possible composite self-
efficacy scores ranged from 1 (indicating the student had no confidence in his own ability to perform the persuasive writing task) to 5 (indicating that the student was 100% confident in his ability to perform the persuasive writing task).

Woodcock-Johnson III Writing Fluency subtest scoring. Students were individually administered the Writing Fluency subtest during baseline testing and during revision posttesting. The test required that students write a complete sentence for each stimulus picture and three corresponding words. Students had seven minutes to write as many sentences as possible.

The researcher scored all students’ Writing Fluency subtests. Each item for which the student wrote a complete sentence that included the required stimulus words received one point. The total number of points earned was the student’s raw score. These raw scores were then converted into standard scores using the computer-based scoring program provided with the test, based on the student’s age. Standard scores are reported for each student and used for analysis.

A second scorer familiar with the Writing Fluency subtest also scored 50% of the subtests to ensure accuracy of scoring. The percentage of interrater agreement was calculated based on raw scores. Interrater agreement on the Writing Fluency subtest was 100%.

On-task behavior scoring. Behavior data was collected during 52.20% of all instructional sessions. An experienced observer was trained to code behavior data. The observer coded student behavior for all selected instructional sessions, and the researcher coded behavior for 50% of the selected instructional sessions to ensure reliability of
coding. Codes were compared to calculate interrater agreement using the interval agreement approach (Kennedy, 2005). Interrater agreement was calculated after each session was coded to ensure high rates of agreement. If interrater agreement fell below 80% for any session, it would be recoded. This did not occur during the course of the study. Interrater agreement for behavioral observations was 98.26%.

Social validity interview coding. Individual interviews were conducted with all participants during the revision posttesting phase to provide feedback about the usefulness and acceptability of instruction. Interviews were transcribed and coded using inductive coding. Students’ responses were compared to identify themes that emerged in response to each question. Additionally, responses across questions were analyzed to look for overall themes in each student’s reactions to instruction.

Fidelity of treatment scoring. Trained observers watched videotapes of all lessons and completed fidelity of treatment checklists, indicating which lesson components were completed. The number of instructional components completed during each lesson was summed and divided by the total number of lesson components to determine the percentages of instructional components completed. For 50% of instructional sessions, a second scorer also assessed fidelity of treatment to ensure reliability of scoring. When there were discrepancies between the two scorers, averages of the scores were used. The percentage of interrater agreement for fidelity of treatment scoring was 89.80%.

Peer-conferencing steps completed scoring. The percentage of revision strategy steps completed by each student was calculated for 100% of the peer-revision
conferences. The primary scorer used videotapes, audio tapes, and student work samples to determine which revision strategy steps were completed by each student. A second scorer observed 40% of the revision lessons and recorded the steps each student completed. When there were discrepancies between the two scorers, averages of the scores were used. The percentage of interrater agreement for peer-conferencing steps completed was 94.90%.

Strategy knowledge questioning. The strategy knowledge questionnaire was used weekly during instruction as a curriculum-based assessment to measure students’ knowledge of the SRSD and revision strategies. The measure was scored by assigning one point for each component correctly identified. For the parts of a good persuasive essay, one point was given for each of the following parts, for a total of ten possible points: (a) pick my idea, (b) organize my notes, (c) write and say more, (d) topic sentence, (e) reasons, (f) explanations, (g) counter reason(s), (h) explanation for the counter reasons, (i) refuting the counter reason, and (j) an ending sentence. Similarly, for the revision question, students could earn a total of eight points for (a) opening compliment, (b) parts, (c) clarity, (d) persuasiveness, (e) ending compliment, (f) capitalization, (g) punctuation, and (h) spelling. The researcher scored student responses on an ongoing basis throughout the instructional phases.

Design

This study employed a single-subject multiple baseline, multiple probe across participants design (Kennedy, 2005). A single-subject design was selected because of the limited number of students with EBD served in the targeted setting. A multiple baseline
approach was necessary because the skills being taught could not be reversed (Kennedy, 2005). That is, once students learned the writing and revision strategies, the skills could not be taken away. Therefore, other single-subject designs were not appropriate for this study. In the multiple baseline model, each student’s baseline performance acts as his or her own control, with a student’s performance after instruction compared to his or her performance on the baseline measures (Horner et al., 2005).

Students were instructed in small groups consisting of two to four students, with a total of four groups participating in the study. There were six phases to the study: (a) baseline testing, (b) self-regulated strategy development (SRSD) instruction, (c) SRSD posttesting, (d) revision instruction, (e) revision posttesting, and (f) maintenance testing. During the baseline phase, students were administered a series of dependent measures, including three persuasive essay writing prompts, the Self-efficacy for Persuasive Writing Scale, and the Woodcock Johnson III Writing Fluency subtest. Once student performance on essay writing stabilized, the instructional phases began. Number of essay parts written was the measure used to determine stability of baseline.

During the second phase of the study, students learned the SRSD strategy for persuasive writing: POW+TREE (Harris et al., 2008). Previous studies have used this approach to successfully teach writing to middle school students with EBD (Mastropieri et al., 2009; 2010). Three post-strategy instruction essay prompts were given during the SRSD posttesting phase. These writing prompts were given before students began revision instruction for two reasons. First, the scores obtained following SRSD instruction can be directly compared to the scores of students in previous studies who
have learned the same strategy. Second, these scores can be used to evaluate the additive impact of revision instruction on students’ writing performance above and beyond the impact of SRSD instruction in planning and drafting.

The fourth phase of the study focused on instruction in revision. Students learned a peer-revision strategy for improving their persuasive essays. Three more post-instruction essay prompts were given during the revision posttesting phase. At that time, other dependent measures were also administered, including (a) Self-efficacy for Persuasive Writing Scale, (b) Woodcock Johnson III Writing Fluency subtest, (c) a generalization writing prompt, and (d) a social validity interview.

The sixth and final phase of the study was the maintenance phase. Six to eight weeks after all posttesting was completed, students were administered a writing prompt to assess maintenance of their persuasive writing skills, and again completed the Self-efficacy for Persuasive Writing Scale.
Summary

This single subject, multiple probes across participants study included 10 eighth-grade students with EBD who received instruction in self-contained language arts classes. The study included six phases: (a) baseline, (b) SRSD instruction in persuasive writing, (c) SRSD posttesting, (d) revision instruction, (e) revision posttesting, and (f) maintenance. Dependent measures included student essays, which were scored on a number of variables, including length, text structure elements, quality, mechanical errors, and revision. Additionally, students were assessed on their self-efficacy for persuasive writing, writing fluency, and their perceptions of the SRSD and revision strategies. On-task behavior data was collected during teacher-led and student-led instruction and during independent work time. Finally, strategy knowledge questioning was completed weekly to document student learning.

After obtaining permission from the school district and university, permission was sought for all students in the targeted language arts classes. Parent consent and student assent were acquired before baseline testing began. The SRSD instructional phase consisted of five lessons administered over eight, 50-minute sessions (Group 1 required nine instructional sessions). Following SRSD posttesting, students received revision strategy instruction. Revision instruction consisted of five lessons administered over from seven to ten, 50-minute sessions. Revision posttesting was completed immediately after instruction; and maintenance testing was completed six to eight weeks after posttesting ended.
4. RESULTS

This chapter begins with fidelity of treatment results, followed by analysis of dependent measures. Dependent measures were analyzed using traditional visual analysis, effect sizes, statistical analysis, and qualitative analysis. A summary of results is provided for each dependent measure.

Fidelity of Treatment

The overall fidelity of treatment was 94.79%. Fidelity of treatment, across groups, for the SRSD instructional phase was 98.94%; for the revision instructional phase it was 88.24%. The lower fidelity of treatment score for the revision instructional phase was primarily due to several elements of lesson 2 (editing) inadvertently being omitted during instruction for Group 4. The other lesson component that was omitted on several occasions was the end-of-lesson wrap-up during which the teacher reviewed what the students had learned in class that day. All other elements of instruction were included. Fidelity of treatment scores for each group during each instructional phase is provided in Table 3.
Table 3

*Fidelity of Treatment Percentages by Group*

<table>
<thead>
<tr>
<th></th>
<th>SRSD Instruction</th>
<th>Revision Instruction</th>
<th>Total Fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group 1</strong></td>
<td>100%</td>
<td>96.67%</td>
<td>98.72%</td>
</tr>
<tr>
<td><strong>Group 2</strong></td>
<td>95.74%</td>
<td>93.10%</td>
<td>94.74%</td>
</tr>
<tr>
<td><strong>Group 3</strong></td>
<td>100%</td>
<td>93.34%</td>
<td>97.44%</td>
</tr>
<tr>
<td><strong>Group 4</strong></td>
<td>100%</td>
<td>70.00%</td>
<td>88.00%</td>
</tr>
<tr>
<td><strong>Total Fidelity</strong></td>
<td>98.94%</td>
<td>88.24%</td>
<td>94.79%</td>
</tr>
</tbody>
</table>

Data Analysis Methods

Student performance on dependent measures was analyzed during each testing phase – baseline testing, SRSD posttesting, revision posttesting, generalization, and maintenance testing. Four analysis methods were used: (a) visual analysis, (b) percent of nonoverlapping data (PND), (c) statistical tests, and (d) qualitative analysis. Visual analysis, PND, and nonparametric statistics were used to evaluate student performance on persuasive essays. Statistical tests were also used to analyze student performance on the Persuasive Writing Self-efficacy Scale, Woodcock-Johnson Fluency Subtest, and on-task behavior measures. Qualitative analysis was conducted on student interview data. These
techniques are described below. Additionally, means and standard deviations were calculated to summarize performance on all measures, with the exception of student interview data.

Visual analysis. Visual inspection techniques, as described by Kennedy (2005), were employed to analyze essay data for the essay parts, holistic quality, essay length, and number of mechanical errors measures. Data points were plotted on a graph with time of measure on the x axis, and score on the y axis. Lines were drawn between data points for each student within each phase. Then, within-phase patterns were analyzed, including level, trend, and variability. Level refers to the average score of the data within the phase. Trend is the slope and magnitude of the data, and is analyzed by looking at the line of data points for each student within a phase. The slope of the data line within a phase may be positive, with scores improving over time; the slope may be negative, with scores declining over time within a phase; or the slope may be flat, indicating consistent performance across data points within a phase. The magnitude of the line may be large (i.e., a steep slope) or small (i.e., a flatter slope), indicating how much growth or decline occurred within a phase. Taken together, then, slope and magnitude give a sense of the growth and movement of student performance within a phase. Finally, variability refers to how much individual data points vary from the overall trend of the data. If a straight line were drawn connecting the first and last data point within a phase, how closely a student’s data points matched the line indicates consistency of performance. Data points that fall along the line show consistent performance. The farther data points fall away from that line indicate a lack of consistency, or variability, in performance.
Additionally, two between-phase patterns were analyzed: immediacy of effect and overlap of data points between phases. Immediacy of effect refers to how quickly student performance changed between phases. That is, effect of instruction is immediate if the first data point in the following phase is noticeably higher or lower than data points during the previous phase. If, on the other hand, the first data point does not indicate a growth or decline in performance, but the growth or decline occurs later in the testing phase, there would be little immediacy of effect of the intervention. Lastly, overlap of data points refers to how similar performance is across phases, and is described in more detail in the next section on PND.

PND. PND was calculated in evaluating the magnitude of effectiveness of the SRSD and revision interventions (Scruggs & Mastropieri, 2001). PND is calculated by dividing the number of nonoverlapping data points between two phases by the total number of data points in the phase. Guidelines provided by Scruggs and Mastropieri (2001) suggest that PND values greater than 90% indicate very effective interventions. PND values between 70% and 90% indicate effective interventions, PND values between 50% and 70% indicate interventions that are questionable, and PND values less than 50% indicate ineffective interventions. PND was calculated for all essay measures across all phases of the study.

Statistical tests. Nonparametric statistics were used to compare results across phases. Nonparametric statistics were selected because they do not require the same underlying assumptions of homogeneity of variance and a normal distribution that parametric statistics require (e.g., t-test, ANOVA) – assumptions that were not met in this
Specifically, the Wilcoxon Matched-Pairs, Signed Ranks test was selected to analyze results across phases because the study employed related samples and difference scores across phases could be calculated (Seigel & Castellan, 1988). The significance level was set at 0.05 for all analyses. The Wilcoxon Matched-Pairs, Signed Ranks test was used to analyze results on the persuasive essay, Persuasive Writing Self-efficacy Scale, and on-task behavior measures.

The W-J III Writing Fluency Subtest was analyzed using a parametric statistical test. This was done because the W-J III is a standardized test and, as such, meets all of the underlying assumptions required of parametric tests: normality, independence of scores, and homogeneity of variance. Specifically, students’ pre- and post-scores were compared using a Paired t-Test.

Qualitative analysis. Qualitative analysis was used to analyze information from student interviews. This was done by comparing and connecting student responses across questions and across students (Maxwell, 2005). First, interview data was transcribed from audio recordings of the interview sessions. Next, student interviews were reviewed individually to look for themes in each student’s perceptions of SRSD and revision instruction. Then, student responses were compiled by interview questions to look for similarities and differences across participants. Similarities in responses were grouped for reporting, while divergent responses were further analyzed to determine why some students responded differently to the instruction provided.

Persuasive Essays. Persuasive essays were scored for length, number of essay parts, overall quality, and number of mechanical errors (specifically, punctuation,
capitalization, and spelling). Overall means for each essay characteristic are provided on Table 4, with individual student scores on Table 5. Essay performance measures are organized by testing phase. For each phase, overall results are presented, followed by detailed results for each instructional group and for individual students.
Table 4

Mean Essay Scores for All Groups

<table>
<thead>
<tr>
<th></th>
<th>Baseline (SD) (N=10)</th>
<th>Post-SRSD (SD) (N=10)</th>
<th>Post-Revision (SD) (N=10)</th>
<th>Generalization (SD) (N=10)</th>
<th>Maintenance (SD) (N=9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Essay Parts</td>
<td>5.49 (1.76)</td>
<td>10.42 (3.05) a</td>
<td>10.13 (1.89) a</td>
<td>10.30 (2.21) a</td>
<td>10.44 (2.19) a</td>
</tr>
<tr>
<td>Holistic Quality</td>
<td>3.63 (0.91)</td>
<td>7.73 (1.65) a</td>
<td>7.77 (1.79) a</td>
<td>8.30 (2.00) a</td>
<td>8.44 (2.07) ab</td>
</tr>
<tr>
<td>Number of Words</td>
<td>148.21 (115.80)</td>
<td>210.10 (103.77) a</td>
<td>207.77 (101.39) a</td>
<td>224.00 (100.07) a</td>
<td>227.33 (79.90) ab</td>
</tr>
<tr>
<td>Mechanical Errors c</td>
<td>7.65 (4.49)</td>
<td>4.07 (3.09) d</td>
<td>3.15 (2.26) d</td>
<td>3.17 (2.40) d</td>
<td>3.17</td>
</tr>
<tr>
<td>Punctuation Errors c</td>
<td>0.90 (0.82)</td>
<td>0.72 (0.77)</td>
<td>0.62 (0.84)</td>
<td>0.49 (0.53)</td>
<td>0.16 (0.33) ab</td>
</tr>
<tr>
<td>Capitalization Errors c</td>
<td>2.02 (2.86)</td>
<td>0.78 (1.07)</td>
<td>0.65 (1.09)</td>
<td>0.65 (0.83)</td>
<td>0.65</td>
</tr>
<tr>
<td>Spelling Errors c</td>
<td>3.63 (3.12)</td>
<td>4.84 (3.97) a</td>
<td>2.66 (2.49) d</td>
<td>2.00 (1.77)</td>
<td>2.36 (1.86)</td>
</tr>
</tbody>
</table>

Note. a Significantly different than baseline, p < .05, according to the Wilcoxon Matched-Pairs, Signed Ranks test. b Significantly different than post-revision instruction, p < .05, according to the Wilcoxon Matched-Pairs, Signed Ranks test. c Proportion of errors per 100 words. d Significantly different than post-SRSD instruction, p < .05, according to the Wilcoxon Matched-Pairs, Signed Ranks test.
| Group | Baseline | | | | Post-SRSD | | | | | Post-Revision | | | | | Maintenance | | |
|-------|----------|---|---|---|----------|---|---|---|---|----------|---|---|---|---|----------|---|
|       | Word (SD) | Part (SD) | Qual (SD) | Word (SD) | Part (SD) | Qual (SD) | Mech (SD) | Word (SD) | Part (SD) | Qual (SD) | Mech (SD) | Word (SD) | Part (SD) | Qual (SD) | Mech (SD) | Word (SD) | Part (SD) | Qual (SD) | Mech (SD) |
|       |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
| Group 1 | 121.89 (92.32) | 5.11 (.150) | 3.89 (.84) | 207.67 (56.05) | 9.44 (.69) | 7.33 (.88) | 6.23 (4.17) | 204.94 (53.45) | 9.22 (1.07) | 7.89 (1.39) | 3.52 (4.35) | 205.67 (50.41) | 10.33 (1.15) | 8.0 (.173) | .95 (.33) | 139.0 | 9.0 | 6.0 | .72 |
| Andy | 228.33 | 5.00 | 4.00 | 266.33 | 9.67 | 6.6 | 3.53 | 173.67 | 8.00 | 6.33 | 1.85 | 139.0 | 9.0 | 6.0 | .72 |
| Carl | 73.67 | 6.67 | 4.67 | 202.00 | 10.00 | 8.33 | 11.05 | 266.67 | 9.67 | 8.33 | 8.47 | 225.0 | 11.0 | 9.0 | 1.33 |
| Tim | 63.67 | 3.67 | 3.00 | 154.67 | 8.67 | 7.00 | 4.13 | 174.50 | 10.00 | 9.00 | .26 | 253.0 | 11.0 | 9.0 | .79 |
| Group 2 | 316.50 (147.79) | 7.75 (.35) | 4.21 (.77) | 305.67 (169.71) | 9.67 (1.41) | 7.50 (1.64) | 5.48 (1.32) | 356.00 (132.94) | 9.83 (.24) | 8.33 (.47) | 4.20 (4.49) | 378.0 | 10.0 | 10.0 | 2.12 |
| Ally | 421.00 | 8.00 | 3.67 | 425.67 | 10.67 | 8.67 | 6.41 | 450.00 | 10.00 | 8.67 | 1.02 |
| John | 212.00 | 7.50 | 4.75 | 185.67 | 8.67 | 6.33 | 4.55 | 262.00 | 9.67 | 8.00 | 7.38 | 378.0 | 10.0 | 10.0 | 2.12 |
| Group 3 | 143.63 (27.40) | 6.33 (.94) | 4.21 (.77) | 276.83 (21.45) | 15.67 (4.7) | 10.00 (0.00) | 5.88 (5.83) | 192.50 (29.93) | 12.67 (1.89) | 9.50 (.71) | 2.20 (1.97) | 268.0 | 13.0 | 10.0 | 3.82 |
| Mario | 124.25 | 7.00 | 4.75 | 261.67 | 15.33 | 10.00 | 10.01 | 213.67 | 11.33 | 9.00 | 3.59 | 260.0 | 12.0 | 10.0 | 6.54 |
| Navid | 163.00 | 5.67 | 3.67 | 292.00 | 16.00 | 10.00 | 1.76 | 171.33 | 14.00 | 10.00 | .81 | 276.0 | 14.0 | 10.0 | 1.09 |
| Group 4 | 65.39 (29.25) | 3.81 (.91) | 2.58 (.22) | 104.33 (17.67) | 8.39 (2.11) | 6.78 (1.84) | 11.70 (4.46) | 121.94 (29.26) | 9.56 (2.22) | 7.11 (2.17) | 5.76 (1.91) | 171.67 (59.91) | 9.0 | 7.33 (2.89) | 5.31 (.56) |
| Brian | 69.00 | 4.33 | 2.67 | 116.00 | 10.67 | 8.67 | 13.17 | 130.00 | 10.67 | 9.00 | 3.87 | 189.0 | 11.0 | 9.0 | 5.29 |
| Clare | 92.67 | 4.33 | 2.33 | 84.00 | 8.00 | 6.67 | 6.69 | 146.33 | 11.00 | 8.33 | 5.70 | 221.0 | 10.0 | 9.0 | 5.88 |
| Mary | 34.50 | 2.75 | 2.75 | 113.00 | 6.50 | 5.00 | 15.23 | 89.50 | 7.00 | 4.00 | 7.70 | 105.0 | 6.0 | 4.0 | 4.76 | Kate |
Baseline. During baseline, students were administered persuasive essay prompts. Each prompt was administered over two days of testing. On the first day, students were given a prompt sheet containing two essay prompt questions. Students were instructed to select one prompt and “write a good persuasive essay.” Students were provided with Neo word processors on which to write their essays. On the second day, students were given the prompt sheets, a hard copy of the essay they had written the day before, and the Neo that contained the essay file. Then, students were instructed to revise their essays. No further instructions were given. After students revised their essays, a hard copy was printed to turn in. Both hard and soft copies of all essays were collected for scoring.

The essay parts measure was used to establish a stable baseline (Little et al., 2010). In order for baseline scores to be considered stable, students had to complete three essays on which the range of number of essay parts was no greater than two. Using this measure as a guideline, Group 1 was administered three essay prompts, while Groups 2, 3 and 4 were administered four essay prompts to establish a stable baseline. Although four baseline observations were included for these groups within the multiple baseline design, there were incomplete baseline data within each group for some individual students. In spite of this, all students within each group met stable baseline criteria in the essay parts before proceeding to treatment phases.

Overall baseline performances on persuasive essays were variable, as can be seen in Figures 1, 2, and 3. At baseline testing, the overall mean number of essay parts across students was 5.49 ($SD = 1.76$; see Table 4). The overall mean score for holistic quality was 3.63 ($SD = 0.91$) out of a possible 10 points. At baseline, the average essay length
was 148.21 words ($SD = 115.80$). The total number of mechanical errors per 100 words could not be calculated during baseline testing because capitalization errors could not be counted. (The automatic error correction feature on Microsoft Word corrected all capitalization errors. This problem was not identified until the end of baseline testing, at which time the automatic correction feature was disabled.) However, mean scores for number of punctuation errors per 100 words ($M = 0.90, SD = 0.82$), and number of spelling errors per 100 words ($M = 3.63, SD = 3.12$) were calculated. Group and individual student baseline performance on all persuasive essay measures is discussed below.
Figure 1. Number of essay parts by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing.
Figure 2. Holistic quality by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing.
Figure 3. Essay length by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing.
Group 1. Andy, Carl and Tim were in Group 1. Each student was administered three baseline essay probes. Students in Group 1 had a mean of 5.11 ($SD = 1.50$) persuasive essay parts per essay, with a mean holistic quality score of 3.89 ($SD = 0.84$). The mean length of students’ essays was 121.89 words ($SD = 92.32$). Individual student performance is presented next.

Andy baseline essay performance. Andy completed three baseline essays. His performance was consistently low. He included an average of 5 essay parts (range = 3-6), with an increasing slope from essay 1 to essay 2, and stabilized performance between essay 2 and essay 3. Essay parts scores on the three essays were 3, 6, and 6. Andy’s holistic quality scores were 3, 5, and 4 over the course of the three baseline essays ($M = 4$, range 3-5), indicating slight variability in performance across essays. The number of words written in Andy’s essays was highly variable, with a steep increasing slope between essays 1 and 2, and a steep decreasing slope between essays 2 and 3 (essay 1 = 176 words, essay 2 = 321 words, and essay 3 = 188 words). Andy had a mean of 228.33 words per essay. The difference in essay length did not correspond to differences in number of essay parts or holistic quality. For example, Andy’s second baseline essay, which contained 321 words, had six essay parts and a holistic quality score of five. His third baseline essay, which contained only 188 words, obtained similar scores with six essay parts and a holistic quality score of four. Finally, at baseline testing Andy had a relatively small number of punctuation errors ($M = 0.59$ per 100 words) and spelling errors ($M = 0.88$ per 100 words).

Carl baseline essay performance. Carl completed three baseline essays. He
included an average of 6.67 essay parts (range = 6-8) in his essays, with a decreasing slope from essay 1 to essay 2, and stabilizing between essay 2 and essay 3. Essay parts scores on the three essays were 8, 6, and 6. Carl’s holistic quality scores were 6, 4, and 4 over the course of the three baseline essays ($M = 4.67$, range 4-6). Again, there was a slight decreasing slope between essay 1 and essay 2, and stable performance between essay 2 and essay 3. Carl was able to write a one-paragraph essay with a topic sentence, two reasons, two explanations, and an ending. In his first baseline essay, Carl was able to provide a third reason, which, in turn, increased his holistic quality score. The length of Carl’s essays was consistent, showing only a slight increasing slope across essays. He had an average of 73.67 words per essay (essay 1 = 60 words, essay 2 = 79 words, and essay 3 = 82 words). While Carl did not make any punctuation errors in his baseline essays, he made a large number of spelling errors ($M = 8.63$ per 100 words).

**Tim baseline essay performance.** Tim completed three baseline essays. He included an average of 3.67 essay parts (range = 3-4) in his essays, with consistent performance from essay 1 to essay 2, and a slight decreasing slope between essay 2 and essay 3. Essay parts scores on the three essays were 4, 4, and 3. Tim’s holistic quality scores were 4, 2, and 2 over the course of the three baseline essays ($M = 2.67$, range 2-4). There was a slight decreasing slope between essay 1 and essay 2, and stable performance between essay 2 and essay 3. This variability occurred because Tim included an ending sentence in his first essay, which he did not include in other essays, thereby improving his quality score. The length of Tim’s essays had a steady decreasing slope throughout baseline testing (essay 1 = 90 words, essay 2 = 61 words, and essay 3 = 40 words) for a
mean essay length of 63.67 words. He had very few punctuation errors ($M = 0.37$ errors per 100 words) and no spelling errors in his baseline essays.

**Group 2.** Ally and John were in Group 2. Both students were administered four baseline essay probes. However, Ally refused to complete a fourth baseline essay. She did not respond to the given prompt, and did not write a persuasive essay, so her response was not scored. An attempt was made to administer a fifth baseline essay prompt, but Ally again did not respond to the given prompt. Therefore, only three baseline data points are available for Ally. Students in Group 2 had a mean of 7.75 ($SD = 0.35$) persuasive essay parts per essay, with a mean holistic quality score of 4.21 ($SD = 0.7$). The mean length of students’ essays was 316.50 words ($SD = 147.79$). Individual student performance is presented next.

**Ally baseline essay performance.** Ally completed three baseline essay prompts. Of all participants, Ally showed the most variability in performance at baseline testing. She included an average of 8.00 essay parts (range = 6-11) in her essays, with a sharply decreasing slope from essay 1 to essay 2, and a continued decrease between essay 2 and essay 3. Essay parts scores on the three essays were 11, 7, and 6. Likewise, the length of her essays was highly variable. She had an average of 421.00 words per essay (range = 227-592 words; essay 1 = 444 words, essay 2 = 592 words, and essay 3 = 227 words). The number of essay parts was not related to the length of her essays. For example, her second essay was the longest, at nearly 600 words, yet it contained only 7 persuasive essay parts. Additionally, the structure of Ally’s essays was not sound. She consistently obtained low holistic quality scores ($M = 3.67$, range = 3-4), relative to the length of her
essays and the number of essay parts they contained. Ally’s holistic quality scores were 4, 4, and 3 over the course of the three baseline essays, showing consistently low performance on the quality measure. Finally, Ally had a small number of punctuation ($M = 0.87$ errors per 100 words) and spelling errors ($M = 1.31$ errors per 100 words) in her baseline essays.

**John baseline essay performance.** John completed four baseline essays. He included an average of 7.50 essay parts (range = 6-10) in his essays, with an increasing slope from essay 1 to essay 2, and a corresponding decreasing slope between essay 2 and essay 4. Essay parts scores on the four essays were 7, 10, 7, and 6. In all baseline essays, John provided a topic sentence, then simply listed reasons. His essays sometimes included an ending sentence, and sometimes they did not. Although his essays had many parts, the overall quality of the essays was not strong. John’s holistic quality scores were 3, 6, 6, and 4 over the course of the four baseline essays ($M = 4.75$, range 3-6). There was an increasing slope between essay 1 and essay 2, and a decreasing slope between essay 3 and essay 4. In spite of the variability in parts and quality of John’s essays, there was no variability the length of his essays. He had an average of 212.00 words per essay (range = 206-217; essay 1 = 217 words, essay 2 = 211 words, essay 3 = 206 words, and essay 4 = 214). Finally, John had minimal punctuation errors in his baseline essay ($M = 0.36$ errors per 100 words), but made a number of spelling errors ($M = 3.07$ per 100 words).

**Group 3.** Navid and Mario were in Group 3. Each student was administered four essay prompts, but Navid completed only three. Students in Group 3 had a mean of 6.33
persuasive essay parts per essay, with a mean holistic quality score of 4.21 (SD = 0.77). The mean length of students’ essays was 143.63 (SD = 27.40) words.

Individual student performance is presented next.

**Navid baseline essay performance.** Navid completed three baseline essays. He included an average of 5.67 essay parts (range = 5-6) in his essays, with a slight increase from essay 1 to essay 2. Essay parts scores on the three essays were 5, 6 and 6. Navid’s holistic quality scores were 4, 4, and 3 over the course of the three baseline essays ($M$ = 3.67, range 3-4). There was a slight decreasing slope between essay 2 and essay 3. The length of Navid’s essays showed a decreasing trend throughout the baseline phase, with the steepest decrease in slope between essay 1 and essay 2. He had an average of 163.00 words per essay (essay 1 = 259 words, essay 2 = 121 words, and essay 3 = 109 words).

In spite of the length of the first essay, Navid did not include more essay parts or have a higher quality essay than on his other baseline essays. Finally, Navid’s baseline essays contained both punctuation errors ($M$ = 2.50 errors per 100 words) and spelling errors ($M$ = 2.29 errors per 100 words).

**Mario baseline essay performance.** Mario completed four baseline essays. He included an average of 7.00 essay parts (range = 5-11) in his essays, with a sharp increase in slope from essay 2 to essay 3, and a corresponding sharp decrease in slope between essay 3 and essay 4. Essay parts scores on the three essays were 5, 6, 11, and 6. Mario’s holistic quality scores were 5, 4, 6, and 4 over the course of the four baseline essays ($M$ = 4.75, range 4-6). While performance on the quality measure was variable from essay to essay, the range of variability was narrow, with only 2 points separating the lowest score
from the highest score. Mario had an average of 124.25 words per essay (range = 88 – 173; essay 1 = 130 words, essay 2 = 130 words, and essay 3 = 173 words, and essay 4 = 88 words). The length of his essays showed an increase in slope between essay 2 and essay 3, with an even larger decrease in slope between essay 3 and essay 4. Mario’s longest essay (essay 3) was also the essay with the most parts and highest holistic quality. However, he was not able to consistently perform at this high level. Finally, Mario made a few punctuation errors on his baseline essays ($M = 1.06$ errors per 100 words), and a large number of spelling errors ($M = 6.38$ errors per 100 words).

**Group 4.** Brian, Clare, and Mary Kate were in Group 4. Each student was administered four essay prompts. Brian refused to complete the second baseline essay and, therefore, only had three baseline essay data points. Additionally, Clare established stable baseline performance on her third essay, and was not required to complete a fourth essay. Students in Group 4 had a mean of $3.81$ ($SD = 0.91$) persuasive essay parts per essay, with a mean holistic quality score of $2.58$ ($SD = 0.22$). The mean length of students’ essays was $65.39$ words ($SD = 29.25$). Individual student performance is presented next.

**Brian baseline essay performance.** Brian completed three baseline essays. He included an average of 4.33 essay parts (range = 4-5) in his essays, with a slight increasing slope between essay 2 and essay 3. Essay parts scores on the three essays were 4, 4, and 5. Brian’s holistic quality scores were 3, 3, and 2 over the course of the three baseline essays ($M = 2.67$, range 2-3). There was a slight decreasing slope on the quality measure between essay 2 and essay 3. Brian had an average of 69.00 words per
essay (range = 52-93; essay 1 = 52 words, essay 2 = 62 words, and essay 3 = 93 words). There was an increasing slope throughout baseline testing on the length measure. No consistent relationship between the length of Brian’s essays and the quality of those essays was evident. Brian had a few punctuation errors ($M = 1.36$ errors per 100 words) and spelling errors ($M = 1.64$ errors per 100 words) in his baseline essays.

**Clare baseline essay performance.** Clare completed three baseline essays. She included an average of 4.33 essay parts (range = 4-5) in her essays, with a small decrease in slope from essay 1 to essay 2. Essay parts scores on the three essays were 5, 4 and 4. Clare’s holistic quality scores were 6, 4, and 4 over the course of the three baseline essays ($M = 4.67$, range 4-6). Again, there was a slightly decreasing slope between essay 1 and essay 2, and stable performance between essay 2 and essay 3. Clare had an average of 92.67 words per essay (range = 45-133; essay 1 = 133 words, essay 2 = 100 words, and essay 3 = 45 words), with a steadily decreasing slope throughout the phase. Regardless of essay length, the number of parts and holistic quality showed little variation. Finally, Clare made a number of errors in both punctuation ($M = 1.91$ errors per 100 words) and spelling ($M = 3.98$ errors per 100 words) during baseline testing.

**Mary Kate baseline essay performance.** Mary Kate completed four baseline essays. She included an average of 2.75 essay parts (range = 2-5) in her essays. There was a decrease in slope from essay 1 to essay 2, and performance was stable on the remaining baseline essays. Essay parts scores on the three essays were 5, 2, 2 and 2. While her first essay contained five parts (topic, three reasons, and an ending), each subsequent essay had only two persuasive essay parts (a topic and one reason) in each
essay. Therefore, Mary Kate received a much higher holistic quality score on her first essay, compared to the remaining three baseline essays. Mary Kate’s holistic quality scores were also 5, 2, 2, and 2 over the course of the four baseline essays ($M = 2.75$, range 2-5). Again, there was a decrease slope between essay 1 and essay 2, and stable performance on essays 2, 3 and 4. Mary Kate had an average of 34.50 words per essay (range = 17-54; essay 1 = 54 words, essay 2 = 17 words, essay 3 = 42 words, and essay 4 = 25 words), with variability in length across all essays. Finally, Mary Kate had few punctuation errors at baseline testing ($M = 0.25$ errors per 100 words). However, she did have a large number of spelling errors ($M = 9.05$, range = 4.0-17.65).

SRSD Posttesting. During SRSD posttesting, students were administered three persuasive essay prompts. Administration procedures were identical to baseline testing procedures. Overall, students’ scores increased on the essay length, essay parts, and holistic quality measures following SRSD instruction (see Table 4). At SRSD posttesting, the overall mean number of essay parts across students was 10.42 ($SD = 3.05$; see Table 4). The overall mean score for holistic quality was 7.73 out of a possible 10 points ($SD = 1.65$); the mean essay length was 210.10 words ($SD = 103.77$).

At SRSD posttesting, the mean number of mechanical errors per 100 words was 7.65 ($SD = 4.49$). When analyzed by type of error, the number of punctuation errors decreased slightly to 0.72 per 100 words ($SD = 0.77$) following instruction in SRSD. On the other hand, the mean number of spelling errors increased at SRSD posttesting to 4.84 errors per 100 words ($SD = 3.97$). Finally, SRSD posttesting yielded a mean capitalization error rate of 2.02 errors per 100 words ($SD = 2.86$). Graphic
representations of each student’s score for number of essay parts (Figure 1), holistic quality (Figure 2), essay length (Figure 3), and number of mechanical errors (Figure 4) are provided.
Figure 4. Mechanical errors per 100 words by student during baseline, SRSD posttesting, revision posttesting, generalization, and maintenance testing.
SRSD posttesting scores were entered into a Wilcoxon Matched-Pairs, Signed Ranks test, which indicated significant differences between baseline and SRSD posttesting performance on the number of essay parts ($p = 0.005$), holistic quality ($p = 0.005$), essay length ($p = 0.022$), and number of spelling errors measures ($p = 0.028$). No significant differences were found in the number of punctuation errors per 100 words ($p = 0.445$). In addition, mean PND scores at SRSD posttesting indicate that SRSD instruction had an overall positive effect on student performance (Overall PND = 75.56 %; see Table 6). The magnitude of the effect varied by type of measure. SRSD instruction had a moderate effect on the number of essay parts students included in their essays (PND = 76.77%), a large effect on the holistic quality of their essays (PND = 90.00%), and little effect on the length of students’ essays (PND = 60.00%). There was variability in the PND scores of individual students, which is discussed in the following sections.
Table 6

Percent of Nonoverlapping Data Across Phases by Student

<table>
<thead>
<tr>
<th></th>
<th>Baseline to SRSD Posttesting</th>
<th>Revision to Baseline Posttesting</th>
<th>Baseline to Maintenance Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ally</td>
<td>0.00%</td>
<td>100%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Andy</td>
<td>100%</td>
<td>66.67%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Brian</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Carl</td>
<td>66.67%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Clare</td>
<td>100%</td>
<td>100%</td>
<td>0.00%</td>
</tr>
<tr>
<td>John</td>
<td>0.00%</td>
<td>33.33%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Mario</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Mary Kate</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Navid</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Tim</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Mean PND</td>
<td>76.67%</td>
<td>90.00%</td>
<td>60.00%</td>
</tr>
</tbody>
</table>

*Note:* PND scores for mechanical errors per 100 words are compared to SRSD posttesting levels, as no score for mechanical errors per 100 words could be calculated at baseline testing.
Group and individual student SRSD posttest performance on all persuasive essay measures is discussed below.

**Group 1.** Students in Group 1 (Andy, Carl, and Tim) showed an increase in scores on all persuasive essay measures following instruction in SRSD. Group 1 had a mean of 9.44 ($SD = 0.69$) persuasive essay parts per essay, an increase of 4.33 essay parts over baseline testing levels. Additionally, their SRSD posttesting mean holistic quality score was 7.33 ($SD = 0.88$), an increase of 3.44 points over baseline. Similarly, the mean length of students’ essays increased to 207.67 ($SD = 56.05$) words, which represented an increase of 85.78 words over baseline levels. The mean total number of mechanical errors per 100 words was 6.23 ($SD = 4.17$). This measure was not calculated at baseline.

Overall PND for Group 1 was 81.48%, indicating that SRSD instruction was an effective intervention for these students. More specifically, SRSD was effective at increasing the number of persuasive essay parts students included in their essays (PND = 88.89%), and the holistic quality of their essays (PND = 88.89%). SRSD instruction had a questionable effect on the length of the essays written by students in Group 1 (PND = 66.67%). Individual student performance is presented next.

**Andy SRSD posttest essay performance.** Andy completed three SRSD posttest essays. He showed some improvement on his persuasive essay scores, but the improvements steadily decreased over time. He included a mean of 9.67 essay parts (range = 7-14) in his essays. This represents an immediate level change of 4.67 parts from baseline testing, when Andy’s mean number of essay parts was 5.00. However,
there was a steep decreasing slope from essay 1 to essay 2 on the essay parts measure. Essay parts scores on the three essays were 14, 7 and 8.

Andy’s holistic quality scores were 9, 7, and 4 over the course of the three SRSD posttest essays ($M = 6.60$, range 4-9), for an overall level change of 2.6 from baseline to SRSD posttesting. Again, there was a steep decrease in slope throughout the phase. Similarly, Andy had a mean essay length of 266.33 words, a slight level increase of 28 words from baseline testing. However, the length of his essays showed a decreasing slope throughout the SRSD posttest phase (range = 226-312; essay 1 = 312 words, essay 2 = 261 words, and essay 3 = 226 words).

Andy had an average of 3.53 mechanical errors per 100 words (range = 2.21-6.13) at SRSD posttesting, with variable results across essays (essay 1 = 2.24 errors per 100 words, essay 2 = 6.13 errors per 100 words, and essay 3 = 2.21 errors per 100 words). These mechanical errors were almost entirely spelling errors.

Andy’s mean SRSD posttest scores indicate that his essays improved overall following SRSD instruction. PND scores, however, indicate more variability in his improvements across essay measures from baseline to SRSD posttesting. SRSD instruction had a large effect on the number of essay parts Andy included in his essays (PND = 100%), a small effect on the quality of Andy’s essays (PND = 66.67%), and no effect on the length of Andy’s essays (PND = 0.00%).

**Carl SRSD posttest essay performance.** Carl completed three SRSD posttest essays. He showed immediate improvement on his persuasive essay scores across all measures following SRSD instruction. Carl included a mean of 10.00 essay parts (range
= 8-11) in his essays. This represents an immediate level change of 3.33 parts from baseline testing, when Carl’s mean number of essay parts was 6.67. While his performance was consistent from essay 1 to essay 2, there was a decreasing trend from essay 2 to essay 3. Essay parts scores on the three essays were 11, 11, and 8.

Carl’s holistic quality scores were 9, 9, and 7 over the course of the three SRSD posttest essays ($M = 8.33$, range 7-9), for an overall level change of 3.66 from baseline to SRSD posttesting. Again, Carl’s performance was consistent on essays 1 and 2, but showed a decreasing trend from essay 2 to essay 3. Carl had a mean essay length of 202.00 words, a significant level increase of 128.33 words from baseline testing levels. There was some variability in length throughout the phase (range = 180-222; essay 1 = 180 words, essay 2 = 222 words, and essay 3 = 204 words). Length did not appear to be correlated with parts or quality measures. In fact, Carl’s shortest and longest posttest essays contained the same number of parts and received the same quality score.

Carl had a large number of mechanical errors at SRSD posttesting ($M = 11.05$ errors per 100 words), with variable results across essays (range = 8.11-12.25; essay 1 = 12.78 errors per 100 words, essay 2 = 8.11 errors per 100 words, and essay 3 = 12.25 errors per 100 words). The vast majority of these errors were spelling errors, although Carl made some punctuation and capitalization errors as well.

Carl’s mean SRSD posttest scores indicate that his essays improved overall following SRSD instruction. PND scores indicate that SRSD had a large effect on both the quality and the length of Carl’s essays (PND = 100% on both measures). However,
instruction had little effect on the number of essay parts Carl included in his essays (PND = 66.67%).

**Tim SRSD posttest essay performance.** Tim completed three SRSD posttest essays. He showed immediate improvement on his persuasive essay scores across all measures following SRSD instruction, and consistently performed at this higher level throughout the SRSD posttesting phase. Tim included a mean of 8.67 essay parts (range = 8-9) in his essays. This represents an immediate level change of 5.00 parts from baseline testing, when Tim’s mean number of essay parts was 3.67. His performance was consistent across essays, with scores of 8, 9, and 9 on the three essays.

Tim’s holistic quality scores also increased at SRSD posttesting, with Tim receiving a quality score of 9 on all essays. This represents a level increase of 6.33 points per essay. The length of Tim’s essays also increased to a mean of 154.67 words, a large level increase of 91.00 words from baseline testing levels. There was a decreasing trend from essay 1 to essay 2, but Tim’s performance stabilized from essay 2 to essay 3 (range = 142-178; essay 1 = 178 words, essay 2 = 142 words, and essay 3 = 144 words).

Tim had a moderate number of mechanical errors at SRSD posttesting (\(M = 4.13\) errors per 100 words; range = 0.56-10.42). There was an increasing trend in number of mechanical errors throughout the phase, with the steepest slope between essay 2 and essay 3 (essay 1 = 0.56 errors per 100 words, essay 2 = 1.41 errors per 100 words, and essay 3 = 10.42 errors per 100 words). On the first two essays, Tim had large numbers of capitalization errors in first drafts of his essays (i.e., he did not capitalize the first words of any sentences), but corrected all of these capitalization errors in his final draft. On the
third SRSD posttest essay, however, Tim did not correct these capitalization errors, and
the number of mechanical errors in his essay increased significantly as a result.

Tim’s mean SRSD posttest scores indicate that his essays improved overall
following SRSD instruction. PND scores support these results. Tim had PND scores of
100% on the essay parts, holistic quality, and length measures, indicating that SRSD
instruction was highly effective at improving Tim’s persuasive essay writing skills.

**Group 2.** Students in Group 2 (Ally and John) showed an increase in scores on
the parts and quality measures following instruction in SRSD, but showed little
difference in essay length. Group 2 had a mean of 9.67 ($SD = 1.41$) persuasive essay
parts per essay, and increase of 1.92 essay parts over baseline testing levels.
Additionally, their SRSD posttesting mean holistic quality score was 7.50 ($SD = 1.64$)
and increase of 3.29 points over baseline. In contrast, the mean length of students’ essays
decreased slightly to 305.67 words ($SD = 169.71$), which represents a decrease of 10.83
words over baseline levels. The mean total number of mechanical errors per 100 words
was 5.48 ($SD = 1.32$). This measure was not calculated at baseline.

Overall PND for Group 2 was 22.22%, indicating that SRSD instruction was not
an effective intervention for these students. While there was a questionable effect of the
intervention on the quality of student essays (PND = 66.67%), the intervention had no
effect on essay parts (PND = 0.00%) or essay length (PND = 0.00%) as measured by
PND.

**Ally SRSD posttest essay performance.** Ally completed three SRSD posttest
essays. During SRSD posttesting, Ally showed notable improvements in the consistency
of her persuasive essays, as compared to baseline testing. She included a mean of 10.67 essay parts in her essays. While this represents a level change of 2.67 parts from baseline testing, there was complete overlap of data points. Ally’s performance following at SRSD posttesting differed from baseline in that her performance was considerably more stable following SRSD instruction. During baseline testing, the range of Ally’s essay parts scores was 6-11 while following SRSD instruction the range of essay parts was 10-11 (essay 1 = 11 parts, essay 2 = 10 parts, and essay 3 = 11 parts).

The level of Ally’s holistic quality scores increased substantially at SRSD posttesting to 8.67. This represents a level increase of 5.00 points over baseline levels. Her performance was consistent across essays (range = 8-9; essay 1 = 9, essay 2 = 8, and essay 3 = 9). On the other hand, the length of Ally’s essays did not improve following SRSD instruction. There was no level change from baseline testing \( (M = 421.00 \text{ words}) \) to SRSD posttesting \( (M = 425.67) \). However, the length of Ally’s essays was considerably more stable following SRSD instruction (range = 404-451; essay 1 = 422 words, essay 2 = 404 words, and essay 3 = 451 words).

Ally had a moderate number of mechanical errors at SRSD posttesting \( (M = 6.41 \text{ errors per 100 words}) \). Her performance was relatively stable across essays (range = 5.94-6.65; essay 1 = 6.64 errors per 100 words, essay 2 = 5.94 errors per 100 words, and essay 3 = 6.65 errors per 100 words). She made a combination of capitalization, punctuation, and spelling errors.

Visual analysis of Ally’s SRSD posttest scores indicate that her performance was far more consistent following instruction in SRSD. PND scores, on the other hand,
indicate that SRSD instruction had a positive effect only on the quality of Ally’s essays (PND = 100%). On the essay parts and essay length measures, the intervention was ineffective (PND = 0.00%), according to PND scores.

*John SRSD posttest essay performance.* John completed three SRSD posttest essays. During SRSD posttesting, John’s performance was uneven, but improvements were noted on most persuasive essay measures. First, John included a mean of 8.67 essay parts in his essays. While this represents a level change of 1.10 parts from baseline testing, there was complete overlap of data points. John’s performance on the essay parts measure differed from baseline in that his performance was considerably more stable following SRSD instruction. During baseline testing, the range of John’s essay parts scores was 6-10, while following SRSD instruction the range of essay parts was 8-9 (essay 1 = 9 parts, essay 2 = 8 parts, and essay 3 = 9 parts).

Second, the level of John’s holistic quality scores increased at SRSD posttesting to 6.33. This represents a level increase of 1.58 points over baseline levels. However, his performance was variable across essays (range = 5-8). While John showed an immediate increase in performance at SRSD posttesting from baseline levels, this was followed by a decreasing slope between essay 1 and essay 2, with only a small increase in slope from essay 2 to essay 3. Holistic quality scores on the three essays were 8, 5, and 6.

Third, the length of John’s essays decreased slightly following SRSD instruction. There was a level change from baseline testing (*M* = 212.00 words) to SRSD posttesting (*M* = 185.67) of −26.33 words. Additionally, the length of John’s essays was more
variable following instruction (range = 161-215). Essay 1 showed a small, immediate decrease in number of words from baseline testing levels. There was a slight increasing slope from essay 1 to essay 2, with a corresponding decease in slope between essay 2 and essay 3 (essay 1 = 181 words, essay 2 = 215 words, and essay 3 = 161 words).

Finally, John had a moderate number of mechanical errors at SRSD posttesting ($M = 4.55$ errors per 100 words). His performance was variable across essays (range = 3.73-6.05). Visual inspection showed an increasing slope from essay 1 to essay 2, with a corresponding decreasing slope from essay 2 to essay 3. John had 3.87 errors per 100 words on his first essay, 6.05 errors per 100 words on his second essay, and 3.73 errors per 100 words on his third essay.

While visual analysis reveals some improvements in the level of John’s persuasive essays following SRSD instruction, PND scores indicate that the intervention had little impact on his writing skills. That is, all PND scores were in the ineffective range: essay parts PND = 0.00%, holistic quality PND = 33.33%, essay length PND = 0.00%.

**Group 3.** Students in Group 3 (Navid and Mario) showed a substantial increase in scores on all persuasive essay measures following instruction in SRSD. Group 3 had a mean of 15.67 ($SD = 0.47$) persuasive essay parts per essay, an increase of 9.34 essay parts over baseline testing levels. Additionally, their SRSD posttesting mean holistic quality score was 10.00 ($SD = 0.00$), an increase of 5.79 points over baseline. Similarly, the mean length of students’ essays increased to 276.83 ($SD = 21.45$) words, which represents an increase of 133.20 words over baseline levels. The mean total number of
mechanical errors per 100 words was 5.88 ($SD = 5.83$). This measure was not calculated at baseline.

Overall PND for Group 3 was 100%, indicating that SRSD instruction was a highly effective intervention for these students. More specifically, SRSD was effective at increasing the number of persuasive essay parts students included in their essays (PND = 100%), the holistic quality of their essays (PND = 100%), and the length of their essays (PND = 100%). Individual student performance is presented next.

*Navid SRSD posttest essay performance.* Navid completed three SRSD posttest essays. He showed immediate improvement on his persuasive essay scores for all measures following SRSD instruction. Navid included a mean of 16.00 essay parts (range = 14-18) in his essays. This represents a significant, immediate level change of 10.33 parts from baseline testing, when Navid’s mean number of essay parts was 5.67. While the number of parts in his essays increased dramatically, there was also more variability in his performance during SRSD posttesting. Specifically, there was a decreasing slope from essay 1 to essay 2, and a smaller increasing slope from essay 2 to essay 3. Essay parts scores on the three essays were 18, 14, and 16.

Navid had holistic quality scores of 10 on all three SRSD posttest essays, the highest possible quality score. This was an overall level change of 6.33 from baseline to SRSD posttesting. His performance was consistent across all essays. Navid had a mean essay length of 292.00 words, a significant level increase of 137.35 words from baseline testing levels. There was a decreasing trend throughout the phase, but it was less steep
than during the baseline phase (SRSD posttesting range = 268-307; essay 1 = 307 words, essay 2 = 301 words, and essay 3 = 268 words).

Navid made few mechanical errors at SRSD posttesting ($M = 1.76$ errors per 100 words). While there was a slight increasing slope from essay 1 to essay 2, there was a steep decreasing slope from essay 2 to essay 3, when Navid had no errors in his essay. There were 2.28 errors per 100 words in his first essay, 2.99 errors per 100 words in his second essay, and 0.00 errors per 100 words in his third essay (range = 0.00-2.99). Navid made capitalization, punctuation and spelling errors on his first two SRSD posttest essays.

PND scores support the finding that SRSD instruction was highly effective in improving Navid’s persuasive essay writing skills. He obtained PND = 100% on measures of essay parts, holistic quality, and essay length.

*Mario SRSD posttest essay performance.* Mario completed three SRSD posttest essays. He showed immediate, substantial improvement on his persuasive essay scores across all measures following SRSD instruction. Mario included a mean of 15.33 essay parts (range = 14-18) in his essays. This represents an immediate level change of 8.33 parts from baseline testing, when Mario’s mean number of essay parts was 7.00. Mario continued to show variability on this measure, with an increasing trend from essay 1 to essay 2, with an equal decreasing trend from essay 2 to essay 3. As during baseline testing, Mario appeared able to write better persuasive essays than his mean scores would suggest, but was unable to do so consistently. Essay parts scores on the three essays were 14, 18, and 14.
Mario had holistic quality scores of 10 on all three SRSD posttest essays, the highest possible quality score. This was an overall level change of 5.25 from baseline to SRSD posttesting. His performance was consistent across all essays. Mario had a mean essay length of 261.67 words, a significant level increase of 137.42 words from baseline testing levels. There was some variability in essay length throughout the phase, with a decreasing slope from essay 1 to essay 2, and a corresponding increasing slope from essay 2 to essay 3 (range = 232-286; essay 1 = 267 words, essay 2 = 286 words, and essay 3 = 232 words). There was no correlation between essay length and the number of essay parts or quality of his essays.

Mario had a large number of mechanical errors at SRSD posttesting ($M = 10.01$ errors per 100 words), with variable results across essays (range = 8.39-10.86; essay 1 = 10.86 errors per 100 words, essay 2 = 8.39 errors per 100 words, and essay 3 = 10.78 errors per 100 words). Again, there was variability in Mario’s performance across essays, with a decreasing trend from essay 1 to essay 2, and a corresponding increasing trend from essay 2 to essay 3. The vast majority of his errors were spelling errors, although Mario made a few capitalization and punctuation errors at SRSD posttesting.

PND scores support the finding that SRSD instruction was highly effective in improving Mario’s persuasive essay writing skills. He obtained PND = 100% on measures of essay parts, holistic quality, and essay length.

**Group 4.** Students in Group 4 (Brian, Clare, and Mary Kate) showed an increase in scores on all persuasive essay measures following instruction in SRSD. Group 4 had a mean of 8.39 ($SD = 2.11$) persuasive essay parts per essay, an increase of 4.58 essay parts
over baseline testing levels. Additionally, their SRSD posttesting mean holistic quality score was 6.78 ($SD = 1.84$), an increase of 4.20 points over baseline. Similarly, the mean length of students’ essays increased to 104.33 ($SD = 17.67$) words, which represents an increase of 38.94 words over baseline levels. The mean total number of mechanical errors per 100 words was 11.70 ($SD = 4.46$). This measure was not calculated at baseline.

Overall PND for Group 4 was 88.89%, indicating that SRSD instruction was an effective intervention for these students. More specifically, SRSD was highly effective at increasing the number of persuasive essay parts students included in their essays (PND = 100%), and the holistic quality of their essays (PND = 100%). SRSD instruction had a questionable effect on the length of the essays written by students in Group 4 (PND = 66.67%). Individual student performance is presented next.

**Brian SRSD posttest essay performance.** Brian completed three SRSD posttest essays. He made significant increases in performance on all persuasive essay measures, which he maintained throughout the posttesting phase. Brian included a mean of 10.67 essay parts (range = 10-11) in his essays. This represents an immediate level change of 6.34 parts from baseline testing, when Brian’s mean number of essay parts was 4.33. While his performance was consistent from essay 1 to essay 2, there was a small decrease in slope from essay 2 to essay 3. Essay parts scores on the three essays were 11, 11, and 10.

Brian’s holistic quality scores were 9, 9, and 8 over the course of the three SRSD posttest essays ($M = 8.67$, range 8-9), for an overall level change of 6.00 from baseline to
SRSD posttesting. Again, Brian’s performance was consistent on essays 1 and 2, but showed a small decreasing slope from essay 2 to essay 3. Brian had a mean essay length of 116.00 words, a level increase of 47.00 words from baseline testing levels. As on the other measures, Brian’s performance was consistent on essays 1 and 2, but showed a small decreasing slope from essay 2 to essay 3 (range = 106-122; essay 1 = 122 words, essay 2 = 120 words, and essay 3 = 106 words).

Brian had a large number of mechanical errors at SRSD posttesting ($M = 13.17$ errors per 100 words), with a slight decreasing trend throughout the phase (range = 12.26-13.93; essay 1 = 13.93 errors per 100 words, essay 2 = 13.33 errors per 100 words, and essay 3 = 12.26 errors per 100 words). While Brian made all three types of errors during this phase, he had a particularly large number of capitalization errors.

PND scores support the finding that SRSD instruction was highly effective in improving Brian’s persuasive essay writing skills. He obtained PND = 100% on measures of essay parts, holistic quality, and essay length.

**Clare SRSD posttest essay performance.** Clare completed three SRSD posttest essays. She showed uneven improvements in her persuasive essay performance. Clare included a mean of 8.00 essay parts (range = 7-9) in her essays. This represents an immediate level change of 3.67 parts from baseline testing, when Clare’s mean number of essay parts was 4.33. She showed some variability in performance, with a decreasing slope from essay 1 to essay 2, but a slight increasing slope from essay 2 to essay 3. Essay parts scores on the three essays were 9, 7, and 8.
Clare’s holistic quality scores were 8, 6 and 6 over the course of the three SRSD posttest essays ($M = 6.67$, range 6-8), for an overall level change of 4.34 from baseline to SRSD posttesting. There was a decreasing slope from essay 1 to essay 2, but consistent performance on essay 2 and essay 3. Clare had a mean essay length of 84.00 words, a level change from baseline testing of $-8.67$ words. There was a decreasing slope throughout the phase (range = 54-109; essay 1 = 109 words, essay 2 = 89 words, and essay 3 = 54 words), with Clare’s last essay being only half the length of her first essay. This trend was similar to her performance during baseline testing on this measure.

Clare had a moderate number of mechanical errors at SRSD posttesting ($M = 6.69$ errors per 100 words). There was a steep, increasing slope between essay 1 and essay 2, and no slope between essay 2 and essay 3 (essay 1 = 1.83 errors per 100 words, essay 2 = 8.99 errors per 100 words, and essay 3 = 9.26 errors per 100 words; range = 1.83-9.26). This was due to a large increase in spelling errors on essays 2 and 3.

PND scores corroborate the findings of the visual analysis. Specifically, SRSD instruction was highly effective at increasing the number of essay parts and the holistic quality of Clare’s persuasive essays, with PND=100% on both measures. The intervention did not have a positive effect on the length of Clare’s essays (PND = 0.00%).

Mary Kate SRSD posttest essay performance. Mary Kate completed two SRSD posttest essays. Mary Kate refused to complete the third posttest essay, having an outburst that resulted in in-school suspension. Given her similar performance on the parts measure on her first two SRSD posttest essays, the decision was made to forgo a third posttest essay. At SRSD posttesting, Mary Kate’s performance improved on all
essay measures. She included a mean of 6.50 essay parts (range = 6-7) in her essays. This represents an immediate level change of 3.75 parts from baseline testing, when Mary Kate’s mean number of essay parts was 2.75. She showed a slight decrease in slope from essay 1 to essay 2, with essay parts scores of 7 and 6 on her two essays.

Mary Kate’s holistic quality scores were 6 and 4 during SRSD posttesting ($M = 5.00$, range 4-6), for an overall level change of 2.25 from baseline to SRSD posttesting. There was a decreasing slope from essay 1 to essay 2, because Mary Kate did not include an ending sentence in her second essay. Mary Kate had a mean essay length of 113.00 words, a notable level change from baseline testing of 78.50 words. There was a decreasing slope throughout the phase (range = 96-130; essay 1 = 130 words, and essay 2 = 96 words).

Mary Kate had a large number of mechanical errors at SRSD posttesting ($M = 15.23$ errors per 100 words). There was a decreasing slope between essay 1 and essay 2 (essay 1 = 16.92 errors per 100 words, and essay 2 = 13.54 errors per 100 words; range = 13.54-16.92). These errors were primarily spelling errors, although she did make punctuation and capitalization errors as well.

Mary Kate’s level increases on all measures were supported by PND scores. SRSD instruction was highly effective in improving the number of essay parts, holistic quality, and essay length in Mary Kate’s persuasive essays, with PND = 100% on all measures.

Revision Posttesting. During revision posttesting students again were administered three persuasive essay prompts. Administration procedures were identical
to baseline and SRSD posttesting procedures. Overall, students’ scores decreased on the mechanical errors measure, but showed no additional improvement on the essay parts, holistic quality, and essay length measures following revision instruction (see Table 4).

At revision posttesting, the mean number of essay parts across students was 10.13 ($SD = 1.89$). The mean score for holistic quality was 7.77 out of a possible 10 points ($SD = 1.79$); the mean essay length was 207.77 words ($SD = 101.39$).

There was a decrease in number of mechanical errors per 100 words at revision posttesting. Specifically, the mean number of mechanical errors per 100 words was 4.07 ($SD = 3.09$), a decrease of 3.58 errors per 100 words from baseline levels (baseline $M = 7.65$ errors, baseline $SD = 4.49$). When analyzed by type of error, the number of punctuation errors decreased slightly to 0.62 per 100 words ($SD = 0.84$) following instruction in revision. The mean number of capitalization errors also decreased to 0.78 errors per 100 words ($SD = 1.07$). Likewise, the number of spelling errors decreased to 2.66 errors per 100 words ($SD = 2.49$). Graphic representations of each student’s score for number of essay parts (Figure 1), holistic quality (Figure 2), essay length (Figure 3), and number of mechanical errors (Figure 4) are provided.

Revision posttesting scores were entered into a Wilcoxon Matched-Pairs, Signed Ranks test to determine significant differences in performance between the SRSD posttesting and the revision posttesting phases. There was a significant difference between the two phases on the mechanical errors measure ($p = 0.022$). When comparing types of errors, a significant difference was found across instructional phases for number of spelling errors ($p = 0.037$), but no significant differences were found for capitalization
errors ($\rho = 0.114$) and punctuation errors ($\rho = 0.374$). Additionally, no significant differences were found between SRSD posttesting and revision posttesting on the number of essay parts ($\rho = 0.678$), holistic quality ($\rho = 0.865$), and essay length ($\rho = 0.959$) measures.

Mean PND scores support these findings. While the overall PND score for revision posttesting indicated that instruction had little effect on student’s persuasive writing abilities above and beyond SRSD instruction (Overall PND = 46.25%), there was a small effect for the mechanical errors measure when compared to SRSD posttesting levels (PND = 63.33%). There was no effect on the other persuasive essay measures (essay parts PND = 26.67%; holistic quality PND = 23.33%; essay length PND = 45.00%).

Group and individual student SRSD posttesting performance on all persuasive essay measures is discussed below.

**Group 1.** Students in Group 1 (Andy, Carl, and Tim) showed little difference in performance on most persuasive essay measures following instruction in revision. They did, however, have a decrease in the number of mechanical errors after revision instruction. Group 1 had a mean of 9.22 ($SD = 1.07$) persuasive essay parts per essay, a small decrease of 0.22 essay parts over SRSD posttest levels. Additionally, their revision posttest mean holistic quality score was 7.89 ($SD = 1.39$), a small increase of 0.56 points over SRSD posttesting. Similarly, the mean length of students’ essays remained consistent at 204.94 words ($SD = 53.45$), which was only 2.73 words shorter than SRSD posttest levels. On the other hand, the mean total number of mechanical errors per 100
words was 3.52 ($SD = 4.35$). This was a decrease of 2.71 errors per 100 words from SRSD posttest levels.

PND for the revision phase was calculated by determining the percent of nonoverlapping data points between SRSD posttest and revision posttest levels. This was done to determine the additive effect of revision instruction, above and beyond the effect of the SRSD instruction on students’ persuasive writing performance. Overall PND for Group 1 was 41.67%, indicating that revision instruction did not have an effect on students’ persuasive writing above and beyond that provided by SRSD. When looking at PND scores more closely by persuasive writing measure, PND = 33.33% for the essay parts, holistic quality, and essay length measures. However, there was a small positive effect of revision instruction on the number of mechanical errors per 100 words that student made in their essays (PND = 66.67%). Individual student performance is presented next.

*Andy revision posttest essay performance.* Andy completed three revision posttest essays. His performance decreased on most essay measures following revision instruction, although it was more consistent than during SRSD posttesting. Andy included a mean of 8 essay parts in his revision posttest essays. This represents a change of -1.67 parts from SRSD posttesting, when Andy’s mean number of essay parts was 9.67. His performance was more consistent (range = 7-9) than during SRSD posttesting, and there was an increasing trend throughout the phase. Essay parts scores on the three essays were 7, 8, and 9.
Andy’s holistic quality scores were 6, 6, and 7 over the course of the three revision posttest essays ($M = 6.33$), for an overall level change of -0.27 from SRSD to revision posttesting. Again, Andy’s performance was more consistent (range = 6-7) than during the SRSD posttest phase. Specifically, his performance was consistent on essays 1 and 2, but showed a slight increasing trend from essay 2 to essay 3. Andy had a mean essay length of 173.67 words, a notable level decrease of -92.66 words from SRSD posttesting levels. There was variability in length throughout the phase (range = 130-205), with a decreasing trend from essay 1 to essay 2, and an increasing trend from essay 2 to essay 3 (essay 1 = 205 words, essay 2 = 130 words, and essay 3 = 186 words).

The number of mechanical errors in Andy’s essays decreased at revision posttesting. He had an average of 1.85 errors per 100 words. There was not an immediate level change at revision posttesting. Rather, there was a gradual, decreasing trend throughout the phase (essay 1 = 2.93 errors per 100 words, essay 2 = 1.54 errors per 100 words, and essay 3 = 1.08 errors per 100 words). His performance was also more consistent following revision instruction (range = 1.08-2.93). Similar to SRSD posttesting, all errors were spelling errors; there were no capitalization or punctuation errors during revision posttesting.

PND scores mirror the results of the visual analysis. Revision instruction had no effect on the number of essay parts, holistic quality, or length of Andy’s essays (PND = 0.00% on all measures). It did, however, have a small, positive effect on the number of mechanical errors in his essays (PND = 66.67%).
**Carl revision posttest essay performance.** Carl completed three revision posttest essays. He showed little change in performance on the parts and quality measures, but showed improvement on the essay length and mechanics measures following revision instruction. Carl included a mean of 9.67 essay parts (range = 8-11) in his revision posttest essays. This represents a slight change of -0.33 parts from SRSD posttesting, when Carl’s mean number of essay parts was 10.00. He showed a steady, decreasing trend throughout the phase. Essay parts scores on the three essays were 11, 10 and 8.

Carl’s holistic quality scores were unchanged from SRSD posttest levels ($M = 8.33$, range = 7-9). His scores on the three revision posttest essays were 9, 9, and 7. As during SRSD posttesting, Carl’s performance was consistent on essays 1 and 2, but showed a decreasing trend from essay 2 to essay 3. Carl had a mean essay length of 266.67 words (range = 167-284), a level increase of 64.67 words from SRSD posttesting levels. While there was an immediate level change between the phases, there was a steep, decreasing trend throughout the revision phase (essay 1 = 284 words, essay 2 = 249 words, and essay 3 = 167 words).

The number of mechanical errors in Carl’s essays decreased at revision posttesting. He had an average of 8.47 errors per 100 words (range = 3.59-16.20). This represents a level decrease of 2.58 errors from SRSD posttest levels. During the phase, Carl showed an immediate level increase (essay 1 = 16.20 errors per 100 words), followed by a steep decreasing trend throughout the phase (essay 2 = 5.62 errors per 100 words, and essay 3 = 3.59 errors per 100 words). When analyzing his revision posttest essays further, it was noted that Carl continued to make large numbers of spelling errors.
in the drafts of his essays, but was able to correct these errors more effectively following instruction in revision.

PND scores mirror the results of the visual analysis. Revision instruction had no effect on the number of essay parts or holistic quality (PND = 0.00% on both measures). Conversely, revision instruction had a small, positive effect on essay length and the number of mechanical errors per 100 words (PND = 66.67% on both measures).

*Tim revision posttest essay performance.* Tim completed three revision posttest essays. His performance improved and became more consistent following revision instruction. Tim included 10 essay parts on each of his revision posttest essays. This represents an increase of 1.33 parts from SRSD testing, when Tim’s mean number of essay parts was 8.67. His performance was very consistent; essay parts scores on the three essays were 10, 10, and 10.

Tim’s holistic quality scores were 9, 9, and 9 over the course of the three revision posttest essays ($M = 9.00$), for an overall level change of 2.00 over SRSD posttesting. Again, his performance was very consistent. Tim had a mean essay length of 174.50 words, an increase of 19.83 words from SRSD posttesting levels. There was a small, immediate level change in essay length on the first revision posttest essay, followed by a decreasing trend throughout the phase (range = 155-194; essay 1 = 194 words, essay 2 = 178 words, and essay 3 = 155 words).

The number of mechanical errors in Tim’s essays decreased at revision posttesting. He had an average of 0.26 errors per 100 words, a decrease of 3.87 errors over SRSD posttest levels. There was a large, immediate level change on the first
revision posttest essay, followed by a decrease in slope from essay 1 to essay 2, and no change from essay 2 to essay 3 when he made no mechanical errors in his essays (essay 1 = 0.52 errors per 100 words, essay 2 = 0.00 errors per 100 words, and essay 3 = 0.00 errors per 100 words). His performance was also more consistent following revision instruction (range = 0.00-0.52).

PND scores provide further indication that revision instruction was effective at improving Tim’s persuasive writing skills. Specifically, revision instruction had a large effect on the number of essay parts (PND = 100%) and the holistic quality (PND = 100%) of Tim’s essays. It had a small effect on the number of mechanical errors in Tim’s essays (PND = 66.67%). It did not, however, have an effect on the length of Tim’s essays (PND = 33.33%).

Group 2. Students in Group 2 (Ally and John) showed improvements in performance on most persuasive essay measures following instruction in revision. Group 2 had a mean of 9.83 (SD = 0.24) persuasive parts per essay, which was consistent with their performance at SRSD posttesting (SRSD M = 9.67, SD = 1.41). Their revision posttest mean holistic quality score was 8.33 (SD = 0.47), an increase of 1.33 points over SRSD posttesting. Similarly, the mean length of students’ essays was 356.00 (SD = 132.94) words, which was an increase of 50.33 words from SRSD posttest levels. Finally, the mean number of mechanical errors per 100 words was 4.20 (SD = 4.49). This was a decrease of 1.28 errors per 100 words from SRSD posttest levels.

PND for the revision phase was calculated by determining the percent of overlapping data points between SRSD posttest and revision posttest levels. Overall
PND for Group 2 was 50.00%, indicating that revision instruction did not have an effect on students’ persuasive writing above and beyond that provided by SRSD. More specifically, PND = 33.33% for the essay parts and holistic quality measures, indicating no effect. However, there was a small positive effect of revision instruction on the length of students’ essays (PND = 66.67%) and the number of mechanical errors in their essays (PND = 66.67%). Individual student performance is presented next.

**Ally revision posttest essay performance.** Ally completed three revision posttest essays. She showed little change in performance on the parts and quality measures, but showed improvement on the length and mechanics measures following revision instruction. Ally’s performance was also more consistent following revision instruction.

Ally included 10 essay parts on each essay. While this is a slight decrease of 0.67 parts from SRSD posttesting levels, her performance was very consistent. Her holistic quality scores were unchanged from SRSD posttest levels ($M = 8.67$, range = 8-9). Scores on the three revision posttest essays were 9, 9, and 8. Ally’s performance was consistent on essays 1 and 2, but showed a small decreasing trend from essay 2 to essay 3. Ally had a mean essay length of 450.00 words (range = 375-572), a level increase of 24.33 words from SRSD posttesting levels. While there was an immediate level change between the phases, there was a steep, decreasing trend throughout the revision phase (essay 1 = 572 words, essay 2 = 503 words, and essay 3 = 375 words). The length of Ally’s essays was not correlated with the number of essay parts or the holistic quality of those essays.
The number of mechanical errors in Ally’s essays decreased at revision posttesting. She had an average of 1.02 errors per 100 words. This represents a notable level decrease of 5.39 errors from SRSD posttest levels (SRSD $M = 6.41$ errors per 100 words). Again, she showed improved consistency in performance (range = 0.80-1.39).

There was an immediate and substantial level change on the first essay. Then, there was a slight increase in slope from essay 1 to essay 2, and a corresponding decrease in slope from essay 2 to essay 3 (essay 1 = 0.87 errors per 100 words, essay 2 = 1.39 errors per 100 words, and essay 3 = 0.80 errors per 100 words). Her mechanical errors were a mix of capitalization, punctuation, and spelling errors.

PND scores indicate that the revision intervention had a mixed effect on Ally’s persuasive writing skills. Revision instruction had a large effect on the number of mechanical errors included in her essays (PND = 100%), and a small effect on the length of her essays (PND = 66.67%). Revision instruction did not have an effect on the number of essay parts (PND = 0.00%) or the holistic quality of her essays (PND = 0.00%).

John revision posttest essay performance. John completed three revision posttest essays. His performance improved across most measures, but became more variable than it was following SRSD instruction. Specifically, John performed more poorly on essay 2 than he did on essay 1 and essay 3. Additionally, John was the only student who had an increase in mechanical errors following revision instruction.

John included 8.67 essay parts on each of his revision posttest essays. This represents an increase of 1.17 parts from SRSD testing, when John’s mean number of
essay parts was 7.50. However, his performance was variable across essays (range = 8-11). While there was an immediate level increase on essay 1, there was a decreasing trend from essay 1 to essay 2, and a corresponding increasing trend from essay 2 to essay 3. Essay parts scores on the three essays were 11, 8, and 10.

John’s holistic quality scores were 9, 6, and 9 over the course of the three revision posttest essays (\(M = 8.00\), for an overall level change of 1.67 over SRSD posttesting. Again, his performance was variable (range = 6-9). There was a sharp decrease in slope from essay 1 to essay 2, with a corresponding sharp increase in slope from essay 2 to essay 3. John had a mean essay length of 262.00 words, a large increase of 76.33 words from SRSD posttesting levels. There was an immediate level change on the first revision posttest essay, but his performance throughout the phase was variable. As on other measures, there was a decrease in trend from essay 1 to essay 2, and an increase in trend from essay 2 to essay 3 (range = 186-312; essay 1 = 288 words, essay 2 = 186 words, and essay 3 = 312 words).

The number of mechanical errors in John’s essays increased at revision posttesting. He had an average of 7.38 errors per 100 words, an increase of 2.83 errors over SRSD posttest levels. Again, John demonstrated variable performance on this measure (range = 2.84-9.68). There was a small, immediate level change on the first revision posttest essay (essay 1 = 2.84 errors per 100 words). However, there was a steep increasing trend from essay 1 to essay 2, and errors remained at that high level on essay 3 (essay 2 = 9.69 errors per 100 words, and essay 3 = 9.62 errors per 100 words).
PND scores support the findings of the visual analysis. The variability in John’s scores resulted in a small, positive effect of instruction on the number of essay parts, holistic quality, and length (PND = 66.67% on these measures). Instruction had no effect on the number of mechanical errors in John’s essays (PND = 33.33%).

**Group 3.** Students in Group 2 (Navid and Mario) showed no improvement in performance on most persuasive essay measures following instruction in revision. In fact, there were decreases in performance over SRSD posttest levels. Group 3 had a mean of 12.67 (SD = 1.89) persuasive parts per essay, which was a decrease of 3.00 parts from SRSD posttesting (SRSD M = 15.67, SD = 0.47). Group 3’s revision posttest mean holistic quality score was 9.50 (SD = 0.71), a decrease of 0.50 points over SRSD posttesting. Similarly, the mean length of students’ essays decreased to 192.50 (SD = 29.93) words, which was a decrease of 84.33 words from SRSD posttest levels. Finally, the only improvement in performance that Group 3 made was on the mechanical errors measure. The mean number of mechanical errors per 100 words was 2.20 (SD = 1.97). This was a decrease of 3.68 errors per 100 words from SRSD posttest levels.

PND for the revision phase was calculated by determining the percent of nonoverlapping data points between SRSD posttest and revision posttest levels. Overall PND for Group 3 was 12.50%, indicating that revision instruction did not have an effect on students’ persuasive writing above and beyond that provided by SRSD. PND scores were in the ineffective range on all measures: essay parts PND = 0.00%; holistic quality PND = 0.00%; essay length PND = 0.00%; and mechanical errors PND = 50.00%. Individual student performance is presented next.
Navid revision posttest essay performance. Navid completed three revision posttest essays. He showed increased consistency on essay parts and mechanical errors measures, while showing a decline in essay length. Navid included 14 essay parts in each essay. While this was a decrease of 2.00 parts from SRSD posttesting levels, his performance was much more consistent. Navid holistic quality scores were unchanged from SRSD posttest levels, with a score of 10 out of 10 possible quality points on each essay. On the essay length measure, Navid had a mean of 171.33 words (range = 157-184), a large level decrease of 120.67 words from SRSD posttesting levels. There was an immediate level decrease between the phases, and a small decreasing trend throughout the revision posttest phase (essay 1 = 184 words, essay 2 = 173 words, and essay 3 = 157 words).

The number of mechanical errors in Navid’s essays decreased at revision posttesting. He had an average of 0.81 errors per 100 words (range = 0.00-1.27). This represents a level decrease of 0.95 errors from SRSD posttesting (SRSD $M = 1.76$ errors per 100 words). There was a small increasing slope from essay 1 to essay 2, and consistency in performance from essay 2 to essay 3 (essay 1 = 0.00 errors per 100 words, essay 2 = 1.16 errors per 100 words, and essay 3 = 1.27 errors per 100 words). Navid’s errors during revision posttesting were only punctuation and spelling errors; he no longer made capitalization errors.

PND scores do not support these apparent gains in Navid’s performance on the persuasive essay measures. Based on PND scores (PND = 0.00% on all measures), revision instruction would appear to have had no effect on Navid’s writing performance.
Mario revision posttest essay performance. Mario completed three revision posttest essays. He showed a decline in performance following revision instruction on parts, quality, and length, but showed improvement on the number of mechanical errors per 100 words. Mario included a mean of 11.33 essay parts in his revision posttest essays. This represents a decrease of 4.00 parts from SRSD posttesting, when Mario’s mean number of essay parts was 15.33. His performance was more consistent (range = 11-12) than during SRSD posttesting. Essay parts scores on the three essays were 11, 12, and 11.

Mario’s holistic quality scores was 9.00 on all essays, a decrease in 1.00 quality point from SRSD to revision posttesting. Like SRSD posttesting, Mario’s performance on this measure was highly consistent. Mario had a mean essay length of 213.67 words, a level decrease of 48.00 words from SRSD posttesting levels. There was variability in length throughout the phase (range = 188-226), with a decreasing trend from essay 1 to essay 2, and a corresponding increasing trend from essay 2 to essay 3 (essay 1 = 226 words, essay 2 = 188 words, and essay 3 = 227 words).

The number of mechanical errors in Mario’s essays decreased substantially at revision posttesting. He had an average of 3.59 errors per 100 words (range = 2.64-4.42), a decrease of 6.42 errors from SRSD posttesting levels (SRSD $M = 10.01$). There was an immediate level change, followed by a decreasing trend throughout the phase (essay 1 = 4.42 errors per 100 words, essay 2 = 3.72 errors per 100 words, and essay 3 = 2.64 errors per 100 words). During this phase, all of Mario’s errors were spelling errors; he no longer made punctuation and capitalization errors.
PND scores support these results, indicating that revision instruction had little positive effect on Mario’s persuasive writing skills. According to PND scores, revision had a large effect on the mechanical errors measure (PND = 100%). However, it had no effect on the essay parts, holistic quality, or length measures (PND = 0.00% on all three measures).

**Group 4.** Students in Group 4 (Brian, Clare, and Mary Kate) improved on all essay measures following instruction in revision. Group 4 had a mean of 9.56 ($SD = 2.22$) persuasive essay parts per essay, an increase of 1.17 essay parts over SRSD posttest levels. Their revision posttest mean holistic quality score was 7.11 ($SD = 2.17$), a slight increase of 0.33 points from SRSD posttesting levels ($SRSD \ M = 6.78, SD = 1.84$). The mean length of students’ essays increased to 121.94 ($SD = 29.26$), an increase of 17.61 words over SRSD posttesting. The mean number of mechanical errors per 100 words was 5.76 ($SD = 1.91$). This was a large decrease of 5.94 errors per 100 words from SRSD posttest levels.

PND for the revision phase was calculated by determining the percent of nonoverlapping data points between SRSD posttest and revision posttest levels. Overall PND for Group 4 was 48.61%, indicating that revision instruction did not have an effect on students’ persuasive writing above and beyond that provided by SRSD. When looking at PND scores more closely by persuasive writing measure, there was no effect of instruction on essay parts (PND = 33.33%) or holistic quality (PND = 22.22%). However, there was a moderate, positive impact of instruction on essay length (72.22%)
and mechanical errors (PND = 66.67%). Individual student performance is presented next.

**Brian revision posttest essay performance.** Brian completed three revision posttest essays. He showed some improvements in his essay writing performance, but did not make improvements on all measures. Brian included a mean of 10.67 essay parts (range = 10-11) in his revision posttest essays, just as he did during SRSD posttesting. As during SRSD posttesting, Brian had consistent performance between essays 1 and essay 2, with a slight decrease in performance at essay 3. Essay parts scores on the three essays were 11, 11, and 10.

Brian’s received holistic quality scores of 9 on all revision posttest essays, for an overall level increase of 0.33 from SRSD to revision posttesting. Brian had a mean essay length of 121.94 words (range = 110-142), a level increase of 17.61 words from SRSD posttesting levels. The change in level was not immediate, but there was an increasing slope between essay 1 and essay 2, and consistent performance between essay 2 and essay 3 (essay 1 = 110 words, essay 2 = 142 words, and essay 3 = 138 words).

The number of mechanical errors in Brian’s essays decreased substantially at revision posttesting. He had an average of 3.87 errors per 100 words (range = 0.00=10.91), a decrease of 9.30 errors from SRSD posttest levels (SRSD $M = 13.17$). There was a small, immediate level change, with a steep decreasing trend between essay 1 and essay 2. The number of mechanical errors continued to decrease between essay 2 and essay 3 (essay 1 = 10.91 errors per 100 words, essay 2 = 0.70 errors per 100 words, and essay 3 = 0.00 errors per 100 words).
PND results support the findings of the visual analysis. According to PND scores, revision instruction had no effect on the essay parts or holistic quality of Brian’s essays (PND = 0.00% on both measures). However, it did have a small effect on the length of his essays (PND = 66.67%), and a large effect on the number of mechanical errors (PND = 100%).

Clare revision posttest essay performance. Clare completed three revision posttest essays. She showed improvement on all persuasive essay measures following instruction in revision. Clare had a mean of 11.00 essay parts (range = 10-13), a level increase of 3.00 from SRSD posttesting levels. There was an immediate and large increase in her score on the first revision posttest essay. However, there was a steep decreasing trend between essay 1 and essay 2, with performance stabilizing between essay 2 and essay 3. Clare had essay parts scores of 13, 10, and 10 on her three revision posttest essays.

Clare’s mean holistic quality score also increased following revision instruction ($M = 8.33$, range = 7-9), a level increase of 1.66 points over SRSD posttest levels. Scores on the three revision posttest essays were 9, 7, and 9. Clare had an immediate level increase on the first revision posttest essay. Then, there was a decreasing trend between essay 1 and essay 2, with a corresponding increasing trend between essay 2 and essay 3. Clare had a mean essay length of 146.33 words, a large level increase of 62.33 words from SRSD posttesting levels. Her performance was also more stable following revision instruction (range = 139-155). There was an immediate, positive level change between the phases, with only a slight increase slope between essay 1 and essay 2, and a
corresponding decrease in slope between essay 2 and essay 3 (essay 1 = 139 words, essay 2 = 155 words, and essay 3 = 145 words).

The number of mechanical errors in Clare’s essays decreased at revision posttesting. She had an average of 5.70 errors per 100 words. This represents a level decrease of 0.99 errors from SRSD posttest levels. Again, Clare showed improved consistency in performance (range = 5.52-5.81). There was an immediate and substantial level change on the first essay, with little change in performance across essays (essay 1 = 5.76 errors per 100 words, essay 2 = 5.81 errors per 100 words, and essay 3 = 5.52 errors per 100 words). Clare made a combination of punctuation, capitalization, and spelling errors.

PND scores show variable effects of revision instruction on Clare’s persuasive writing performance. Revision instruction had a large effect on the number of essay parts (PND = 100%) and the length of Clare’s essays (PND = 100%). Instructional had a small effect on the quality of Clare’s essays (PND = 66.67%). According to PND scores, revision instruction had no effect on the number of mechanical errors in her essay (PND = 0.00%).

*Mary Kate revision posttest essay performance.* Mary Kate completed two revision posttest essays. Her performance improved on mechanics following instruction in revision, but did not improve on other essay measures. She was more consistent in her performance as well. Mary Kate had 7.00 essay parts on both revision posttest essays, a level increase of 0.50 from SRSD posttesting levels. Mary Kate’s received a holistic quality score of 4.0 on both essays, which was equal to the quality score on her final
SRSD posttest essay. This was a level decrease of 1.00 points from SRSD posttest levels. Mary Kate had a mean essay length of 89.50 words, a level decrease of 23.50 words from SRSD posttesting levels. Again, her performance was more consistent than during SRSD posttesting (range = 83-96 words). There was an immediate, small level change between the phases, followed by a slight increasing slope between essay 1 and essay 2 (essay 1 = 83 words, essay 2 = 96 words).

The number of mechanical errors in Mary Kate’s essays decreased notably at revision posttesting. She had an average of 7.70 errors per 100 words (range = 6.02-9.38). This represents a level decrease of 7.53 errors from SRSD posttest levels. There was an immediate and substantial level change on the first essay. Then, there was an increase in slope from essay 1 to essay 2 (essay 1 = 6.02 errors per 100 words, and essay 2 = 9.38 errors per 100 words). These errors were evenly spread across punctuation, capitalization, and spelling errors.

PND scores echo these findings. According to PND scores, revision instruction had a large effect on the number of mechanical errors in Mary Kate’s essays (PND = 100%). It had no effect on the number of essay parts (PND = 0.00%), holistic quality (PND = 0.00%), or essay length (PND = 50.00%).

Generalization Testing. Students were administered one generalization prompt following revision posttesting. As with persuasive essay prompts, students were given a page with two content area prompts. The prompts were related to science or social studies content student had learned earlier in the school year. On the first day of generalization testing, students were asked to select a prompt and “write a good
persuasive essay.” Again, they used Neo word processors for their work. On the second day of testing, students were given the Neos and hard copies of their essays, and asked to revise their essays and print final copies for scoring.

At generalization testing, students’ overall mean number of essay parts was 10.30 ($SD = 2.21$), which was similar to revision posttesting levels. Holistic quality improved slightly ($M = 8.30$, $SD = 2.00$), as did essay length ($M = 224.00$ words, $SD = 100.07$). The overall mean number of mechanical errors fell to its lowest point at 3.15 errors per 100 words ($SD = 2.26$). There was a decrease in each type of mechanical error at generalization testing, including punctuation ($M = 0.49$, $SD = 0.53$), capitalization ($M = 0.65$, $SD = 1.09$), and spelling errors ($M = 2.00$, $SD = 1.77$). Group and individual student generalization testing performance on all persuasive essay measures is discussed below.

**Group 1.** Students in Group 1 (Andy, Carl, and Tim) were able to generalize their persuasive writing skills to write essays on science and social studies content. On the generalization essay, students in Group 1 had a mean of 10 parts per essay ($SD = 1.73$), a slight increase of 0.88 parts over revision posttest levels. Similarly, their mean holistic quality score rose to 8.33 ($SD = 1.15$), an increase of 0.44 from revision posttesting. Group 1 did not make gains on the essay length measure ($M = 188.00$, $SD = 30.61$). Instead, their generalization essays were, on average, 16.94 words shorter than their revision posttest essays. Lastly, the mean number of mechanical errors per 100 words increased to 2.95 ($SD = 1.25$) on the generalization essay, an increase of 0.99 errors from revision posttesting levels.
Andy generalization essay performance. Andy was able to transfer his persuasive writing skills to write about science or social studies content, performing similarly to revision posttest levels on the generalization measure. There was complete overlap of scores between the revision posttest phase and generalization testing on all essay measures. Specifically, Andy included 8 persuasive parts in his essay, had a holistic quality score of 7, included 185 words in his essay, and had 1.62 errors per 100 words at generalization testing.

Carl generalization essay performance. Carl was also able to transfer his persuasive writing skills to complete content-related writing. In fact, his performance on the generalization essay was similar to his best revision posttest scores. Carl included 11 essay parts in his generalization essay, had a holistic quality score of 9, included 220 words in his essay, and had 4.09 errors per 100 words.

Tim generalization essay performance. Tim’s performance on the generalization essay was similar to his performance at revision posttesting on most measures. His number of essay parts score increased slightly to 11 parts. Holistic quality remained the same at 9 out of 10 points, as did the length of the essay (159 words). However, the number of mechanical errors increased notably at generalization, from near zero at revision posttesting to 3.14 errors per 100 words at generalization testing.

Group 2. Students in Group 2 (Ally and John) were also able to generalize their persuasive writing skills to write essays on science and social studies content. They did the same as, or better than, their performance at revision posttesting. On the generalization essay, students in Group 2 had a mean of 10 parts per essay ($SD = 1.41$),
similar to the mean of 9.83 parts they earned during revision posttesting. Their mean holistic quality score rose to 9.00 (SD = 0.00), an increase of 0.67 points from revision posttesting. Group 2 performed consistently on the essay length measure (M = 350.00, SD = 134.35), a difference of only 6.00 words from their revision posttest essays. Lastly, their performance on the mechanical errors measure improved. The mean number of mechanical errors per 100 words decreased to 2.61 (SD = 0.20) on the generalization essay, a decrease of 1.59 errors from revision posttesting levels.

*Ally generalization essay performance.* Ally was able to transfer her persuasive writing skills to write about science or social studies content, performing similarly to revision posttest levels on most generalization measures. Ally had slightly more persuasive essay parts in her generalization essay (11 parts). She received the same holistic quality score (9 out of 10), and had an essay of similar length (445 words), when compared to revision posttesting. Ally’s performed more poorly on the mechanical errors measure, however, with 2.47 errors per 100 words in her generalization essay.

*John generalization essay performance.* John was also able to transfer his knowledge of essay writing to write about science and social studies content. There was overlap of scores with revision posttesting on most measures. Specifically, John had similar scores on the essay parts measure (9 parts), holistic quality measure (9 out of 10 points), and essay length measure (255 words). His score on the mechanical errors measure improved on the generalization essay. That is, he had only 2.75 errors per 100 words on the generalization essay, a significant decrease from his revision posttest level of errors.
**Group 3.** Students in Group 3 (Navid and Mario) were able to generalize their persuasive writing skills to write essays on science and social studies content. Their performance was very similar to their revision posttest performance. On the generalization essay, students in Group 3 had a mean of 12.5 parts per essay ($SD = 2.12$), similar to the mean of 12.67 parts they earned in revision posttesting. Their mean holistic quality score stayed the same at 9.50 ($SD = 0.71$). Group 3 performed consistently on the essay length measure ($M = 207.00$, $SD = 38.18$), a small increase of only 14.50 words from their revision posttest essays. Lastly, their performance on the mechanical errors measure improved slightly. The mean number of mechanical errors per 100 words decreased to 1.48 ($SD = 0.28$) on the generalization essay, a decrease of 0.72 errors from revision posttesting levels.

*Navid generalization essay performance.* Navid was able to generalize his persuasive writing skills to write content-related essays. His generalization essay performance overlapped with his revision posttest performance on most measures. Navid performed similar to revision posttesting on the number of essay parts (14 parts), holistic quality (10 out of 10 points), and essay length (180 words). Navid had a slight increase in number of mechanical errors (1.67 errors per 100 words) from revision posttesting.

*Mario generalization essay performance.* Mario was able to apply his persuasive writing skills to write about grade-level content, and even increased his performance on some measures on the generalization essay. Mario had overlap in his performance from revision posttesting to generalization on the parts (11 parts) and quality measures (9 out of 10 parts). He had a small increase in the length of his essay (234 words), and a
decrease in the number of mechanical errors (1.28 errors per 100 words) in his
generalization essay.

*Group 4.* Students in Group 4 (Brian, Clare, and Mary Kate) were able to
generalize their persuasive writing skills to write essays on science and social studies
content, and even improved their performance on some essay measures. On the
generalization essay, students in Group 4 had a mean of 9.33 parts per essay \(SD = 3.06\),
similar to their mean revision posttest level of 9.56 parts. Additionally, their mean
holistic quality score was 7.00 \(SD = 3.46\), similar to their revision posttesting score of
7.11. Group 4 made gains on the essay length measure \(M = 187.33.00, SD = 119.37\).
Their generalization essays were, on average, 65.39 words longer than their revision
posttest essays. Finally, they also made gains on the mechanical errors measure, with a
mean of 4.81 errors per 100 words \(SD = 3.75\), a decrease of 0.95 errors from revision
posttesting levels.

*Brian generalization essay performance.* Brian was able to generalize his essay
writing skills to write about grade-level content. Most of his scores on the generalization
essay overlapped with his performance during revision posttesting. Brian had 10 parts in
his generalization essay, earned 9 out of 10 holistic quality points, and had 0.63
mechanical errors per 100 words. Additionally, Brian increased the length of his essay at
generalization testing to 160 words, 30 words over revision posttest levels.

*Clare generalization essay performance.* Clare had uneven performance on the
generalization essay, when compared to revision posttest levels. She performed similarly
to revision posttesting on the essay parts measure (10 parts), and the holistic quality
measure (9 out of 10 points). Her essay length improved dramatically to 318 words, more than double the number of words at revision posttesting. Clare did more poorly, however, on the mechanical errors measure. She made more errors in the generalization essay, with 7.86 errors per 100 words. This was an increase of 2.36 errors per 100 words over revision posttesting.

*Mary Kate generalization essay performance.* Mary Kate’s performance decreased slightly on the generalization essay, as compared to revision posttesting. Mary Kate’s essay parts (6 parts) and holistic quality scores (3 out of 10 points), decreased from revision posttesting. The length of her essay overlapped with revision posttest levels, with 84 words. The number of errors in her essay was also similar to revision posttesting (5.95 errors per 100 words).

**Maintenance Testing.** Maintenance testing was conducted six weeks (Groups 1 and 2) or eight weeks (Groups 3 and 4) after revision posttesting was completed. During maintenance testing, students were administered one persuasive essay prompt. Again, administration procedures were identical to testing procedures at baseline, SRSD posttesting, and revision posttesting. Maintenance testing scores showed slight improvements over revision posttesting levels for number of essay parts ($M = 10.44, SD = 2.19$), holistic quality ($M = 8.44, SD = 2.07$), length ($M = 227.33, SD = 79.90$), and number of errors per 100 words ($M = 3.17, SD = 2.40$). However, the types of errors students made shifted somewhat. Students made fewer punctuation errors at maintenance testing ($M = 0.16, SD = 0.33$), but made more spelling errors ($M = 2.36, SD = 1.86$). The number of capitalization errors remained the same ($M = 0.65, SD = 0.83$). Graphic
representations of each student’s score for number of essay parts (Figure 1), holistic quality (Figure 2), essay length (Figure 3), and number of mechanical errors (Figure 4) are provided.

Maintenance testing scores were entered into a Wilcoxon Matched-Pairs, Signed Ranks test, which indicated significant differences between revision posttesting and maintenance testing on the holistic quality ($\rho = 0.046$) and essay length measures ($\rho = 0.038$), as well as on the number of punctuation errors per 100 words ($\rho = 0.028$). No significant differences were found in the number of essay parts ($\rho = 0.320$), number of mechanical errors ($\rho = 0.594$), number of capitalization errors ($\rho = 0.249$) or number of spelling errors ($\rho = 0.953$) between revision posttesting and maintenance testing.

Maintenance testing scores were also compared to baseline scores using the Wilcoxon Matched-Pairs, Signed Ranks test. Significant differences were found for the number of essay parts ($\rho = 0.008$), holistic quality ($\rho = 0.008$), and essay length ($\rho = 0.015$). While the number of mechanical errors could not be compared to baseline levels, maintenance levels were compared to SRSD posttest levels. There was a significant difference in number of mechanical errors per 100 words between the two phases ($\rho = 0.008$). In addition, mean PND scores indicated a moderate to large effect of SRSD and revision instruction on students’ persuasive essay performance (see Table 6). Mean PND for essay parts between baseline and maintenance testing was 88.89%; mean PND for holistic quality was 100%; and mean PND for essay length was 88.89%.

Group and individual student maintenance testing performance on all persuasive essay measures is discussed below.
**Group 1.** At maintenance testing, Group 1 maintained revision posttest levels on the essay quality and essay length measures, and improved their performance on the essay parts and mechanical errors measures. They had a mean of 10.33 essay parts \((SD = 1.15)\), an increase of 1.11 essay parts over revision posttest levels. Their maintenance mean holistic quality score was 8.00 \((SD = 1.73)\), virtually unchanged from their mean revision posttest score of 7.89 \((SD = 1.39)\). Group 1 also maintained revision posttest levels on the essay length measure, with a mean of 205.67 words per essay \((SD = 50.41)\). Finally, Group 1 improved their performance on the mean number of mechanical errors per 100 words \((M = 0.95, SD = 0.33)\), a decrease of 2.57 errors per 100 words from revision posttesting levels.

Maintenance test scores were also compared to baseline testing levels to determine overall growth in student performance resulting from the combination of SRSD and revision instruction. Group 1 outperformed all baseline measures at maintenance testing. At maintenance testing, the number of essay parts increased by 5.22 from baseline testing. The overall quality of students’ essays increased by 4.11 points on a 10-point scale. The length of students’ essays increased by 83.78 words. While mechanical errors were not calculated at baseline testing, Group 1 had 5.28 fewer mechanical errors per 100 words at maintenance testing than they did at SRSD posttesting.

PND scores were also calculated, comparing student performance at maintenance testing to their baseline testing levels. Overall, the combination of SRSD and revision instruction was highly effective at improving students’ persuasive writing skills (overall
PND = 88.89%). It was highly effective in improving the number of parts in students’ essays (PND = 100%) and holistic quality (PND = 100%). It was somewhat effective in improving the length of students’ essays (PND = 66.67%) and the number of mechanical errors (PND = 66.67%). Individual student performance is presented next.

*Andy maintenance essay performance.* During maintenance testing, Andy performed similarly to his revision posttest levels on essay parts (9 parts), holistic quality (6 out of 10 points), and essay length (139 words), with complete overlap between these scores and revision posttest scores. Andy’s performance on the mechanical errors measure improved at maintenance testing to its lowest level (0.72 errors per 100 words).

Instruction in SRSD and revision improved Andy’s performance on most essay measures. At maintenance testing, Andy had 4.0 more essay parts than he did at baseline testing, and his holistic quality score increased by 2.0 points. Additionally, the number of mechanical errors in his essays decreased by 2.81 errors per 100 words from SRSD posttesting to maintenance. In contrast to these gains, the length of Andy’s essays decreased over the course of the study. His maintenance essay was 89.33 words shorter than his mean baseline essay length. In fact, his maintenance essay was one of the shortest essays he wrote during the study. These findings are corroborated by PND scores, which indicate that SRSD and revision instruction were very effective at improving the number of parts (PND = 100%), holistic quality (PND = 100%), and number of mechanical errors (PND = 100%) in Andy’s essays, compared to baseline levels. Andy’s essay length scores at maintenance testing, however, showed complete overlap with baseline levels (PND = 0.00%).
**Carl maintenance essay performance.** During maintenance testing, Carl performed at a level equal to his highest performances during SRSD and revision posttesting on the essay parts (11 parts) and holistic quality (9 out of 10) measures. His essay was not as long as revision posttest essays (222.5 words), but it remained longer than baseline levels. Additionally, Carl’s performance on the mechanical errors measure decreased to the lowest levels of the study during maintenance testing (1.33 errors per 100 words).

Instruction in SRSD and revision improved Carl’s performance on all essay measures. At maintenance testing, Carl had 4.33 more essay parts than he did at baseline testing. His holistic quality score also increased by 4.33 points, and the length of his essays increased by 151.33 words. Additionally, the number of mechanical errors in Carl’s essays decreased dramatically by 9.72 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were very effective at improving all aspects of Carl’s persuasive essays, with PND = 100% on the essay parts, holistic quality, and essay length measures from baseline to maintenance testing. Likewise, PND = 100% on the mechanical errors measures from SRSD posttesting to maintenance testing.

**Tim maintenance essay performance.** During maintenance testing, Tim’s performance was similar to revision posttest levels. Tim performed better on the essay parts measure at maintenance testing, including 11 parts in each essay. His performance on the holistic quality measure was equal to revision posttest performance (9 out of 10). Additionally, his essay was considerably longer at maintenance testing (253 words). Tim
showed a slight increase in the number of mechanical errors in his essay at maintenance testing (0.79 errors per 100 words).

Instruction in SRSD and revision improved Tim’s performance on all essay measures. At maintenance testing, Tim had 7.33 more essay parts than he did at baseline testing. His holistic quality score increased by 6.00 points, and the length of his essays increased by 189.33 words. Additionally, the number of mechanical errors in his essays decreased by 3.34 errors per 100 words from SRSD posttesting to maintenance testing. Overall, SRSD and revision instruction were highly effective at improving Tim’s persuasive writing skills, with PND = 100% for parts, quality, and length from baseline testing to maintenance testing. However, PND scores indicated that instruction had no effect on the number of mechanical errors in his essays (PND = 0.00%).

**Group 2.** John was the only student in Group 2 to complete maintenance testing. Although several attempts were made to administer maintenance testing to Ally, she refused to complete the essay task. Therefore, no group means are available for Group 2. John’s individual performance is discussed next.

**John maintenance essay performance.** During maintenance testing, John’s performance improved on the holistic quality, essay length, and mechanical errors measures. John included 10 essay parts in his maintenance essay, which overlapped with his revision posttesting performance. The level of his holistic quality score rose from 8.0 at revision posttesting to 10.0 at maintenance testing, with no overlap in scores between the two phases. John wrote his longest essay of the study during maintenance testing, with 378.0 words. This was an increase of 116 words over revision posttest levels.
Lastly, the number of mechanical errors in John’s essays decreased substantially to 2.12 errors per 100 words, 5.26 errors fewer than during revision posttesting. This represents John’s lowest level of errors in the entire study.

Instruction in SRSD and revision improved John’s performance on all essay measures. At maintenance testing, John had 2.50 more essay parts than he did at baseline testing. His holistic quality score increased by 5.25 points, and the length of his essays increased by 166 words. Additionally, the number of mechanical errors in his essays decreased by 2.43 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were highly effective at improving the quality, length, and number of mechanical errors in John’s essays, with PND = 100% on these measures. However, the instructional phases had no impact on the number of parts of John’s essays (PND = 0.00%). During baseline testing, John was able to write one essay with 10 parts, which overlapped with his maintenance testing performance.

**Group 3.** At maintenance testing, Group 3 maintained revision posttest levels on the essay parts and quality measures, improved their performance on the essay length measure, and performed more poorly on the mechanical errors measure. They had a mean of 13.00 essay parts ($SD = 1.41$), little changed from their revision posttest level of 12.67. Their maintenance mean holistic quality score was 10.00 ($SD = 0.00$), a slight increase of 0.50 points over revision posttesting. Group 3 increased the length of their essays at maintenance testing with a mean of 268.00 words per essay ($SD = 11.31$). This represents an increase of 75.50 words over revision posttesting levels. Finally, Group 3
performed more poorly on the mechanical errors measure, with a mean of 3.82 errors per 100 words ($SD = 3.85$). This is an increase of 1.62 errors per 100 words over revision posttest levels.

Maintenance test scores were compared to baseline testing levels to determine overall growth in student performance resulting from the combination of SRSD and revision instruction. Group 3 outperformed all baseline measures at maintenance testing. At maintenance testing, the number of essay parts increased by 6.67 from baseline testing. The overall quality of students’ essays increased by 5.79 points on a 10-point scale. The length of students’ essays increased by 124.37 words. While mechanical errors were not calculated at baseline testing, Group 3 had 2.06 fewer mechanical errors per 100 words at maintenance testing than they did at SRSD posttesting.

PND scores were also calculated, comparing student performance at maintenance testing to their baseline testing levels. Overall, the combination of SRSD and revision instruction was effective at improving students’ persuasive writing skills (overall PND = 87.50%). It was highly effective in improving the number of parts in students’ essays (PND = 100%), holistic quality (PND = 100%), and essay length (PND = 100%). It did not have an effect on the number of mechanical errors in students’ essays (PND = 50.00%), as compared to SRSD posttesting levels. Individual student performance is presented next.

*Navid maintenance essay performance.* During maintenance testing, Navid’s performance was consistent with his revision posttesting performance on the essay parts (14 parts) and holistic quality (10 out of 10) measures. The length of his maintenance
The essay returned to SRSD posttest levels (276.0 words), a large increase over revision posttesting levels (revision posttest mean = 171.33 words). Lastly, the number of mechanical errors in Navid’s essay increased slightly from revision posttest levels to 1.09 errors per 100 words, an increase of 0.28 errors.

Instruction in SRSD and revision improved Navid’s performance on all essay measures. At maintenance testing, Navid had 8.33 more essay parts than he did at baseline testing. His holistic quality score increased by 6.33 points, and the length of his essays increased by 113 words. Additionally, the number of mechanical errors in his essays decreased by 0.67 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were very effective at improving the number of essay parts, quality and length of Navid’s essays, with PND = 100% on all measures. Instruction had no effect on the number of mechanical errors in his essays (PND = 0.00%).

**Mario maintenance essay performance.** During maintenance testing, Mario’s performance maintained, or improved upon, revision posttest levels of performance. As during revision posttesting, Mario included 12 persuasive essay parts in his maintenance essay. His holistic quality score improved to SRSD posttest levels of performance, with 10 out of 10 possible quality points. The length of his essay also improved to SRSD posttest levels at 260.00 words. This represents an increase over revision posttesting levels of 46.33 words. On the other hand, Mario’s performance on the mechanical errors measure showed an increase in errors over revision posttesting (6.54 errors per 100
words). However, this is still well below the number of mechanical errors during SRSD posttesting.

Instruction in SRSD and revision improved Mario’s performance on all essay measures. At maintenance testing, Mario had 5 more essay parts than he did at baseline testing. His holistic quality score increased by 5.25 points, and the length of his essays increased by 135.75 words. Additionally, the number of mechanical errors in his essays decreased by 3.47 errors per 100 words from SRSD posttesting to maintenance testing. PND scores further indicate that SRSD and revision instruction was highly effective at improving Mario’s persuasive writing skills, with PND = 100% on all measures (parts, quality, length, and mechanics).

**Group 4.** At maintenance testing, Group 4 maintained revision posttest levels on the essay parts, quality, and mechanics measures, and improved their performance on the essay length measure. They had a mean of 9.00 essay parts ($SD = 2.65$), a small decrease from their revision posttest level of 9.56. Their maintenance mean holistic quality score was 7.33 ($SD = 2.89$), a small increase of 0.22 points over revision posttesting. Group 4 increased the length of their essays at maintenance testing with a mean of 171.67 words per essay ($SD = 59.91$). This represents an increase of 40.73 words over revision posttesting levels. Finally, Group 4 performed similarly on the mechanical errors measure, with a mean of 5.31 errors per 100 words ($SD = 3.85$). This was a decrease of 0.45 errors per 100 words from revision posttest levels.

Maintenance test scores were compared to baseline testing levels to determine overall growth in student performance resulting from the combination of SRSD and
revision instruction. Group 4 outperformed all baseline measures at maintenance testing. At maintenance testing, the number of essay parts increased by 5.19 from baseline testing. The overall quality of students’ essays increased by 4.75 points on a 10-point scale. The length of students’ essays increased by 106.28 words. While mechanical errors were not calculated at baseline testing, Group 4 had 6.39 fewer mechanical errors per 100 words at maintenance testing than during SRSD posttesting.

PND scores were also calculated, comparing student performance at maintenance testing to their baseline testing levels. Overall, the combination of SRSD and revision instruction was highly effective at improving students’ persuasive writing skills (overall PND = 91.67%). It was highly effective in improving the number of parts in students’ essays (PND = 100%), holistic quality (PND = 100%), and essay length (PND = 100%). It had a small effect on the number of mechanical errors in students’ essays (PND = 66.67%), as compared to SRSD posttesting levels. Individual student performance is presented next.

*Brian maintenance essay performance.* During maintenance testing, Brian’s performance was consistent with his revision posttesting performance on the essay parts (11 parts) and holistic quality (9 out of 10 point) measures. His performance on the essay length measure improved to 189 words. In fact, Brian wrote his longest essay during maintenance testing. On the other hand, Brian performed more poorly on the mechanical errors measure than he did during revision posttesting, with 5.29 errors per 100 words. Although this was not as low as his revision posttest scores, it still represented marked improvement over SRSD posttest levels.
Instruction in SRSD and revision improved Brian’s performance on all essay measures. At maintenance testing, Brian had 6.67 more essay parts than he did at baseline testing. His holistic quality score increased by 6.33 points, and the length of his essays increased by 120 words. Additionally, the number of mechanical errors in his essays decreased substantially by 7.88 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were very effective at improving Brian’s persuasive writing, with PND = 100% on all measures (parts, quality, length, and mechanics).

*Clare maintenance essay performance.* During maintenance testing, Clare’s performance was consistent with her revision posttesting performance on most measures. Similar to her highest revision posttest scores, Clare included 10 persuasive essay parts in her essay, and earned 9 out of 10 points on the holistic quality rubric. Her maintenance essay was longer than her revision posttest essays, with 221 words. This represents an increase of 74.67 words over revision posttesting. Finally, she maintained her performance on the mechanical errors measure, with 5.88 errors per 100 words.

Instruction in SRSD and revision improved Clare’s performance on all essay measures. At maintenance testing, Clare had 5.67 more essay parts than she did at baseline testing. Her holistic quality score increased by 6.67 points, and the length of her essays increased significantly by 128.33 words. Additionally, the number of mechanical errors in her essays decreased by 0.81 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were very effective at improving the number of essay parts, quality, and length of Clare’s essays,
with PND = 100% on those measures. Instruction had no effect on the number of mechanical errors in her essays (PND = 0.00%).

Mary Kate maintenance essay performance. During maintenance testing, Mary Kate performed more poorly than revision posttesting on the essay parts measure, similarly to revision posttesting on the quality and length measures, and better than revision posttesting on the mechanics measure. To begin, the number of persuasive essay parts Mary Kate included in her maintenance essay decreased slightly to 6 parts. However, she maintained a holistic quality score of 4, and included 105 words in her essay. Finally, Mary Kate improved her performance on the mechanical errors measure, with 4.76 errors per 100 words.

Instruction in SRSD and revision improved Mary Kate’s performance on all essay measures. At maintenance testing, Mary Kate had 3.25 more essay parts than she did at baseline testing. Her holistic quality score increased by 1.25 points, and the length of her essays increased significantly by 70.50 words. Additionally, the number of mechanical errors in her essays decreased by 0.81 errors per 100 words from SRSD posttesting to maintenance testing. PND scores indicate that SRSD and revision instruction were very effective at improving Mary Kate’s persuasive writing, with PND = 100% on all measures (parts, quality, length, and mechanics).

Revisions. Revisions were scored in two ways. First the total number of revisions was calculated. This score was converted to reflect the total number of revisions per 100 words. Next, each revision was classified as either a change in
meaning or a surface change. Table 7 details the mean number of revisions per 100 words by type.

Table 7

<table>
<thead>
<tr>
<th></th>
<th>Mean Number of Revisions by Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Revisions per 100 words (SD)</td>
</tr>
<tr>
<td></td>
<td>X(SD) % of total X(SD) % of total</td>
</tr>
<tr>
<td>Baseline (N=10)</td>
<td>2.60 (3.32) 0.23 (0.29) 8.85% (3.26) 91.15%</td>
</tr>
<tr>
<td>SRS SD Posttest (N=10)</td>
<td>3.60 (3.06) 0.42 (0.66) 11.67% (2.50) 88.33%</td>
</tr>
<tr>
<td>Revision Posttest (N=10)</td>
<td>6.12 (4.72) 0.48 (0.70) 7.84% (4.88) 92.32%</td>
</tr>
<tr>
<td>Generalization (N=10)</td>
<td>5.85 (5.30) 0.19 (0.36) 3.25% (5.33) 96.75%</td>
</tr>
<tr>
<td>Maintenance (N=9)</td>
<td>3.85 (3.15) 0.58 (0.96) 15.06% (3.46) 85.19%</td>
</tr>
</tbody>
</table>

Note. There were no significant differences in type of revision made from baseline to SRS SD posttesting, from SRS SD posttesting to revision posttesting, or from baseline to maintenance testing, according to the Wilcoxon Matched-Pairs, Signed Ranks test (p < 0.05).

The mean number of revisions per 100 words at baseline testing was 2.60 (SD = 3.32). This number went up slightly following SRS SD instruction to 3.60 (SD = 3.06). Following instruction in revision, the number of revisions per 100 words nearly doubled to 6.12 (SD = 4.72). This number fell slightly at generalization testing to 5.85 (SD = 5.30), and returned to post-SRS SD levels at maintenance testing (M = 3.85, SD = 3.15).

Instruction in revision appears to have had an immediate, positive impact on the number
of revisions made. However, these gains were not maintained over time. The Wilcoxon Matched-Pairs, Signed Ranks test found no significant difference in performance across phases: baseline to SRSD posttesting, SRSD posttesting to revision posttesting, and baseline to maintenance.

When analyzing revisions by type, the clear majority of revisions during each phase were surface change revisions. Surface change revisions accounted for from 85.19% to 96.75% of all revisions throughout the course of the study (range = 2.37-5.66 surface change revisions per 100 words). Few meaning change revisions were made during any phase (range = 3.25-15.06%). While there appears to be a slight increase in the number of meaning change revisions at SRSD posttesting ($M = 0.42$) and maintenance testing ($M = 0.58$), these changes were small and were not significant. Students made the fewest number of meaning change revisions on the generalization essay ($M = 0.19$). Given the relatively small number of meaning change revisions overall, the following analysis focuses on the total number of revisions per 100 words.

Figure 5 displays individual student data by revision type for each testing phase. Two students, John and Mario, showed a steady increase in the number of revisions over the course of the study. Two other students, Andy and Navid, showed an increase in number of revisions following instruction in the peer-revision strategy, which they were able to maintain over time. Other students (i.e., Carl, Tim, and Mary Kate) increased the number of revisions following revision instruction, but these revisions were not maintained over time. The final three students (i.e., Ally, Brian, and Clare) showed an increase in number of revisions following SRSD instruction, but did not show an
additional increase after revision instruction. Overall, 60% of students continued to make more revisions at maintenance testing than they had during baseline testing (i.e., Andy, Tim, Ally, John, Mario, and Brian). The remaining students (i.e., Carl, Navid, Clare, and Mary Kate) made fewer revisions at maintenance testing than during baseline testing.
**Figure 5.** Type of revision by student during baseline, SRSD posttesting, revision posttesting, and maintenance testing.
Self-efficacy for Persuasive Writing Scale. The mean self-efficacy composite score at baseline for all students was 3.25 ($SD = 0.91$), indicating a moderate level of self-efficacy for writing persuasive essays. Following instruction in the SRSD persuasive writing strategy, the mean composite score increased slightly to 3.48 ($SD = 1.14$). The self-efficacy score increased slightly again following instruction in the revision strategy to 3.68 ($SD = 1.22$), and remained at that level during maintenance testing ($M = 3.61$, $SD = 1.23$). However, the Wilcoxon Matched-Pairs, Signed Ranks test revealed no differences in scores across phases (baseline to SRSD posttest $p = 0.169$; baseline to revision posttest $p = 0.153$; baseline to maintenance $p = 0.314$). Both instruction in SRSD and instruction in revision had no impact on student’s self-efficacy for persuasive writing.

Writing fluency. At baseline, participants’ mean standard score for writing fluency was 89.22 ($SD = 11.88$). As a group, participants were in the low average range for writing fluency when compared with their same-age peers. Following fluency instruction the group mean standard score was 90.78 ($SD = 11.49$), indicating no change in writing fluency as a result of instruction. A paired sample t-test confirmed this finding ($t(8) = 0.382$, $p = 0.712$). While writing instruction did not focus specifically on building writing fluency, there is some evidence to suggest that writing fluency may increase simply from increase time spent writing (e.g., Mastropieri et al., 2010). This does not appear to be the case for students in this study.

Peer-conferencing student performance. During the course of the study, there were 12 different student peer-conference pairings during revision lessons 4 and 5 (i.e., 2
pairings in both Group 1 and Group 2, and 4 pairings in both Group 3 and Group 4). In
11 of these pairings, students worked well together, with no negative student-to-student
interactions. In one Group 3 pairing, students had a verbal dispute stemming from an
earlier incident outside of class, which required teacher intervention.

Student performance during peer-conferences was analyzed in two additional
ways. First, academic performance was evaluated by identifying the steps of peer-
conferences completed by each student. Second, behavioral performance was evaluated
by calculating on-task behavior during student-led instruction, as compared to teacher-led
instruction and independent work time.

*Steps of peer-conferencing completed.* The number of steps completed by each
student during each peer-conferencing session was calculated to provide a more nuanced
look at how effectively students were able to meet academic goals when peer-
conferencing during the revision instructional phase. Overall, students completed
87.94% of the required steps of peer-conferencing.

The step that was most consistently omitted was the second, editing conference.
While students consistently edited each other’s essays for capitalization, punctuation, and
spelling errors, they rarely held a second conference to go over the edits. Instead, they
simply handed the edited paper back to their partners to make the corrections. In
approximately half of all instances (9 out of 19), students edited for their peers during the
first conference, which was supposed to focus exclusively on revisions.

Peer-conferencing steps completed varied by instructional group, with Groups 1
and 4 having the highest completion rates (92.86% and 91.07%, respectively), and Group
3 having the lowest rate (80.36%). Group 1 consistently completed all peer-revision components, with the exception of the editing conference. Group 4 was the only group that completed the editing conference. Brian and Clare completed 100% of the steps of peer-revision. However, the group’s overall score was negatively impacted because Mary Kate refused to participate in the editing portion of the second conference. Group 2 completed 83.72% of the steps of the peer-revision conferences. Ally missed several student-led sessions due to avoidance behaviors. During the revision portion of one lesson, John edited his own essay, rather than asking his partner to edit it. Finally, Group 3 completed the steps of peer-revision conferences with 80.36% fidelity. Their score was negatively impacted by their behaviors. During one class, Navid said he was not feeling well, put his head down on the desk, and did not do any writing work. During another class, Mario did not revise his partner’s essay, even though his partner revised Mario’s essay. Instead, Mario played on the computer.

On-task behavior. Overall on-task behavior across participants was 75.86%. Levels of on-task behavior varied based upon the type of instruction being delivered. The overall group mean for on-task behavior during teacher-led instruction was 78.09% (SD = 14.01). During independent work time, this average fell to 67.69% (SD = 19.33). On-task behavior was highest when students were engaged in student-led instruction (i.e., peer tutoring) with a mean of 82.11% (SD = 11.60). Results of the Wilcoxon Matched-Pairs, Signed Ranks test found a significant difference between on-task behavior during teacher-led instruction and independent work time, but did not find significant differences between teacher-led instruction and student-led instruction, or between student-led
instruction and independent work time. Individual student on-task behavior is detailed on Table 8.

Table 8
Percent of On-task Behavior by Type of Instruction

<table>
<thead>
<tr>
<th>Participant</th>
<th>Teacher-led instruction</th>
<th>Independent Work</th>
<th>Student-led instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ally</td>
<td>71.60%</td>
<td>29.33%</td>
<td>75.58%</td>
</tr>
<tr>
<td>Andy</td>
<td>94.19%</td>
<td>89.83%</td>
<td>86.67%</td>
</tr>
<tr>
<td>Brian</td>
<td>81.67%</td>
<td>77.78%</td>
<td>90.00%</td>
</tr>
<tr>
<td>Carlos</td>
<td>86.36%</td>
<td>83.33%</td>
<td>94.87%</td>
</tr>
<tr>
<td>Clare</td>
<td>86.67%</td>
<td>55.06%</td>
<td>75.56%</td>
</tr>
<tr>
<td>John</td>
<td>81.18%</td>
<td>76.00%</td>
<td>98.72%</td>
</tr>
<tr>
<td>Mario</td>
<td>69.88%</td>
<td>77.78%</td>
<td>76.67%</td>
</tr>
<tr>
<td>Mary Kate</td>
<td>86.67%</td>
<td>65.56%</td>
<td>74.44%</td>
</tr>
<tr>
<td>Navid</td>
<td>78.65%</td>
<td>78.89%</td>
<td>64.44%</td>
</tr>
<tr>
<td>Tim</td>
<td>44.04%</td>
<td>43.33%</td>
<td>84.15%</td>
</tr>
<tr>
<td>Total</td>
<td>78.26%</td>
<td>67.42%</td>
<td>81.71%</td>
</tr>
</tbody>
</table>

Differences in on-task behavior across differing types of instruction was most pronounced for a few students. First, Tim had a great deal of difficulty maintaining on-task behavior during teacher-led and independent work time (44.04% and 43.33%, respectively). During these times, Tim was observed turning his body away from instruction, picking at his body, staring at the floor or walls, or eating candy that he kept in his pocket. He required a great deal of teacher redirection to complete tasks, particularly during independent work time. However, during peer-revision conferences,
Tim was able to remain on-task 84.15% of the time. He actively engaged in discussions with his peers and provided relevant feedback to his revising partners.

Similarly, Ally demonstrated extreme off-task behavior during independent work time. While she was able to maintain moderate levels of on-task behavior during teacher-led instruction (71.60%) and student-led instruction (75.58%), most independent work time was spent off-task (on-task behavior = 29.33% for independent work time). This lack of on-task behavior resulted in Ally writing less than her peers during instruction. Her refusal to do work during independent work times was so extreme that she had to be removed from the room by the school’s Crisis Resource Specialist on one occasion.

Mary Kate and Clare also had more difficulty than their peers during independent work time, but their off-task behavior was not as extreme as Ally’s.

Finally, unlike his peers, Navid had the most difficulty remaining on-task during student-led instruction. While he maintained consistent levels of on-task behavior during teacher-led and independent work times (78.65% and 78.89%, respectively), he showed a relatively large decrease in on-task behavior when working with peers (64.44%). A review of observation notes does not identify any externalizing behavior problems that may explain this decreased score. Rather, during two consecutive student-led sessions, Navid appeared extremely tired and was often observed closing his eyes or putting his head down on the desk.
Strategy knowledge questioning. Students were asked to recall the components of the SRSD writing strategy (POW+TREE) and the revision strategy (e.g., Does it have all the parts? Is it clear? Is it persuasive?) during baseline testing, weekly during SRSD and revision instruction, during revision posttesting, and at maintenance testing. Figure 6 show results, by group, on the two strategy questions. When students were asked to list the parts of a good persuasive essay during baseline testing, on average, they were able to list 1.4 parts. The most common part identified was the topic, or “state your opinion.” At baseline, students were able to identify 2 correct items when asked, “What do you look for when you revise a persuasive essay?” The items most often identified were spelling and capitalization.

During SRSD instruction, there was an immediate level increase for all groups on the POW+TREE parts probe, while the number of revision parts remained at baseline levels. During revision instruction, then, the number of persuasive essay parts identified by students increased slightly, and the number of revision parts identified showed an immediate level increase. The pattern was similar for all instructional groups. At revision posttesting, students identified an average of 8.4 components of SRSD’s POW+TREE strategy, and an average of 4.8 components of the revision strategy. At maintenance testing, the level of SRSD components identified was consistent with posttest levels ($M = 8.6$), while the level of revision components identified increased to 6.13. The increase in the number of revision components identified at maintenance testing was due to students in group 3 identifying more parts of the strategy. Students in Groups 1, 2, and 4 showed consistent performance between post-testing and maintenance
testing on both strategy knowledge questions. These results indicate that students were able to memorize most of the components of the SRSD and revision strategies.
Figure 6. Group mean scores on the strategy knowledge questionnaire.
Social Validity. Student responses to the social validity questionnaire indicated overall positive perceptions of the SRSD and revision instruction. To begin, all students said that SRSD instruction helped them become better writers. The most often cited reason was that it helped them organize their thoughts before writing (4 students). Three students said that the clear guidelines helped. As Brian explained, “It gave a clear guideline for what you needed to do or how you needed to do it.” Two students cited the repetition of the strategy throughout instruction, and two students said that receiving rewards for writing was what helped them become better writers.

When asked how SRSD might help other students, five students said that it would help them organize their writing. Similarly, two students said it would help others plan. Two other students said that SRSD would help their peers stay focused on their writing goals. When asked how they would improve instruction, only one student offered a suggestion – having more time to work on their writing. At the time of the interviews during revision posttesting, no participants reported using SRSD in other classes, but suggested that they could use the strategy during science, civics, or physical education.

The social validity interview also asked questions about revision instruction. When asked what about the revising strategy helped them become better writers, students provided a range of responses. Eight of the ten responses were positive – the revision instruction helped them become better writers. Two students, Ally and Mary Kate, did not think the revision instruction helped them become better writers. They did not like working with partners – Ally because she preferred to work alone, and Mary Kate
because she did not like other people finding errors in her work. As Mary Kate put it, “Every time I write an essay, people find a lot of errors and I like it just the way it is.”

Another question on the interview asked students how working with their peers helped them become better writers. The most common response, provided by three students, was that other people could find mistakes that the writer could not. Two students said it was helpful to hear someone else’s point of view. Summaries of each students’ responses to the social validity interview follow.

*Andy felt both SRSD instruction and revision instruction helped him become a better writer.* When speaking about SRSD instruction, he said, “I’m able to write more clearly because I know what I am supposed to be writing. I organize it. So instead of being a blob of words, it actually means something.” The graphic organizer was a particularly helpful tool for him. He felt SRSD could help other students organize their writing, too. Andy was equally positive about revision instruction, saying it helped him, “...go back and find my mistakes. I could make the sound of the words flow more evenly, and I could persuade more people.” He found working with other students helpful because they could find mistakes that he could not.

*Carl found both SRSD instruction and revisions instruction useful.* He found instruction on counter reasons to be the most helpful part of SRSD instruction because it was something he had not learned before. It also helped him organize his writing, and work through the writing process. He felt SRSD could help other students organize their ideas and plan before writing. He would have appreciated more time spent on writing instruction. Carl also valued the peer-revision process, even though he did not find it
necessary at first. Several times throughout the course of the interview, Carl mentioned that it was helpful to work with others. When asked why it was helpful, Carl said, “Sometimes you don’t see everything. You might miss something. It helps to have the person next to you look at it sometimes.”

*Tim felt that both the SRSD instruction and the revision instruction helped him become a better writer,* noting, “It helped me become an okay writer to a pretty good writer.” Specifically, Tim noted that the clear expectations helped him know what he was supposed to write, the graphic organizer helped him organize his thoughts, and the ending was helpful. Additionally, he found the entire revising process helpful. As he said, “It helped me know what other people thought of my essay. Then I could correct it so it made sense to other people, not just me.”

*Ally thought the SRSD instruction was somewhat helpful, but did not find revision instruction useful.* She felt that the benefit of SRSD instruction was that she could “get into the essays quickly,” even if she wasn’t “feeling up to it.” Ally thought it was a good organizing strategy that others might find useful. The most helpful aspect of SRSD instruction was learning about counter reasons, which she had never needed to include in her essays before. Ally was less positive about revision instruction. She felt it was “dragging on” her. While she acknowledged that working with peers helped her, she preferred working alone.

*John felt both the SRSD instruction and the revision instruction improved his writing.* SRSD instruction helped him write better because it helped him organize his thoughts. He said, “It’s a way of organizing your thoughts neatly so they are separated.
So when you go to write it is not all jumbled together. So you can’t find your reasons and stuff.” John thought it could help other students organize their writing, too. John appreciated getting other students’ perspectives on his writing during the peer-revision conferences. He felt it helped his essays “make more sense.”

Navid felt that both the SRSD instruction and the revision instruction helped him improve his writing. SRSD helped him know exactly what he was supposed to do, to the point that it became automatic. He also liked charting his progresses and comparing the number of parts in his essay to others. As he said, “I like being the best writer.” Navid thought SRSD could help other students become better writers, although he did not offer specific ways it could help them. Lastly, Navid recommended that the teacher develop a mnemonic for the editing portion of revision task. During revision instruction, Navid liked working with a partner. His partner noticed mistakes that he did not notice, which was helpful.

Mario found both the SRSD instruction and the revision instruction very helpful for improving his writing. SRSD was most helpful to him because it made his writing organized and clear. He described it as, “. . . going straight, and not writing the same thing over again.” Writing a lot also helped him improve his writing. Mario felt SRSD could help other students learn how to organize their writing. During revision instruction, Mario found it helpful to see how other people wrote and their ideas. As he said, “What was good about [the revision strategy] was that I get to talk to other people . . . and see how they write.” Furthermore, working with a peer helped him find his errors.
Brian felt that SRSD instruction and revision instruction set him on the path to improve his writing skills. As he said, “It gave me the basic idea. [It] basically started a foundation for helping me with writing.” He appreciated that SRSD gave a clear guideline for what he needed to do, and that helped him keep his writing focused on the topic. Brian thought it could help other students stay focused in their writing, as well. The most useful tools for him were the graphic organizer and the transition word list. During the interview, Brian described how the organization process was now automatic for him – he could do it in his head, rather than writing his ideas in a graphic organizer. When speaking of revision instruction, Brian felt it was helpful to get feedback from his peers, and hear what his writing “sounded like from a different stand point.” He expressed some concern about working with a partner, however, because he would not always be able to work with a partner on his writing (e.g., in college), so he needed to learn to do it on his own.

Clare liked SRSD instruction, and felt revision instruction helped her become a better writer. The aspect of SRSD instruction that stood out to her the most was repeating the POW+TREE mnemonic every day, and she mentioned this several times throughout the interview. Clare felt that other students could benefit from learning the mnemonic. Additionally, she enjoyed getting rewards for remembering the strategy. Lastly, Clare felt her writing improved because of the peer-revision strategy, and she enjoyed working with a partner to revise her essay.

Mary Kate liked the SRSD instruction, but was frustrated by the peer-revision process. While her favorite aspect of instruction was receiving rewards for completing
essay, Mary Kate stated that SRSD instruction did help her become a better writer. As she said, “At the beginning, I didn’t even know we had to organize and stuff. But now I know how to do it and it’s easy.” She found the graphic organizer, the transition word list, and the POW+TREE mnemonic to be the most helpful tools. Mary Kate also said that SRSD instruction could help other students stay focused in their writing. On the other hand, Mary Kate did not find the revision instruction helpful for improving her writing skills. She preferred to work alone, and did not like other people finding mistakes in her essays.

Summary

A number of dependent measures were administered and scored to determine growth in students’ writing abilities and on-task behavior as a result of instruction in SRSD and the peer-revision strategy, including: (a) persuasive essay writing prompts, (b) Persuasive Writing Self-efficacy Scale, (c) on-task behavior data, (d) social validity interview, (e) Woodcock-Johnson Writing Fluency subtest (Woodcock et al., 2001), (f) strategy knowledge questionnaire, and (g) fidelity of treatment checklist. These data sources were analyzed using a variety of methods, including: (a) traditional visual analysis, (b) PND, (c) nonparametric statistical analysis, and (d) qualitative analysis. Fidelity of treatment was 94.79%.

Persuasive essay results indicate large gains on the number of persuasive essay parts, the holistic quality, and the length of students’ essays following SRSD instruction. Students maintained these levels of performance following revision instruction, but no additional gains were seen on these measures. However, following revision instruction,
the number of mechanical errors in students’ essays decreased substantially. Students were able to generalize these skills to write about science and social studies content. Moreover, they were able to maintain these gains at 6 and 8 week maintenance testing, outperforming baseline levels on all measures.

No change was seen in the number of revision students made following instruction in SRSD, but there was a large increase in number of revisions following revision instruction. However, the number of revisions students made at maintenance testing fell to pre-revision instruction levels. During all phases, students primarily made surface-level changes and few meaning change revisions.

Students were able to meet both the academic and behavioral expectations of peer-conferencing. On average, they completed 87.94% of the required steps on the peer-conferencing checklist. Additionally, on-task behavior data revealed that students were most on-task during student-led instruction (81.71%), and least on-task during independent work time (67.42%). On-task behavior during teacher-led instruction was 78.26%.

In social validity interviews, students reported that they liked the SRSD POW+TREE strategy and found it useful. Eight out of 10 students reported finding the revision strategy helpful for improving their writing, and enjoyed working with their peers. Two students reported that they did not like working with their peers to revise their essays.
Finally, no significant differences were found on the Persuasive Writing Self-efficacy Scale (range of $\rho = 0.153 - 0.314$) and the Writing Fluency subtest ($\rho = 0.712$) from baseline testing to revision posttesting.
5. DISCUSSION

Students’ academic and behavioral performance improved as a result of instruction in SRSD and the peer-revision strategy. This chapter first discusses the findings of the study in relation to each research questions. Next, a discussion of the single-subject quality indicators (Horner et al., 2005) is presented, followed by discussion of the limitations of the study, implications for practice, and directions for future research.

Research Questions

This study posed six research questions about the effectiveness of the SRSD and revision interventions. The answers to these questions are discussed below.

Persuasive essay performance. The first research question asked: Does persuasive writing improve (as measured by length, content, quality, and mechanics) after instruction in SRSD and after instruction in revision? Results from this study indicate that students made considerable gains in their ability to write persuasive essays following instruction. Different types of gains were made after instruction in SRSD than were made after instruction in revision. To begin, following SRSD instruction, students wrote persuasive essays that had more essay parts, were of higher quality, and were longer. These results mirror findings of similar studies in which adolescents with EBD were taught the POW+TREE strategy for persuasive writing (e.g., Cuenca Sanchez et al.,
2012; Mastropieri et al., 2009). Table 9 compares the results of studies that taught multi-paragraph persuasive essays using SRSD to adolescents with EBD. The present study provided among the least amount of instructional time (412.5 minutes) of all studies, which was well below the average instructional time of 1033.44 minutes across studies. While students in this study made somewhat smaller gains on the number of essay parts, holistic quality, and essay length measures than students in other studies, they were able to maintain their gains at higher rates at maintenance testing. For example, students in this study had an overall gain on the holistic quality measure of 4.1 points following SRSD instruction. This was 0.27 points less than the overall average across studies. However, students in this study increased their quality score at maintenance testing by 0.71 points, while students’ quality scores across studies fell by an average of 1.23 points at maintenance testing. Similar patterns were seen on the essay parts and essay length measures.

Another area in which students made gains similar to students in other SRSD studies was in their ability to generalize their writing skills to other content areas, and to maintain their skills over time. Students in this study were able to generalize their skills to write about science and social studies content, just as students in the Mastropieri et al. (2009, 2010, 2012) studies did. Additionally, students were able to maintain these skills at six and eight week maintenance testing. In fact, students in this study outperformed students in other studies on maintenance measures. With an average of 7 weeks between posttesting and maintenance, the present study had the average length of maintenance testing compared to other studies. Two studies (Cuenca Sanchez et al., 2012; Mastropieri
et al., 2012) had only two weeks between posttesting and maintenance testing, while two other studies (Mastropieri et al., 2009, 2010) had 11.5 and 12 week, respectively, between posttesting and maintenance. Across studies, student performance declined between posttesting and maintenance on the parts, quality, and length measures. In this study, however, student maintained their scores on the essay parts measure, and increased their performance on the holistic quality and essay length measures. The only other studies during which students made gains at maintenance testing were the Hauth (2012) and Mastropieri et al. (2012) studies. In those studies, students made gains on the essay parts and holistic quality measures. Additionally, in the Mastropieri et al (2009) study, students maintained their levels of performance on the quality measure. All other measures in all other studies showed a decline between posttesting and maintenance.

When comparing results across SRSD studies, it is important to remember that students in this study were higher functioning than students in most of the Mastropieri et al. (2009, 2010, In press) studies and the Cuenca Sanchez (2012) study. That is, students in the present study were educated in self-contained classes in a general education middle school, rather than in a self-contained school, indicating that they were higher functioning emotionally and behaviorally. They were also higher functioning in terms of their writing skill when entering the study. Students in the present study wrote longer essays
<table>
<thead>
<tr>
<th>Study</th>
<th>Instruction Provided</th>
<th>Mean Min. of SRSD Inst.</th>
<th>Mean Weeks Maint.</th>
<th>Number of Essay Parts</th>
<th>Holistic Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Study</td>
<td>SRSD + Revision</td>
<td>412.5</td>
<td>7</td>
<td>+4.93</td>
<td>+0.02</td>
</tr>
<tr>
<td>Cuenca Sanchez (2012)</td>
<td>SRSD + Self-determination</td>
<td>990</td>
<td>2</td>
<td>+6.64</td>
<td>-1.82</td>
</tr>
<tr>
<td>Hauth (2012)</td>
<td>SRSD + Civics Content</td>
<td>280</td>
<td>5</td>
<td>+9.0</td>
<td>+0.26</td>
</tr>
<tr>
<td>Irby Cerar (2012)</td>
<td>Fluency + SRSD</td>
<td>1320</td>
<td>5</td>
<td>+7.14</td>
<td>-0.68</td>
</tr>
<tr>
<td>Mastropieri et al. (2009)</td>
<td>SRSD + Fluency</td>
<td>1,740</td>
<td>11.5</td>
<td>+3.7</td>
<td>-0.57</td>
</tr>
<tr>
<td>Mastropieri et al. (2010)</td>
<td>SRSD only</td>
<td>2,532</td>
<td>12</td>
<td>+3.71</td>
<td>-1.11</td>
</tr>
<tr>
<td>Mastropieri et al. (In press)</td>
<td>SRSD + Fluency</td>
<td>693</td>
<td>2</td>
<td>+6.75</td>
<td>+0.25</td>
</tr>
<tr>
<td>Mastropieri et al. (2012)</td>
<td>SRSD only</td>
<td>300</td>
<td>8</td>
<td>+2.72</td>
<td>-1.43</td>
</tr>
<tr>
<td>Mean Range</td>
<td></td>
<td>1033.44</td>
<td>6.56</td>
<td>5.57</td>
<td>-0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>280</td>
<td>2</td>
<td>+2.72</td>
<td>-1.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,532</td>
<td>12</td>
<td>+9.0</td>
<td>+0.26</td>
</tr>
</tbody>
</table>

**Note.** The minutes of SRSD instruction reported here represent only the number of minutes students received instruction in the SRSD intervention. Instructional time in other instructional phases of a study (e.g., revision, fluency, civics content) are not represented in this comparison.
with more essay parts at baseline testing than students in the Mastropieri et al. and Cuenca Sanchez studies. In these respects, they were most similar to students in the Hauth (2012) and Mastropieri et al. (2012) studies. As in the present study, students in the Hauth and Mastropieri et al. (2012) studies required shorter intervention periods and were better able to maintain their gains over time than their lower functioning peers.

Unlike previous studies that included adolescents with EBD in SRSD instruction, the present study included additional instruction in revision. Following revision instruction, students did not make further gains on the essay length, essay parts, or holistic quality measures. They did make gains on the mechanical errors measure, however. That is, students had fewer capitalization, punctuation, and spelling errors per 100 words following revision instruction.

A number of earlier revision studies reported on the number of mechanical errors and the overall quality of student essays, which can be compared to the present study. When comparing the number of capitalization, punctuation, and spelling errors per 100 words, students in this study performed similarly their peers in other revision studies (see Table 10). Following revision instruction, students in this study decreased the number of spelling errors and punctuation errors, but saw no change in the number of capitalization errors. At maintenance testing, students in the present study further decreased the number of errors in their essays, as compared to post-revision instruction. Only one other study (Stoddard & MacArthur, 1993) included information about mechanical errors at
maintenance testing. In that study, students made more errors in capitalization and spelling at maintenance testing, but made fewer errors in punctuation.

A number of revision studies can also be compared to the present study in terms of overall essay quality (see Table 11). Students in this study showed little change in essay quality following revision instruction, while student in other studies had an average gain in quality of 1.56 points following instruction. Few studies included a maintenance measure of essay quality. Of the three that included such a measure, students in the present study had the largest quality gains at maintenance testing, as compared to posttest levels.

Finally, instruction also resulted in more consistency in performance for students who were highly variable at the beginning of the study. Tim, Ally, and John showed improved consistency following SRSD instruction. Ally continued to become more consistent following revision instruction, as did Andy, Navid, Mario and Mary Kate. (Carl and Brian performed similarly in both posttesting phases.) Mason et al. (2010) found similar benefits of SRSD quick write instruction for eighth grade students who were already able to meet grade level expectations in writing. These findings support existing research that demonstrates the benefits of SRSD instruction for writers at a variety of levels (e.g., De La Paz & Graham, 2002).
Table 10

**Mechanical Errors Difference Scores Across Revision Studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Revision Instruction Provided</th>
<th>Capitalization Pre-to Post-Treatment&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Capitalization Post-Treatment to Maint.</th>
<th>Punctuation Pre-to Post-Treatment&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Punctuation Post-Treatment to Maint.</th>
<th>Spelling Pre-to Post-Treatment&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Spelling Post-Treatment to Maint.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Study</td>
<td>Peer revision</td>
<td>0.00</td>
<td>-0.13</td>
<td>-0.10</td>
<td>-0.46</td>
<td>-2.18</td>
<td>-0.30</td>
</tr>
<tr>
<td>Graham &amp; MacArthur (1988)</td>
<td>SCAN</td>
<td>-0.35&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>-0.33&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>-0.02&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
</tr>
<tr>
<td>MacArthur et al (1991)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Peer revision</td>
<td>-0.12</td>
<td>N/A</td>
<td>-0.14</td>
<td>N/A</td>
<td>-0.60</td>
<td>N/A</td>
</tr>
<tr>
<td>Stoddard &amp; MacArthur (1993)</td>
<td>Peer revision</td>
<td>3.60</td>
<td>0.65</td>
<td>2.55</td>
<td>-4.45</td>
<td>-1.45</td>
<td>1.55</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.78</td>
<td>0.26</td>
<td>0.50</td>
<td>-2.46</td>
<td>-1.06</td>
<td>0.63</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>(-0.35) – (-0.13) –</td>
<td>(0.65)</td>
<td>(2.55)</td>
<td>(-4.45) – (-0.46)</td>
<td>(-2.18) – (-0.02)</td>
<td>1.55</td>
</tr>
</tbody>
</table>

*Note.*<sup>a</sup>For the present study, change scores were calculated by subtracting revision posttest scores from SRSD posttest scores, rather than from baseline, in order to determine the additive effect of the peer revision intervention on the number of mechanical errors in student essays. <sup>b</sup>Graham and MacArthur (1988) averaged number of mechanical errors for post-instruction and maintenance testing, calling it the proportion of errors “post-treatment.” <sup>c</sup>No maintenance testing was included in this study.
Table 11

Holistic Quality Difference Scores Across Revision Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Instruction Provided</th>
<th>Pre- to Post-Treatment</th>
<th>Post-Treatment to Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Study(^a)</td>
<td>Peer-revision</td>
<td>0.04</td>
<td>0.67</td>
</tr>
<tr>
<td>Graham (1997)(^bc)</td>
<td>CDO</td>
<td>0.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Graham &amp; MacArthur (1988)</td>
<td>SCAN</td>
<td>1.38</td>
<td>0.07</td>
</tr>
<tr>
<td>Graham et al (1995)(^c)</td>
<td>Goal setting</td>
<td>1.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Graham et al (1995)(^c)</td>
<td>Goal setting + procedural facilitation</td>
<td>0.97</td>
<td>N/A</td>
</tr>
<tr>
<td>MacArthur et al (1991)(^c)</td>
<td>Peer revision</td>
<td>1.50</td>
<td>N/A</td>
</tr>
<tr>
<td>Saddler &amp; Assar (2007)(^c)</td>
<td>SRSD + Revising through dialogue</td>
<td>3.96</td>
<td>N/A</td>
</tr>
<tr>
<td>Stoddard &amp; MacArthur (1993)</td>
<td>Peer revision</td>
<td>3.63</td>
<td>-0.75</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.56</td>
<td>-0.01</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>0.00 – 3.96</td>
<td>(-0.75) – 0.67</td>
</tr>
</tbody>
</table>

Note. \(^a\)In the present study, change scores were calculated by subtracting revision posttest scores from SRSD posttest scores, rather than from baseline, in order to determine the additive effect of the peer revision intervention on the number of mechanical errors in student essays. \(^b\)For experimental and quasi-experimental studies, difference scores were calculated by subtracting the mean for the control group from the mean from the experimental group. \(^c\)No maintenance testing was included in this study.
Revisions. The second research question asked: *What types of revisions do students make using the peer-revision process? What impact do these revisions have on the length, content, quality, and mechanics of the essay?* In this study, students increased the number of revisions they made following instruction in the peer-revision process. They primarily made surface-level changes during all phases of the study. There was no increase in the number of meaning-level changes following revision instruction. As such, revisions had no impact on the length, content or quality of student essays. Revisions did have a positive impact on the number of mechanical errors students made. That is, the surface-level revisions students made reduced the number of punctuation, capitalization, and spelling errors in their essays. However, students were not able to maintain the increased number of revision over time. At maintenance testing, the number of revisions students made fell to SRSD posttesting levels. It is important to note that, although students did not make as many revisions during maintenance testing, their essays had even fewer mechanical errors than revision posttest essays. This means that students had fewer mechanical errors in the *first* drafts of their essays during maintenance testing.

Table 12 compares the number and types of revisions students made to the results of other revision studies. As in other studies (e.g., Stoddard & MacArthur, 1993), students made more revisions from first drafts to final drafts following revision instruction. In the present study, these results were primarily surface-level changes. Other revision studies by MacArthur and colleagues (Graham, MacArthur, & Schwartz, 1995; MacArthur, Schwartz, & Graham, 1991; Stoddard & MacArthur, 1993) found
contradictory results. In those studies, students made more meaning-change revisions following instruction in a revision strategy, which improved the quality of their essays.

One reason for the difference in results between this study and previous studies might be that, in a number of previous studies, students were explicitly taught different types of revisions they could make (e.g., add more information). No instruction on types of revisions was included in the present study. Rather, students were guided to look for essay parts, persuasiveness, and clarity, but were not explicitly taught how to remedy any issues raised by their partners. While the teacher modeled and provided guided practice in revising, students were not given a list of possible revision types to choose from or were not given explicit instruction to add a certain amount of information to their essays, as was done in other studies. Such explicit instruction and guidance may further enhance the efficacy of this peer-revision strategy.
Table 12

**Revision Difference Scores Across Studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Instruction Provided</th>
<th>Total Revisions per 100 Words</th>
<th>Surface Changea Revisions per 100 Words</th>
<th>Meaning Change Revisions per 100 Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre- to Post-Treatment</td>
<td>Post-Treatment to Maintenance</td>
<td>Pre- to Post-Treatment</td>
<td>Post-Treatment to Maintenance</td>
</tr>
<tr>
<td>Present Studyb</td>
<td>Peer-revision</td>
<td>2.52</td>
<td>2.27</td>
<td>2.47</td>
</tr>
<tr>
<td>Graham (1997)c,d</td>
<td>CDO</td>
<td>4.16</td>
<td>N/A</td>
<td>1.99</td>
</tr>
<tr>
<td>Graham &amp; MacArthur (1988)</td>
<td>SCAN</td>
<td>3.05</td>
<td>-1.19</td>
<td>-0.48</td>
</tr>
<tr>
<td>Graham et al. (1995)d</td>
<td>Goal setting</td>
<td>7.39</td>
<td>N/A</td>
<td>0.19</td>
</tr>
<tr>
<td>Graham et al. (1995)d</td>
<td>Goal setting + procedural facilitation</td>
<td>2.1</td>
<td>N/A</td>
<td>-5.93</td>
</tr>
<tr>
<td>Saddler &amp; Assaro (2007)d</td>
<td>SRSD + revising through dialogue</td>
<td>3.44</td>
<td>N/A</td>
<td>3.44</td>
</tr>
<tr>
<td>Stoddard &amp; MacArthur (1993)</td>
<td>Peer revision</td>
<td>4.30</td>
<td>2.75</td>
<td>1.42</td>
</tr>
<tr>
<td>Mean</td>
<td>Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.85</td>
<td>3.44</td>
<td>0.44</td>
<td>-0.45</td>
</tr>
</tbody>
</table>

**Note.** aThe definition of surface change revisions in the present study was similar to the definition of “non-meaning change revisions” in the Graham and MacArthur studies. bIn the present study, change scores were calculated by subtracting revision posttest scores from SRSD posttest scores, rather than from baseline, in order to determine the additive effect of the peer revision intervention on the number of mechanical errors in student essays. cFor experimental and quasi-experimental studies, difference scores were calculated by subtracting the mean for the control group from the mean from the experimental group. dNo maintenance testing was included in this study.
Another reason for the lack of findings may be because, unlike SRSD instruction, no mastery criteria were set for students during the revision phase. In the SRSD instructional phase, students had to be able to write an essay that included 11 essay parts (topic, 3 reasons, 3 explanations, counter reason, explanation for counter reason, refute the counter reason, and ending) before they could enter the SRSD posttesting phase. In the revision instructional phase, however, no criteria were established to move students to posttesting. Rather, students were moved into revision posttesting when they completed the five revision lessons. This was done because this study was testing the feasibility of the researcher-developed revision strategy. Setting a criterion level for mastery, such as requiring that students include three meaning-change revisions in their essays, before moving on to revision posttesting would likely have resulted in more meaning-change revisions during revision posttesting and, in turn, may have had a positive impact on essay quality.

Self-efficacy for persuasive writing. The third research question was: Does student self-efficacy for writing improve following SRSD and revision instruction? Are changes in self-efficacy maintained over time? Students’ self-efficacy for the persuasive writing task was measured at baseline testing, in conjunction with the first persuasive essays at SRSD posttesting and revision posttesting, and at maintenance testing, using the researcher-developed Persuasive Writing Self-Efficacy Scale. Results showed no difference in scores following instruction in SRSD, following instruction in revision, or at maintenance testing. Therefore, in this study, instruction had no effect on students’ self-efficacy for the persuasive writing task. This was true for all students.
Only two other studies focused on SRSD and students with EBD (Cuenca Sanchez et al., 2012; Irby Cerar, 2012) reported on self-efficacy. In these studies, students in the experimental condition made larger gains on the researcher-developed self-efficacy measure than students in the control condition. Given the limited number of studies that have reported self-efficacy measures, and the contradictory results of these studies, no conclusion can be drawn about the effectiveness of SRSD in improving students’ self-efficacy for the writing task.

Writing fluency. The fourth research question was: Does students’ writing fluency improve as a result of instruction in the SRSD and revision strategies? Students’ writing fluency was measured at baseline and revision posttesting using the Woodcock Johnson III Writing Fluency subtest (Woodcock et al., 2001). There was no difference in student scores following writing instruction. In some previous SRSD studies, which explicitly taught a strategy to develop writing fluency, students’ writing fluency increased (e.g., Mastropieri et al., 2009, 2012; Mason et al., 2010). In other SRSD studies that measured fluency, but did not explicitly teach it, results were mixed. In one study (Mastropieri et al., 2010), students showed gains in raw scores on the Writing Fluency subtest, while in another study (Cuenca Sanchez et al., 2012) students showed no improvement following instruction. Cuenca Sanchez theorized that this difference may be a function of the amount of instructional time in the two studies. Mastropieri et al. (2010) provided 2,532 minutes of instruction, while Cuenca Sanchez et al. provided only 990 minutes of instruction. This may be the reason that students in this study, who received
an average of 862.5 minutes of instruction, did not show gains in writing fluency as a by-product of writing instruction.

**Working with peers.** The fifth research question was: *Are students with EBD able to accurately conduct all steps of the peer-revision strategy? Are students with EBD socially able to work together as measured by time on task (as compared to time on task during teacher-led instruction)?* One of the most interesting findings of this study was how well students with EBD were able to work together to complete the peer-revision process. Students were able to achieve academic goals, as measured by the number of peer-conferencing steps completed by student pairs. Students conducted nearly all components of the peer revision conferences. The component that was omitted most often – the second, editing conference – was not completely ignored. Rather, students more naturally edited their partners’ work during the first revision conference. Additionally, students were more on-task when working with each other during peer conferences, than they were during teacher-led instruction or during independent work time. In fact, when tasked to complete work on their own, as was typically the case for these students, they exhibited the lowest levels of on-task behavior. This finding suggests that students with EBD may have more success behaviorally when they have structured opportunities to work with their peers. Literature reviews on peer-mediated instruction for students with EBD (Ryan et al., 2004; Spencer et al., 2009) reported similar improvements in academics and behavior for those who participated in peer-mediated learning activities.
Social validity. The sixth research question was: *Do students find the SRSD and revision strategies easy to implement, useful and enjoyable?* Overall, students found SRSD and revision instruction useful in improving their persuasive essay performance. Students reported that SRSD instruction in the POW+TREE strategy was useful because it made expectations for how to write persuasive essays clear. Additionally, students felt that the organizational tools were the most helpful aspects of the intervention for struggling writers. Similar responses were obtained in other SRSD studies with adolescents (e.g., Mastropieri et al., 2009, 2010). When asked about revision instruction, nearly all students reported that they found peer-revision helpful. In particular, they noted that it was helpful to get another person’s perspective on their writing, whether to help them find errors or to hear another point of view. Additionally, they enjoyed working with partners during the peer-revision phase. Spencer et al. (2003) received similar feedback from middle school students with EBD when using peer tutoring during reading instruction. Such feedback further supports the use of peer-mediated instruction with these students as they enjoy it and find it beneficial.

**Study Quality**

Horner et al. (2005) outlined the criteria for high-quality single-subject research. The current study met most of these criteria. First, the instructional practice was clearly defined, and implemented with a high degree of fidelity. All of the materials used in the SRSD and revision interventions are provided in this study, as well as detailed descriptions of each lesson. Fidelity of treatment data indicates that both interventions were implemented as intended, with 98.94% fidelity of treatment for SRSD instruction,
and 88.24% fidelity of treatment for revision instruction. Therefore, this study meets Horner et al.’s criteria for (a) operationally defining the practice, and (b) implementing the practice with fidelity.

Another criteria of high-quality single-subject research is that the context and outcomes are clearly defined. The current study meets these criteria by: (a) providing detailed information about participant characteristics and how those participants were selected, (b) describing the setting in which the intervention took place, (c) describing and providing copies of each dependent measure, as well as detailed information about administration and scoring of each measure, and (d) thoroughly describing baseline conditions. The quality of this study is further enhanced because of the high social validity of the interventions. That is, improving writing skills and increasing on-task behavior of adolescents with EBD are socially important goals. Furthermore, the study was implemented in a typical classroom setting over an extended period of time, and was positively received by participants, making the intervention a practical and cost-effective strategy.

One of the most important criteria for high-quality, single-subject research is that the intervention is functionally related to a change in the desired outcome (Horner et al., 2005). This is achieved through demonstrating experimental control. The current study only partially met this expectation. It met this expectation by measuring student essay writing performance repeatedly over time. That is, students wrote at least three essays during the baseline phase, three essays during the SRSD posttesting phase, three essays
and a generalization essay during the revision posttesting phase, and one essay during the maintenance phase. Moreover, the effect was replicated across twelve participants in four groups. Adding to the quality of the study, reliability of scoring was reported for all variables and was consistently above the minimum standard of 80% inter-observer agreement established by Horner et al. (2005). However, this study failed to meet expectations for establishing a clear and predictable baseline for all students, which would adequately demonstrate the effect of the intervention on student performance. Specifically, three students did not produce the same number of baseline essays as others in their groups. In addition, several students had one essay on which they performed outside of the range set for a stable baseline on the essay parts measure, including Andy, Ally, John, Mario, and Mary Kate. While erratic academic performance is common for students with EBD, it makes it difficult to rule out the possibility that they could have performed at a higher level without the writing intervention.

Limitations

There were a number of limitations to this study. First, students could have benefited from more explicit instruction, and more instructional time, in the revision phase. This conclusion is drawn for several reasons. To begin, students only made slight increases in meaning-change revisions following instruction. Previous revision research suggests that instruction in revision results in more meaning-change revisions. Students may not have had enough explicit instruction in how to improve these deficiencies, or enough practice in the revision strategy, to allow them to fully develop this skill. Additionally, the number of revisions students made dropped considerably at
maintenance testing. With additional practice, students may have been able “over-learn” this skill in order to maintain it over time. Moreover, students continued to make improvements in essay quality and length at maintenance testing, which may signal that they needed more time to fully develop these writing skills.

Additionally, while participants all met the criteria of performing below their peers in writing on a daily basis, resulting in their placement in a self-contained class, there were a variety of skill levels across students at the beginning of the study. These different skill levels may have led to different impacts of instruction on students’ persuasive writing skill. That is, for students who entered the study with higher levels of skill, such as John and Ally, instruction seems to have made their performance more consistent over time. For other students, such as Tim, Brian, and Clare, instruction resulted in steady gains over time, even continuing into maintenance testing. The inclusion criteria were not as stringent as they could have been to obtain more consistent results across students over time. With so few students in the study, generalizable conclusions about the differences in the impact of instruction on various subgroups of students cannot be made.

Lastly, missing data points for some students during baseline testing raise questions of the validity of the study. Conducting research with students with EBD who struggle to complete assigned tasks, as these students all did, poses challenges. Unfortunately, in this case, students’ refusal to complete written tasks, or their choice not to follow given directions, negatively impacted this study. On a more positive note, these
refusals and the lack of following directions decreased over time as students’ writing improved.

Implications for Practice

The findings from this study inform teacher practice in several ways. First, including revision instruction as an integral part of writing instruction is important. Students had fewer mechanical errors as a result of instruction. Additionally, for some students, such as Tim and Clare, revision instruction may also improve other aspects of students’ writing. Or, like Andy, Ally, Navid, Mario, and Mary Kate, revision instruction may make students more consistent in their ability to write high-quality essays.

Next, although writing studies for students with disabilities typically show large gains for students whose writing skills are well below grade level, this and other studies (e.g., Mason et al., 2010) suggest that SRSD instruction may benefit stronger writers as well. For these writers, like Ally and John, instruction in SRSD may improve their ability to consistently produce higher quality written work. Therefore, targeted interventions such as this need not be reserved for only the lowest performing students; many students benefit from such direct, explicit writing instruction.

Lastly, teachers should provide structured opportunities for students with EBD to work with their peers. Not only can this improve the amount of time students are on task, it can also improve their writing skills. Furthermore, teachers should consider alternative arrangements for grouping students for such work. In this study, grouping students by
their ability to work well together, rather than by skill level, may have contributed to the positive behavioral results.

Directions for Future Research

This study raises additional questions about revision instruction for future research to address. First, more work needs to be done to refine the revision strategy. One suggestion provided by students was to develop a mnemonic for the revision strategy. Another necessary change is to combine the revising and editing conferences, as this seemed to be a more natural fit for students. On a deeper, more important level, the strategy should be refined to encourage more meaning-change revisions that improve the overall quality of students’ essays. While the revision strategy was effective in decreasing the number mechanical errors in students’ essays, and increased the number of surface-level revisions, a more powerful revision strategy would have a positive impact on the content and quality of essays as well. Perhaps this can be achieved by including more explicit instruction on how to make revisions, in addition to instruction on identifying what needs to be revised.

Second, future research on revision strategies implemented as part of a writing package should explore how instruction might be differentiated for groups of students within the same classroom. As in most intact classrooms, students entered this study with a range of writing skills. Revision instruction greatly improved the level of performance of low writers, while increasing the consistency of performance for writers with more skill. These differences suggest that writers may need differentiated instruction based on
their skill level. Future research should explore the essential components of revision
instruction for lower level writers as well as higher level writers, and fine tune
instructional strategies to meet their different needs.

Finally, while this study provides evidence that students with EBD are capable of
working positively and effectively with one another, more research is needed to support
this claim. Replication of the academic effectiveness of grouping students together based
on social compatibility is necessary. Additionally, the impact of other types of grouping
(e.g., by ability, with general education peers) on on-task behavior and academic
achievement should be explored.

Summary

Students made both academic and behavioral gains as a result of instruction in
SRSD and the revision strategy. Following SRSD instruction, students showed
improvement on the essay length, essay parts, and holistic quality measures, similar to
students in previous SRSD studies. Following revision instruction, students made
additional gains on the mechanical errors measure. These gains were on par with gains
made by students in other revision studies. When analyzing the number and types of
revisions students made, students in the present study primarily made surface-level
revisions following revision instruction. In contrast, students in previous revisions
studies made more meaning-change revisions following instruction. As in previous
SRSD studies, students in this study found the interventions useful and enjoyable.
On the self-efficacy and fluency measures, students in this study did not make any gains. In previous studies, students made gains on the fluency measure when they had been explicitly taught a writing fluency strategy or when they were in the writing intervention for a considerable amount of time. Very little information is available on the impact of SRSD instruction on the self-efficacy for writing for students with EBD. Finally, students were most on-task when working with their peers, and were able to complete nearly all of the steps of the peer-revision process consistently.

The results of this study point to several implications for teacher practice. That is, revision instruction is an important component of writing instruction for students with EBD. Furthermore, a wide range of students can benefit from instruction, not only the most struggling writers. Additionally, peer learning opportunities are beneficial for middle school students with EBD, who benefit both academically and behaviorally from such arrangements. Future research is needed to further refine the revision strategy to maximize its effectiveness. More research is also needed to establish peer-mediated instruction as an evidence based practice for students with EBD.
APPENDIX A

Sample Persuasive Essay Writing Prompts

Directions: Choose one of the topics below and circle your choice. Write a good persuasive essay about the topic.

Should your family do a staycation (i.e., vacation at home with day trips) instead of going away for a vacation?

OR

Should all kids your age be required to play a sport to stay physically active?
APPENDIX B

Essay Scoring Conventions

Number of words: Use the word count feature on Microsoft Word to determine number of words.

- If a student has copied the prompt question verbatim, do not include that in the word count.
- Do not include words in [brackets]. They are meant to tell the reader what the previous word is, if the spelling is difficult to decipher.

Number of Parts: Assign one point for each of the following:

- Topic sentence
- Reasons (1 point for each unique reason)
- Explanations for the reasons (1 point if the author explained the reason by providing more detail or an example. Explanations may be multiple sentences long.)
- Counter reasons
- Explanation for counter reason
- Refute
- Ending

Holistic Quality Score: See holistic quality rubric in Appendix C.

Number of Punctuation Errors: Count the number of sentences that do not end in a period, question mark, or exclamation point. That is the ONLY punctuation error that will be counted.

- Run-on sentences can be split into two sentence, resulting in a punctuation error where a period or question mark should have been placed. A run-on sentence will be defined as three or more independent clauses put together without proper punctuation or conjunctions.
- Note: This number will be divided by the number of words to get a proportion of errors per 100 words.
Number of Capitalization Errors: Count the number of sentences that do not begin with a capital letter. That is the ONLY capitalization error that will be counted.

- Run-on sentences can be split into two sentence, resulting in a capitalization error where a new sentence should have begun. A run-on sentence will be defined as three or more independent clauses put together without proper punctuation or conjunctions.
- Note: This number will be divided by the number of words to get a proportion of errors per 100 words.

Number of Spelling Errors: Count all words spelled incorrectly EXCEPT for people’s names.

- Count as errors the wrong form of a word with multiple spellings (e.g., their/there/they’re, your/you’re, hear/her). 
- Count as error if apostrophe not included in a contraction (e.g., cant, wont).
- The total number of errors will be divided by the number of words to get a proportion of errors per 100 words.

Total Number of Errors: Add up the following to determine the total number of errors:  
*Number of punctuation errors + number of capitalization errors + number of spelling errors.*

- This number will be divided by the number of words to get a proportion of errors per 100 words.
APPENDIX C

Holistic Quality Scoring Rubric

Score of 10. Persuasive essay includes topic sentence, more than three reasons with at least three explanations, and an ending sentence. Essay is written in a logical sequence that strengthens the writer’s argument. Writer uses more than one counter argument/point in the essay.

Score of 9. Persuasive essay includes topic sentence, more than three reasons, at least 3 explanations, and an ending sentence. Essay is written in a logical sequence that strengthens the writer’s argument. Writer uses 1 counter argument/point in the essay.

Score of 8. Persuasive essay includes topic sentence, more than three reasons, at least 2 explanations, and an ending sentence. Essay is written in a logical sequence that strengthens the writer’s argument.

Score of 7. Persuasive essay includes topic sentence, three reasons with at least two explanations, and ending sentence. Essay is written in a logical sequence that strengthens the writer’s argument.

Score of 6. Persuasive essay includes topic sentence, three reasons with at least 1 explanation, and ending sentence. Essay’s sequence is weak, therefore limiting the writer’s argument.

Score of 5. Persuasive essay includes topic sentence, three reasons, and ending sentence.

Score of 4. Persuasive essay includes four of the following parts: topic sentence, reasons, and ending sentence.

Score of 3. Persuasive essay includes three of the following parts: topic sentence, reasons, and ending sentence.

Score of 2. Persuasive essay includes two of the following parts: topic sentence, reasons, and ending sentence.
Score of 1. Persuasive essay includes one of the following parts: topic sentence, reason, and ending sentence.

Score of 0. No essay parts.
APPENDIX D

*Persuasive Writing Self-efficacy Scale*

**Directions:** Read the essay prompts below and circle your choice. Answer the questions below based on how you feel about your writing on this topic.

Should your family do a staycation (i.e., vacation at home with day trips) instead of going away for a vacation?  

OR  

Should all kids your age be required to play a sport to stay physically active?

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I will be able to start writing this essay without any difficulty.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2.</td>
<td>Even if I don’t like this topic, I will still be able to write a good essay about it.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>I can write a good plan for this essay.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>I can write a good introduction sentence on this topic.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5.</td>
<td>I will be able to correct all the grammatical errors in my essay.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6.</td>
<td>In writing an essay on this topic, I will be able to decide what goes first, second, third, and so on.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
7. If I get stuck while I am writing, I will be able to find ways to overcome the problem.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

8. When writing an essay on this topic, I can think of good reasons to persuade the reader.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

9. I can write a good essay on this topic.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

10. I will be able to rewrite my confusing sentences clearly.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

11. When writing this essay I will be able to think of agreements for both sides of the topic.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

12. I will be able to revise my first draft of this paper to make a better-organized essay.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

13. I will write good transitional sentences from one idea to another.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

14. I will be able to manage my time effectively to finish this paper during this class period.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

15. I will be able to concentrate on my writing even if there are many distractions around me.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

16. I will be able to find someone to give me ideas about how to make my paper better.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral or Maybe</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

249
APPENDIX E

Woodcock Johnson III Writing Fluency Subtest Sample Questions

<table>
<thead>
<tr>
<th>Sample A:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cake</td>
<td></td>
<td></td>
</tr>
<tr>
<td>is</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The cake is good.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample B:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pig</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>is</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample C:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>this</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td>big</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample D:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>bell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ringing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX F

On-task Behavior Coding Sheet

date:  
Group:  
Start:  
end:  
observer:  

|    | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 30 |
|----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

CODES: 1= on task, 2= off task, 3= out of the room, and 4= when student comes back into the room from being out.

Notes:  
Today's lesson covered…

Brief description of students and classrooms…
APPENDIX G

*Social Validity Interview*

Directions: Tell students you are going to ask them some questions about what they learned about writing.

1. Tell me the writing strategy that you learned to use.
2. Draw a picture of the graphic organizer we used
3. What did you like most about this strategy?
4. Has using the POW+TREE strategy helped you become a better writer? How?
5. How do you think POW+TREE could help other students?
6. If you were the teacher, would you add anything to help students learn to write?
7. If you were the teacher, what would you change in the POW+TREE lessons? Why?
8. From the POW+TREE lessons, what things have most helped you become a better writer?
9. Have you used POW+TREE in any other classes? If yes, ask, what other classes or assignments and how has it helped?
10. Tell me how you have used counter arguments in your writing. Why are counter arguments important?
11. Tell me about the revision strategy you learned.
12. What about the revising strategy helped you make you writing better? What didn’t?
13. How did working with a partner help you revise?

252
APPENDIX H

POW + TREE Mnemonic
## Graphic Organizer

<table>
<thead>
<tr>
<th>T</th>
<th>TOPIC</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>What do I believe?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TW</th>
<th>REASONS</th>
<th>- 3 or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Why do I believe this?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Will my readers believe this?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>EXPLANATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Say more about each reason.</td>
</tr>
<tr>
<td></td>
<td>What details will persuade my reader?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CR</th>
<th>Counter Reason</th>
<th>- 1 or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Who might disagree and why?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>EXPLANATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Say more about the counter reason(s).</td>
</tr>
</tbody>
</table>

**Refute it!** Tell why your side is better?

<table>
<thead>
<tr>
<th>E</th>
<th>ENDING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What do I want my reader to remember?</td>
</tr>
</tbody>
</table>
Self-statement Sheet

My Self-Statements

To think of good ideas:

While I work:

To check my work:
APPENDIX K

Self-monitoring Graph

Goal
13
12
11
10
9
8
7
6
5
4
3
2
1
Date
Trans
Words

256


### Transition Word List

#### Transition Words

- **To show a reason**
  - First
  - Second
  - Third
  - In addition
  - Another
  - To begin
  - Also
  - Furthermore
  - Next
  - Finally
  - My final
  - Lastly

#### Show a counter reason

- However
- Nevertheless
- Conversely
- On the contrary
- Yet
- Instead
- On the other hand

#### Conclusion

- In conclusion
- In short
- On the whole
- To summarize
- In general
- In sum
Good Writing Handout

Good Writing . . .

Looks Like
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Sounds Like
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Makes Me Feel
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
APPENDIX N

Revision Strategy Guide

Author ______________________ Reviewer ______________________

1. COMPLIMENT What did the author do well?

2. Does it have all the PARTS? Mark-up the essay. Then, put a "v" next to the parts you see.

3. Is it CLEAR? + = yes; 0 = no

4. Is it PERSUASIVE? + = yes; 0 = no

<table>
<thead>
<tr>
<th>T</th>
<th>Topic - What does the author believe?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>v</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TW</th>
<th>Reason E Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CR</th>
<th>Counter Reason E Explanation - Refute the CR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>Ending - Restate the topic and reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. COMPLIMENT What did the author do well?

6. Editing

   CAPITALIZATION Does every sentence start with a capital letter?
   PUNCTUATION Does every sentence end with an end mark (.,?))?
   SPELLING Did you check your spelling?
APPENDIX O

Peer-revision Checklist

Peer-Conferencing Checklist

☐ Find a place in the classroom to conference where you will not distract other students.

☐ Turn on the audio recorder to record your conference.

☐ Use quiet voices. Be respectful of other students who are working.

☐ Take turns reading your essays to each other.

☐ Re-read your partner’s essay to yourself and fill out the POW+TREE Revision sheet.

☐ Conference with your partner about your suggestions.

☐ Revise your own essay on the computer. Save your work. Print out a copy.

☐ Meet with the teacher to talk about your revisions.

☐ When both partners are done revising on the computer, meet again to edit your partner’s essay for capitalization, punctuation, and spelling errors. Remember to record your conference.

☐ Make your final corrections on the computer. Save your work. Print out a final copy and turn it in to your teacher.

EXPECTATIONS:

❖ TRY HARD
❖ DO YOUR BEST
❖ BE POSITIVE
❖ BE KIND AND ENCOURAGING TO PARTNERS
❖ WHEN IN DOUBT, RAISE YOUR HAND FOR ASSISTANCE
Conferencing Feedback Guide

1 & 5. Give a COMPLIMENT.
   + I like the way you . . .
   + I noticed that your essay . . .
   + One thing that good writers do that you did is . . .
   + Your essay is well written.

2. Does it have all the PARTS?
   + You have all the parts.
   + It was very easy to find all of your essay parts.
   0 I cannot find your topic sentence. Can you show me your topic sentence?
   0 I did not see three reasons. I only saw _____ (1 or 2).
   0 You had a reason, but I did not see an explanation.
   0 I did not see a counter reason. Who would disagree with your opinion?
   0 I did not notice transition words for all your reasons or for your ending.
   0 You did not include an ending.
   + You are only missing one part.

3. Is it CLEAR?
   + You stated your topic clearly.
   + When I read this reason and explanation, I got a clear picture in my head of what
     you were trying to say.
   0 This sentence does not make sense to me. Is there a better way to say it?
   + Your writing flows very well.
   0 I did not understand this part. What were you trying to say?
   0 I am confused but maybe you can explain this to me?

4. Is it PERSUASIVE?
   + This reason is very persuasive.
   0 This might be a strong reason, but it needs more details to support it.
   0 It seems like you repeated the same reason twice.
   • This reason does not match the topic sentence that you wrote.
Strategy Knowledge Questionnaire

What do you look for when you revise a persuasive essay?
### APPENDIX R

**Sample Fidelity of Treatment Checklist**

**Lesson 1: Introduction to Revision**

<table>
<thead>
<tr>
<th>Group</th>
<th>Observer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date began</td>
<td>Date completed</td>
</tr>
<tr>
<td>Total number of session to complete lesson</td>
<td>Comments:</td>
</tr>
</tbody>
</table>

1. Discuss what revision is and why it is important. Discuss the following questions: What is revision? Why is it important to revise? When should you revise? What should you look for when you revise?

2. Brainstorm list of what good writing looks like. Teacher writes items on the Good Writing poster while students write them on their individual handouts.

3. Introduce revision strategy. Teacher posts Revision Strategy poster and explains each item on the poster. Teacher then hands out copies for students to put in their folders.

4. Practice completing peer-revision feedback form. Students read *Skateboarding at the Mall* essay and class goes through the first part of the revision strategy (i.e., compliment, parts, clarity, reasons, compliment). (Note that capitalization, punctuation, and spelling are covered in lesson 2.)

5. Make revision on the computer. Students take turns making suggested changes on the computer.

6. Review parts of the revision strategy. Verbally review parts of the revision strategy.

7. Lesson Wrap-up. Teacher verbally reviews what was covered during the lesson.

Total number of lesson parts accurately completed (out of 7):

_____
APPENDIX S

*Peer-revision Steps Completed Checklist*

<table>
<thead>
<tr>
<th>Step</th>
<th>Student 1</th>
<th>Student 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Students read their essays to each other</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Students filled out the peer-revision checklist individually</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Students reviewed the peer-revision checklist with their partner-author, including:</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Compliment</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Parts</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Clarity</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Reasons</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Compliment</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Students revised essays on computer</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Students conference individually with teacher</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Students met again to edit work, including reviewing:</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Capitalization</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Punctuation</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Spelling</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Students produced final copies of essays to turn in</td>
<td></td>
</tr>
</tbody>
</table>

Total number of components completed (out of 26) ________
REFERENCES
REFERENCES


269


273


CURRICULUM VITAE

Sara J. Mills earned her Bachelor of Science degree from Northwestern University and her Master of Teaching degree from the University of Virginia. Before enrolling in the PhD in Education program at George Mason University, Ms. Mills taught special education to elementary-aged students in Fairfax County, Virginia. She also served as a Mentor Resource Teacher, providing intensive, job-embedded professional development to first year special education teachers.

As a doctoral student, Ms. Mills was selected to receive the Special Education Doctoral Leadership Cohort Program Fellowship, under the direction of Dr. Margo Mastropieri and Dr. Thomas Scruggs. The fellowship was funded by the U.S. Department of Education, Office of Special Education Programs Grant (H324D070008). She also worked with Dr. Mastropieri on the GMU-PSU Writing Project Grant (R324A070199-07) from the U.S. Department of Education, investigating writing instruction for middle school students with emotional and behavioral disorders.

Ms. Mills currently works as a Manager in the Instructional Coaching Program in District of Columbia Public Schools. In this capacity, she supervises, trains, and supports Instructional Coaches in 10 elementary and middle schools as they work to improve teaching and learning. Additionally, Ms. Mills works as an adjunct instructor at George Mason University’s College of Education and Human Development.