

TEACHING STUDENTS WITH EMOTIONAL AND BEHAVIOR DISORDER HOW  
TO WRITE PERSUASIVE ESSAYS FLUENTLY

by

Nancy Irby Cerar  
A Dissertation  
Submitted to the  
Graduate Faculty  
of  
George Mason University  
in Partial Fulfillment of  
The Requirements for the Degree  
of  
Doctor of Philosophy  
Education

Committee:

Margaret Mastropieri

Chair

The G. Gage

Christopher H. Kenton

Gary R. Galluzzo

Program Director

Marjorie J. Smith

Dean, College of Education  
and Human Development

Date: April 11, 2012

Spring Semester 2012  
George Mason University  
Fairfax, VA

Teaching Students with Emotional and Behavior Disorder How to Write Persuasive  
Essays Fluently

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

By

Nancy Irby Cerar  
Master of Education  
George Mason University, 2007  
Bachelor of Science  
George Mason University, 2005

Director: Margo A. Mastropieri, Professor  
Graduate School of Education

Spring Semester 2012  
George Mason University  
Fairfax, VA

## DEDICATION

This is dedicated to my beloved husband, Chad, our child, Everett, and the rest of my family. Everything I am and have accomplished in my life would not have been possible without you all by my side. Thank you for always believing in me and for the endless love, encouragement, and support you have provided me throughout this process.

## ACKNOWLEDGEMENTS

Many people crossed my path and were an enormous support during this process. I want to express my deepest appreciation to those who walked with me during this doctoral journey and helped me succeed. First, I want to express my appreciation to Dr. Margo A. Mastropieri for always guiding me in the right direction and making sure that I stayed on the track. Margo, thank you for working and mentoring me for the last five years. You have taught me so much over the years. I cherish the work we have done together, and the words of encouragement you always gave to me. You challenged me to understand and conduct single subject research to a high quality standard. In this, you had me answer my own question about single subject research while providing me with guidance. I have become a stronger researcher because of this and it has been a great honor working with you.

I also want to express my appreciation to my other committee members, Dr. Thomas E. Scruggs and Dr. Anastasia Kitsantas. Thank you both for your support during my years in the doctoral program and during the dissertation phase. Also, thank you both so much for giving me the opportunity to work with you.

The generous financial support I received made it possible to fund my doctoral studies and my dissertation research. Margo and Tom, thank you for selecting me as one of the recipients for funding through your Doctoral Leadership Grant in Special Education. I also want to thank the Graduate School of Education for awarding me fellowships, and the dissertation fellowship awards.

I am appreciative to several friends who assisted and supported me at important stages during my study: Mary, Catherine, Stephanie, Shannon, Sara, Dani, and Clara. First, I wanted to specifically thank both Mary and Catherine for helping me tremendously with my dissertation. The study would not have been possible without both of you. Thank you to everyone else for being there for me and scoring dependent measures, conducting observations, and for providing emotional support and words of encouragement along the way. It has been a pleasure working with you all.

Partial support for this research was provided by grants: Grant No. H325D070008 from the U.S. Department of Education, Office of Special Education Programs to George Mason University and Grant No. R324A070199-07 from the U.S. Department of Education, Institute of Educational Sciences awarded to Pennsylvania State University with a subcontract to George Mason University. This represents the views of the author and is not an endorsement by either funding agency.

## TABLE OF CONTENTS

	Page
List of Tables.....	vi
List of Figures.....	vii
Abstract.....	viii
 1. Introduction.....	 1
Statement of the Problem.....	2
Purpose.....	8
Research Questions.....	9
Definition of Terms.....	9
 2. Literature Review.....	 12
Students with Emotional Behavioral Disorders.....	12
Characteristics.....	14
Academic Setting.....	15
Research on Academic Outcomes and Students with Emotional Behavioral Disorders.....	16
Fluency Instruction and Automaticity .....	18
Reading Fluency .....	19
Writing Automaticity.....	23
Writing Instruction and Self-Regulated Strategy Development .....	24
Self-Regulated Strategy Development with Students with Emotional Behavioral Disorders or at-risk for EBD.....	30
Research Questions .....	41
 3. Methodology .....	 42
Research Design.....	42
Setting .....	44
Participants.....	44
Materials .....	47
Data Sources and Scoring Procedures .....	49
Procedures.....	55
 4. Results.....	 66
Data Analyses .....	66
Treatment Fidelity.....	66
On-Task Behavior.....	67

Standardized Test.....	68
Writing Performance.....	68
Self-Efficacy .....	124
Social Validity, Student Interviews, and Strategy Report .....	125
5. Discussion .....	129
Standardized Testing.....	129
Post-Fluency Instructional Performance.....	130
Post-Multiple Paragraph Instruction Performance.....	138
Fluency Maintenance and Generalization Performance .....	142
Multiple Paragraph Maintenance and Generalization Performance .....	145
Self-Efficacy .....	147
Social Validity .....	148
On-Task Behavior.....	148
General Discussion .....	149
Educational Implications .....	151
Limitations .....	151
Recommendations for Future Research .....	152
Appendices.....	153
Appendix A: POW + TREE Mnemonic Chart .....	153
Appendix B: POW + TREE Graphic Organizer .....	154
Appendix C: Transition Word List .....	155
Appendix D: Self-Statement Sheet .....	156
Appendix E: POW + TREE Record Sheet.....	157
Appendix F: Sample Lesson Plan: Modeling Fluency Instruction .....	158
Appendix G: Sample Lesson Plan: Modeling Multiple Paragraph Instruction .....	161
Appendix H: Persuasive Writing Self-Efficacy Measure .....	167
Appendix I: Social Validity Questionnaire.....	168
References .....	170

## LIST OF TABLES

Table	Page
1. Student Demographic Characteristics.....	46
2. Fluency Results for the Overall Student Essay Results of timed essay by phase: baseline, post-fluency, maintenance, and generalization .....	70
3. Multiple Paragraph Results for the Overall Student Essay Results of untimed essay by phase: baseline, post-multiple paragraph, maintenance, and generalization .....	72
4. Percentage of Nonoverlapping data by participant by testing phase: post-fluency, post-multiple paragraph, fluency maintenance, fluency generalization, multiple paragraph maintenance, and multiple paragraph generalization compared to baseline.....	83
5. Fluency percent of increase or decrease based on overall means for number of parts, number of words, and holistic quality from baseline to post-test, maintenance, and generalization.....	132
6. Percent of increase or decrease based for multiple paragraph studies on overall means for number of parts, number of words, and holistic quality from baseline to post-test in previous and current studies.....	139

## LIST OF FIGURES

Figure	Page
1. Total number of words by participant by phase: baseline, instruction, post-fluency, and post-multiple paragraph, maintenance, and generalization.....	73
2. Total number of persuasive essay parts by participant by phase: baseline, training, post-fluency, and post-multiple paragraph, maintenance, and generalization.....	74
3. Holistic quality score by participant by phase: baseline, instruction, post-fluency, and post-multiple paragraph, maintenance, and generalization.....	75



## ABSTRACT

### TEACHING STUDENTS WITH EMOTIONAL AND BEHAVIOR DISORDER HOW TO WRITE PERSUASIVE ESSAYS FLUENTLY

Nancy Irby Cerar, Ph.D.

George Mason University, 2012

Dissertation Director: Dr. Margo A. Mastropieri

A multiprobe, multiple baseline design was used to evaluate the effectiveness of persuasive writing strategy instruction. Six middle school students with emotional and behavioral disabilities (EBD) received two instructional phases of Self-Regulated Strategy Development (SRSD) instruction for writing persuasive essays over 33 days of intervention. A single paragraph essay was introduced first, followed by multiple paragraph essay instruction, followed by a generalization lesson. Assessments included: (a) at least 14 essays across: baseline, post instruction phase one, phase two, maintenance, and generalization; (b) Woodcock Johnson fluency subtests; (c) a self-efficacy measure; (d) sociality validity and student interviews; and (e) time on task during instruction measures. Assessments were scored and evaluated in several ways including: (a) essays by holistic quality, length, and number of persuasive essay elements; (b) Woodcock Johnson fluency subtest at pre- and post- testing; (c) self-efficacy at pre-, post-, and maintenance testing. Findings revealed positive effects for: (a) all essay

measures at post-instructional phases, maintenance, and generalization testing periods; (b), the Woodcock Johnson fluency subtest; and (c) on the self-efficacy measure. These findings replicated and extended previously conducted written expression research with middle school students with EBD. Most importantly, results revealed that instructional order of single or multiple paragraphs appeared to work equally well when findings are compared with previous research and that students with severe EBD require extensive, intensive instruction. Implications for education of students with EBD and future research are also presented.

## 1. INTRODUCTION

Numerous researchers have been studying writing instruction to improve poor writer's ability since the 1980's (e.g., De La Paz, 1999; De la Paz, 2001; Harris & Graham, 1999; Graham & Harris, 2003; Lane, et al., 2008). The research has ranged in the writing instruction focus from planning, to paragraph writing, to organizing, to different strategies, to mnemonics, to peer editing, to revising strategies, to attributions, and to self-regulation. A recent strategy instruction meta-analysis found 15 studies published in the 1980's, 34 studies published in the 1990's, and 13 studies published from 2000-2005 (Graham, 2006). The increase focus on writing could be due to No Child Left Behind (NCLB) Act of 2007; or due to the increased number of students with learning disabilities, that increased by 1.5% of the total population (6.5 million) of individuals with disabilities (U.S. Department of Education, 2009). NCLB focuses teachers' teaching ability to meet student needs in subject areas such as, English, mathematics, and writing.

NCLB requires the performance of students with disabilities on the standardized state assessment be categorized separately. The category of students with disabilities is composed partly of individuals with specific learning disabilities in areas such as reading, writing, English, mathematics, science, and other disability areas. This category has

forced schools to examine their ability to meet the academic needs of students with disabilities.

### **Statement of the Problem**

The purpose of the present study is to address the academic needs of students with emotional disabilities in writing. Addressing the area of writing will help students on state standardized assessments and academically across subjects, since writing is an interdisciplinary skill that crosses subjects.

**Overall writing and writing interventions.** Strategy instruction is used to help students in a number of different subjects. One of those areas is writing. Reviews and meta-analyses that compare different ways of teaching writing have found that strategy instruction is one of the most effective ways to teach writing to students. For example, Graham's (2006) meta-analysis found there were 39 studies on writing, with 19 single-subject design studies and 20 group comparison studies. A majority of the studies targeted students with learning disabilities (LD) or students who are struggling writers in elementary or secondary school. Results found that strategy instruction produced the largest effects on writing for students with LD (effect size (ES) = 1.32) compared to all students (ES = 1.15). He also found a difference in the different types of students and how they performed on writing measures. The effect of strategy instruction on the writing measures for quality and elements for all students had a moderate to large effect on their performance (ES = .80 and .60 correspondingly). Length of essays varied across studies. It was found that students with LD had the greatest impact from the strategy instruction on length (ES = .97), while poor and average students obtained a moderate

improvement on length ( $ES = .54$  and  $.78$ ), with a negative to no improvement on good writer's essay length ( $-.002$ ). SRSD was used 45% for group studies and 68% for single subject studies. These SRSD studies yielded greater improvements than studies that used other types of writing strategy instruction.

For instance, another recent review of writing with adolescents (Graham & Perin, 2007) found 132 studies that fit their inclusion criteria and reported that 19% of writing studies used strategy instruction to teach writing and 40% of the 20 writing studies used SRSD to teach writing.

Meta-analysis (Graham & Perin, 2007) supports the findings of Graham (2006). They included writing studies that: (a) had participants in grades four through twelve, (b) attended a regular or private school, (c) included a writing measure on quality, (d) reported a quality measure's reliability, (e) used an experimental or quasi-experimental design, and (f) provided ES or data to calculate ESs. The most effective writing instruction for adolescents was strategy instruction and summarization strategies ( $ES = .82$ ), followed by peer assistance ( $ES = .75$ ), setting product goals ( $ES = .70$ ), word processing ( $ES = .55$ ), sentence combining ( $ES = .50$ ), and inquiry, prewriting activities, and process writing approach ( $ES = .32$ ). In conclusion, the most impact on adolescent students' writing performance included strategy instruction or summarization strategies.

**Self-regulated strategy development in writing.** Dr. Steve Graham and Dr. Karen Harris developed the specific strategy called self-regulated strategy development (SRSD) in the 1980's. Graham and Harris found that good writers spend time planning, monitor their progress, evaluate their work, revise their work, and regulate through the

writing process. Struggling writers either skip parts of the writing process like planning, monitoring their progress, and revising or find these processes to be difficult for them (Graham & Harris, 2003). SRSD has been used for narrative writing, story writing, persuasive writing and report writing for students who are average writers, poor writers, gifted, and with and without disabilities in grades second through eighth (De La Paz, 1999; 2001; Graham & Harris, 2003; Harris, Graham, & Mason, 2006; Mason, Snyder, Sukhram, & Kedem, 2006).

Two recent meta-analyses examined SRSD studies with students with LD (Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009; Graham & Harris, 2003). Each synthesis had different inclusion criteria. Graham and Harris (2003) included 18 studies in their analyses and Baker et al. (2009) included 21 studies.

Graham and Harris's (2003) found that 72% of the studies involved students with LD in grades fourth through eighth grade. Furthermore, Graham and Harris (2003) looked at the effects of SRSD on different genres of writing. They found seven studies taught story writing with two group comparisons studies and five single subject studies that had a large effect on (ES = 1.47-3.52 or 71-100% percent of nonoverlapping data (PND)) students' length, elements, grammar, and quality of writing. For opinion essay writing six studies with two group comparison studies and four single subject studies obtained moderate to high effects on (ES = .32-5.18 or 70-100% PND) students' length, elements, coherence, and quality. In conclusion, this review of the SRSD literature illustrates the effectiveness of SRSD instruction for story writing and opinion essays for

students with LD, but the effectiveness of the interventions varied from medium to very effective.

Additional analyses looked at student type (overall, LD, and poor writers) on writing measures on quality, elements, story grammar, and length. Interventions for all students were highly effective for: (a) quality (ES = 1.47 or 97% PND), (b) elements (ES = 1.97 or 92% PND), (c) story grammar (ES = 3.52 or 100% PND), and (d) length (ES = 2.07 or 82% PND). For students with LD interventions were highly effective for: (a) quality (ES = 1.14 or 89% PND), (b) elements (ES = 2.15 or 93% PND), (c) story grammar (ES = 3.52 or 100% PND), and (d) length (ES = 1.86 or 89% PND). Interventions for poor writers were also highly effective for: (a) quality (ES = 1.67 or 100% PND), (b) elements (ES = 1.42 or 100% PND), and (c) length (ES = 2.02 or 100% PND). Baker's et al. (2009) meta-analysis found similar effect sizes for the five group comparison studies that Graham and Harris (2003) found. In conclusion, Graham and Harris (2003) meta-analysis found that SRSD instruction was very effect at improving the writing performance of all students with and without disabilities.

**SRSD instruction for students with or at risk for EBD.** SRSD instruction for students with or at risk for EBD had been studied 10 times (Adkins, 2005; Graham, Harris, & Mason, 2005; Harris et al., 2006; Lane, Harris, et al., 2008). Of those 10 studies, only three examined the effectiveness of SRSD on secondary students with EBD. These three studies' methodologies, instruction, and focuses varied from one another, even though all investigated the effects of SRSD instruction for persuasive writing using the mnemonic POW+TREE. The mnemonic POW+TREE stands for (P) pick your idea,

(O) organize your notes, (W) write and say more, (T) topic sentence, (R) reasons, (E) ending, and (E) examine.

The first study that was conducted using SRSD for students with EBD (Mastropieri et al. 2010) used a design experiment methodology to explore the effects of the SRSD intervention for persuasive writing. This study taught students in a large group of 8-10 students that was conducted by the teacher initially, but transferred to the project staff. Students' persuasive writing abilities were measured pre-intervention and post-intervention with the intervention lasting approximately four months. Students had a mean on-task behavior of 86% and missed instruction a mean of 32%, yet students made statistically significant improvement from pre to post fluency subtest of the Woodcock Johnson-III ( $ES = .67$ ). Students also performed statistically significantly higher from pre- to post- essay measures ( $ES = 1.28 - 2.46$ ) on the number of words, number of parts, number of paragraphs, number of transition words, and holistic scoring. Students were able to maintain their higher performance at maintenance testing ( $ES = .66 - 2.11$ ) than pre-testing, but maintenance testing performance was lower than the posttest essay measures.

Another study, Mastropieri et al. (2009) taught twelve eighth grade students with EBD to write better persuasive essays. Mastropieri et al. (2009) replicated and extended Mastropieri (2010), but used a multiple baseline design to teach students with EBD to write persuasive essay using the SRSD instructional model. The intervention taught students in small groups of three to five students who were taught by trained graduate student researchers. Students received instruction for approximately 50 days four days a



week for 30-minutes. The SRSD intervention was twofold in that strategy first phase taught students to plan and write a multiple paragraph essays followed by a second phase that taught students to plan and write a one paragraph essay in 10-minutes. All students improved on all writing measures from baseline to post intervention on words, sentences, paragraphs, transition words, persuasive essay elements, and quality after each instructional phase. Fluency results were very effective ( $ES = 1.53 - 2.47$ ) and somewhat higher than after the first instruction phase of multiple paragraph writing ( $ES = .87 - 1.74$ ). More specifically, students improved the most on: (a) parts ( $ES = 2.47$ ), (b) transition words ( $ES = 2.46$ ), (c) quality ( $ES = 2.22$ ), and (d) words ( $ES = 2.22$ ) for the students' 10-minute timed writing compared to their first phase of multiple paragraph untimed posttest essays interventions.

Mastropieri et al. (2009) adds to the findings of Mastropieri et al. (2010) in that SRSD instruction for multiple paragraph persuasive writing was highly effective for middle school students with serious EBD. However, instruction to teach students with EBD to write multiple paragraph instruction was very intense and required 50-55 days of instruction.

A third study, Mason, Kubina, Valasa, and Cramer (2010) also taught middle school students with EBD to write persuasive essays using the SRSD model. The focus of this study was on the quick write or fluency instruction for students to plan and write a one paragraph essay in 10-minutes. The participating students were five Caucasian students that were four males and one female that ranged in age from 12.1 to 14.5 years old who attended an alternative school. The SRSD intervention for persuasive writing

was taught one-to-one by trained graduate researchers for five lessons for SRSD and three additional lessons for quick writing. Students improved from baseline testing to post-testing on holistic quality scores (84.0% PND), number of elements (42.9% PND), and number of words (40% PND). These results were not maintained over time when students completed maintenance testing two weeks after post-testing (holistic quality score 60% PND, number of elements 0% PND, number of words 0% PND).

Mason et al. (2010) adds to the Mastropieri and colleagues (Mastropieri et al., 2010; 2009) studies on the effectiveness of SRSD for persuasive writing for students with EBD. Mason et al. (2010) illustrated students with EBD were able to improve their one paragraph quick writing skills in a small number of sessions while Mastropieri and colleagues illustrated that multiple paragraph persuasive writing instruction take more intensive instruction. The students in the Mason studies appeared higher functioning, however, but Mastropieri et al. (2009) showed quick writing or fluency skills were taught quickly after learning the multiple paragraph essays.

## **Purpose**

The purpose of this study is to replicate and extend the literature in four important ways. First, the research seeks to replicate the effectiveness of SRSD instruction for persuasive writing with middle school student with EBD. Second, this study will extend the research by investigating the effects of the order of instructional phases for middle school students with EBD, by first teaching students to write a one paragraph essay followed by a second phase to teach student to write a multiple paragraph essay. Third, further examine the effects of SRSD on fluency or quick writing skills. Lastly, this study

will examine the students' self-efficacy beliefs about their persuasive writing abilities and the effects of SRSD for persuasive writing has on the students' self-efficacy beliefs.

### **Research Questions**

- 1) Does SRSD for persuasive writing, using POW+TREE, improve the persuasive writing abilities (as measured by length, content, and quality) of students and fluency abilities (including WJ) after instruction (instructional phase 1) and after instruction in expanding essay length (phase 2), and are they able to maintain and generalize learning?
- 2) Are students with EBD able to accurately conduct all steps of the POW+TREE strategy? Are students with EBD maintaining good time on task behavior? Does learning how to write persuasive essays using SRSD improve students' self-efficacy?
- 3) Do students find the POW+TREE strategies easy to implement, useful and enjoyable?

### **Definition of Terms**

**Students with Emotional and Behavior Disorders (EBD).** Many terms are used to describe emotional, behavioral or mental disorder. For this proposal, students with EBD will be identified for this study using the Federal definition and their identification by the school system. Students are categorized as EBD if they exhibit one or more of the following characteristics over a long period of time and to a marked degree that interferes with their academic performance: (a) an inability to learn that cannot be explained by intelligence, sensory, or health factors; (b) an inability to build or

maintain interpersonal relationships; (c) inappropriate behaviors or feelings under normal circumstances; (d) a general persistent mood of unhappiness or depression; or (e) a tendency to develop physical symptoms or fears associated with personal or school problems [Code of Federal Regulations, Title 34, Section 300.7(b)(9)]. Due to large variation of students with EBD descriptions of participants will be provided in the study.

**Self-Regulated Strategy Development (SRSD).** Instructional lessons incorporate the six stages of SRSD: (a) develop background knowledge, (b) discuss it, (c) model it, (d) memorize it, (e) support it, and (f) perform it. The SRSD approach incorporates scaffolding during the support it phase where the objective is to wean students off of all support material to perform the strategy independently without support material. Therefore the SRSD strategy incorporates instruction in self-regulation process of goal setting, self-instruction, self-monitoring, and self-reinforcement.

**Self-Efficacy.** Self-efficacy is our belief in our ability to succeed or achieve a goal in a specific situation, in the present study the situation is persuasive writing (Bandura, 2006). Therefore students are likely to engage in activities to the extent that they perceive themselves to be competent at those activities. In regards to education, this means that learners will be more likely to attempt, to persevere, and to be successful at tasks at which they have a sense of efficacy (Bandura, 2006).

**Persuasive Essay.** A persuasive essay is an essay that's objective is to convince the reader of some point of view, that includes a topic, reasons, explanations, and an ending.

**Mastery Performance.** The student becomes proficient in a topic to the best of his or her ability. For instruction mastery performance is obtained when a student writes two consecutive that includes a topic, at least three reasons, at least one counter reason, explanations for each reason, refute of counter reasons, and ending.

## 2. LITERATURE REVIEW

This chapter presents an overview of the literature on students with emotional and behavioral disorders (EBD) in relation to student characteristics, academics, and writing instruction. The first section provides an overview of students with EBD. The next section provides an overview of fluency instruction provided for reading and writing for students with disabilities and EBD. The final section provides an overview of self-regulated strategy development (SRSD) provided to students with disabilities and EBD and the connection of fluency to SRSD.

### **Students with Emotional Behavioral Disorders**

The classification of a student with *emotional disturbance* (which will be referred to as EBD) is defined by the U.S. government in the Individuals with Disabilities Educational Act (IDEA). IDEA defines emotional disturbance as follows:

- (i) Emotional disturbance means 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 or 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 of fear 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. (Code of Federal Regulations, Title 34, Section 300.7(b) (9))

However, students with emotional and behavioral disorder (EBD) are initially characteristics and seen by their social and behavioral characteristics: externalizing behavior, such as aggression, defiant, and compulsive, and internalizing behavior, such as anxiety, depression, and social withdrawal (Donovan & Cross, 2002; Little et al., 2010). Students with EBD struggle to self regulate their own behavior, let alone their academics. There is indication that students with EBD are at high risk for a wide range of negative outcome besides poor academic achievement, such as truancy, school dropout, motor vehicle accidents, unemployment, substance abuse, criminality, and welfare recipients (Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005; Walker et al., 1996; Walker & Severson, 2001).

Students with EBD were the fifth largest disability category in special education in 2004 which is less than 1% of the entire K-12 population who are served under the Twenty-eighth Annual Report to Congress on the Implantation of the Individuals with Disabilities Education Act (IDEA) (U.S. Department of Education, 2009). Students with EBD are continuously the disability with the highest dropout rates from 1994-1995 to 2003-2004. Every year the dropout rate is significantly higher than the second highest dropout rate. This population is the least likely to graduate than students with any other

disability and has a 52% dropout rate compared to a mean of 31% dropout rate across all disability categories (U.S. Department of Education, 2009).

**Characteristics.** The characteristics of students with EBD vary across sources. There are both government sources and research studies that examined the characteristics of students with EBD. The governmental sources included a congressional mandated committee that reviewed data from two national data sets (the Office of Special Education Programs and Office for Civil Rights) to provide a picture of the special education population by gender, ethnic, and disability categories (Donovan & Cross, 2002). Donovan and Cross (2002) found that all ethnic groups had equal chances of being identified with mental retardation and learning disabilities, but ethnic groups did not have equal chances for being identified with EBD. They reported that African American students were two times more likely to be identified with EBD than Caucasian students, but both Asian/Pacific Islander and Hispanics were identified as EBD at lower rates than Caucasians students (Donovan & Cross, 2002). In addition, Donovan and Cross (2002) also found the greatest disparity in gender in students with EBD. They found that 80% of the students with EBD were male who were identified at a three times higher rate than females.

Researchers have also examined the characteristics of students with EBD that participated in empirical studies. One study (Reid, Gonzalez, Nordness, Trout, & Epstein, 2004) reported the general overall characteristics of research participants with EBD to be on average 11.22 years old, with an IQ of 94.89, 80% male, with ethnicities of 69% Caucasian, 27% African American, 3% Hispanic, and 1% mixed background.



Mooney, Epstein, Reid, and Nelson (2003) examined 25 studies that revealed the general overall characteristics of students with EBD were between the ages of five and 11 years old 54.5% of the time and students who were 12 years and older were participants 36.4%, with 81.8% male participants. Over half the studies used male participants only. Furthermore, Mooney et al. (2003) reported that only 27.3% of the studies examined and reported student ethnicity, while only 4.5% of the studies reported socioeconomic status of the participants.

Examination of both governmental (Donovan & Cross, 2002) and research synthesis (Mooney et al., 2003; Reid et al., 2004) shows a clear picture of the characteristics of students with EBD. All three sources come to similar findings on the gender of students with EBD, while two sources revealed differences in the ethnicity proportions of students with EBD (Donovan & Cross, 2002; Reid et al., 2004). The research synthesis conducted by Reid et al. (2004) only looked at the population of students with EBD who were targeted by researchers for specific studies while the whole population of students with EBD served under IDEA by Donovan and Cross (2002). Donovan and Cross (2002) found that African American students were the highest population of students with EBD, but Reid et al. (2004) found studies used more than twice as many Caucasian students than African Americans students. These differences could also be due to the fact that approximately 27% of research studies report participant ethnicity or the geographic region of the study.

**Academic setting.** The students' with EBD educational setting varies across several educational settings including general education classrooms, resource rooms, self-

contained classrooms, separate schools, residential treatment centers, home schools, or university/clinics. The 25 studies investigated by Mooney et al. (2004) that looked at empirical studies that provided academic intervention to students with EBD and found the academic setting varied across all of the educational settings. Mooney et al. (2004) synthesis found that students with EBD performed lower on the academic intervention than their comparison groups regardless of educational setting.

**Research on academic outcomes of students with EBD.** Students with EBD tend to struggle academically across content areas (Trout, Nordness, Pierce, & Epstein, 2003). Deficits in academic subjects do not appear to decrease over time, but increase with time (Nelson, Benner, Lane, & Smith, 2004). There is research on students with EBD focuses on self-management skills and on academic performance. Self-management refers to an individual's ability to control his/her behavior to accomplish an objective or goal. Students with EBD have a hard time controlling their behavior, which is why self-management skills are so valuable for these students. Academic achievement is another important area for these students, as well. First, self-management will be explained more in-depth and some of the components of self-management will be described. Then the research finding on self-management and academic achievement of students with EBD will be reported.

Hughes and Lloyd (1993) article titled "*An analysis of self-management*" is a theoretical analysis of self-management skills. They defined self-management to included "self-control, self-regulation, and self-determination (p.406)." They stated, "Self-management refers to instances in which the individual forgoes immediate

reinforcement in favor of highly valued, long-term benefits (p.417).” This review looks at individuals making behavior choices for either long term or short term rewards from their behavioral choice. For example an individual can choose to watch television, put off their paper, but their paper will still weigh on their thoughts. Hughes and Lloyd (1993) concluded by stating, “We would like to know, what are the key ingredients in self-management and how do they interact with other variables to enhance or inhibit effects and if an individual does not want to change a targeted behavior, will self-management be effective in changing the targeted behavior? (p.421)”

Since 1993 researchers have conducted empirical research on self-management and examined academic outcomes of students with EBD. Mooney, Ryan, Uhing, Reid, and Epstein (2005) conducted a review of self-management interventions that targeted academic outcomes for students with EBD. They found 20 empirical studies that used self-management skills to target the academic outcome of students with EBD. Mooney et al. (2005) reported the 20 studies were highly effective self-management interventions (effect size (ES) = 1.80). Most of the studies focused on self-monitoring as their primary self-management skill. The self-management foci included: (a) self-instruction (ES = 2.71); (b) multiple-component interventions (ES = 2.11); (c) self-monitoring (ES = 1.90); (d) strategy instruction (ES = 1.75); (e) self-evaluation (ES = 1.13). Mooney et al. (2005) found that researchers focused self-management intervention on math calculation skills than any other subject. The academic outcomes of self-management interventions were most effective for social studies (ES = 2.66) followed by reading (ES = 2.28), math interventions (ES = 1.97), and writing (ES = 1.13).

Reid, Gonzalez, Nordness, Trout, and Epstein's (2004) examined the academic status of students with EBD. The purpose of Reid et al. (2004) was to quantify the magnitude of difference in academic performance between students with EBD and their nondisabled peer counterparts. Overall, Reid et al. (2004) found that there was an overwhelming difference in the academic performance of students with EBD, compared to students without disabilities. Students with EBD compared to their nondisabled peers had a significantly lower academic performance ( $ES = -.69$ ). More specifically, students with EBD performed the lowest in mathematics ( $ES = -.81$ ) and spelling ( $ES = -.81$ ) compared to students without disabilities. The setting of the intervention also played a significant impact on students with EBD, in that students with EBD at a residential facilities performed the worst ( $ES = -1.19$ ), while students with EBD in a resource rooms performed closer to their peers ( $ES = -.33$ ).

In summary, students with EBD have serious academic needs. These students' academic deficits are not decreasing over time, but falling further behind their peers' performance. Students with EBD academic assessment on states standardized assessments count for No Child Left Behind (NCLB) and the schools' accreditation ratings. There needs to be academic interventions for these students that can be utilized across academic content areas and utilized what we already know from researchers like Reid et al. (2004) and Mooney et al. (2005).

### **Fluency Instruction and Automaticity**

The term fluency is described by being able to do something smoothly and easily at a reasonable pace without hesitation. Further the definition of automaticity is the

ability to do things fluently without thinking about the basic mechanics and focusing one's attention on more complex things. Automaticity occurs with learning, repetition, and over practice. For example, after driving for numerous years most people do not think about how to drive their car and instead think about how to get to where they are going.

In education fluency is taught in reading, writing, and mathematics. In elementary school students are expected to develop more fluency in reading as they progress in school.

**Reading.** Reading fluency is a multifaceted concept that entails a person to accurately read words, at a conversational rate, and with prosody or expression (Hudson, Lane, & Pullen, 2005). First for students to read accurately they have to know how to recognize and decode words, which requires the knowledge of the alphabet with the letter sounds and the ability to blend the letter sounds together to become a words. Then with lots of practice students will learn sight words, which are words student do not have to decode, but can read automatically.

Reading fluency is very important in education. Once students can read fluently they are able to focus their attention on the content of what they are reading instead of focusing their attention on decoding and reading the words. Reading comprehension is a critical skill for all students.

Morgan and Sideridis (2006) conducted a meta-analysis on reading fluency interventions for students with or at risk for learning disabilities. They found thirty studies on reading fluency, but these studies secondary focus also included keyword and

previewing, goal setting with performance feedback, contingent reinforcement, word recognition, and tutoring. Overall these studies included 107 students with or at risk for learning disabilities who were 21 females and 86 males, and represented 94 whites, 8 blacks, 4 Chinese, and 1 Hispanic. Participant grade levels included 74 in kindergarten to fourth graders and 33 fifth to twelve graders who were in 92 general or integrated settings, and 15 in special education settings. The results reported included likely moderators to intervention effectiveness and intervention differences that led to the greatest gains in fluency.

The moderators to intervention effectiveness included gender, age, and placement. First, gender found that girls read more words per minute than boys. Secondly, older students read more than younger students. The intervention for kindergarten through fourth grade revealed that the most effective fluency reading interventions were those that included both goal and reinforcement interventions. Listening and repeated readings might be more effective interventions for grades five through twelve. The third like moderator to intervention effectiveness was education setting, since students in general education settings typically read more than students in special education settings. All the interventions in general education setting was highly effective.

It was reported that goal setting with feedback and reinforcement was more effective than keywords and previewing, listening and repeated readings, reinforcement, tutoring, and word recognition. In addition, they reported that goal setting interventions produced the highest effects in boys while reinforcement interventions produced highest effects in girls.

Another meta-analysis examined repeated reading studies and the impact on fluency and comprehension (Therrien, 2004). This analysis examined eighteen studies on repeated readings between the years of 1977 to June 2001. Therrien (2004) reported a large effect for fluency ( $ES = .83$ ) and comprehension ( $ES = .67$ ). In addition, the repeated reading modeling, corrective feedback, and charting were found as most effective interventions for reading fluency, but were slightly less effective for comprehension.

Chard, Vaughn, and Tyler (2002) also conducted a meta-analysis on interventions for reading fluency for elementary school students with learning disabilities. This study analyzed twenty-four studies published between January of 1975 and December of 2000. A total of 128 participants were reported in 21 of the studies reviewed. The overall effectiveness of reading fluency interventions ranged from having a no effect ( $d = .02$ ) to large effects ( $d = 3.02$ ) with an average large effect for reading fluency interventions ( $d = .68$ ).

An empirical study by Oakes, Mathur, and Lane (2010) focused on reading fluency interventions for students with challenging behavior. This study included both students and teachers. The students were nine second graders, six boys and four girls with a mean age of 7.67 years old. Participants, who all had reading difficulties, included four Caucasians, four Hispanics, and one African American (Oakes, Mathur, & Lane, 2010). In addition to this intervention one student received services for English language learner and four students demonstrated challenging behavioral patterns with demonstrated reading difficulties. Three second grade teachers with a range of teaching

experience from 0-12 years with a mean of 7.67 years delivered the instruction. Experience ranged from one year, eleven years, and twelve years. There was also a reading specialist who had twenty-six years as a speech/reading specialist, first grade teacher, reading specialist and expertise training in reading instruction, program implementation, and data-based decision making. The Oral Reading Fluency (ORF) was implemented with primary reading instruction with Harcourt Trophies series. Instruction was delivered 45 minutes a day, four days a week, with a teacher focus on vocabulary, daily reading practice of stories, and whole class discussions. A second instructional component included whole class phonemic awareness and phonics instruction for an additional 30 minutes a day that was taught four days week. Then another secondary intervention was provided when students' reading rates did not improve after six weeks of small group instruction. The secondary intervention was delivered by the reading specialist for an extra 20 minutes a day, four times a week with an additional 10 minutes spent on read-aloud.

This study found that the nine students (described above) did not progress in the primary instruction and advanced to the secondary instructional phase. The secondary intervention consisted of three phases: (a) reading accuracy instruction and behavioral support, (b) reading accuracy instruction, behavioral support with fluency building, (c) then back to reading accuracy instruction and behavioral support, which was provided in addition to regular class reading instruction. The first phase of the secondary instructional helped improve four of the student's accuracy of correct words per minute, while three students' maintained stable performance and two student's performance



continued to decline. In the second phase of the secondary instruction six students showed positive growth and three students showed slight decreases in performance. Then in the final phase of instruction all students' experienced large gains, except for one student.

**Writing Automaticity.** Writing automaticity is also a multifaceted concept that entails a person to form letters, spell words, use grammar and punctuation at a reasonable rate. Writing just like reading starts with knowing the alphabet, but for writing students need to know the shapes of the letters and replicate the shape of the letter in particular sequences to form words on paper. Then the multiple formed words separated by a space together make sentences with the use of proper grammar and punctuation. Once a student can form multiple sentences together they form a paragraph.

Writing fluency is also important. First, students have to compose text in a certain amount of time. Second, students have to think while writing. Third, writing automaticity like reading fluency allows students to focus on high-order thinking skills that improve their written expression.

Two studies used self-regulated strategy development (SRSD) that report fluency instruction or timed writing. The first study (Mastropieri et al., 2009) reported the second phase of instruction as fluency instruction for persuasive writing. Mastropieri et al. (2009) found that after initial instruction to post-fluency testing three students improved in the number of words written. Also one low performing student improved on number of persuasive essay parts from post-testing to post-fluency testing. Furthermore, all students showed growth from post-testing to post-fluency testing in essay holistic quality,

but several students essay holistic quality was the highest at post-fluency testing. The second study with a primary focus was a study on timed persuasive writing (Mason, Kubina, Valasa, & Cramer, 2010). The Mason et al. (2010) study found that: (a) holistic quality of students' essays level and trend improved and stabilized above baseline level, (b) the number of persuasive essay parts was at a more stable performance level at instruction and post-testing, (c) the number of written words showed that the overall appearance is more stable in instruction and post-testing.

### **Writing Instruction and Self-Regulated Strategy Development**

The self-regulated strategy development (SRSD) was developed by Dr. Steve Graham, Dr. Karen Harris, and colleagues (e.g., Harris & Graham, 1999; Harris, Graham, Mason, & Friedlander, 2008). This strategy was designed to develop for the needs of struggling writing and found that good writers spend time planning, monitor their progress, evaluate their work, revise their work, and regulate through the writing process (Graham & Harris, 2003; Harris & Graham, 1999). Struggling writers either skip parts of the writing process like planning, monitoring their progress, and revising or they report these processes to be difficult for them (Graham, 1990).

The SRSD model of instruction embeds self-regulation strategies within the explicitly taught strategy that facilitates the writing process (Harris & Graham, 1999; Graham & Harris, 2003). The self-regulation components of the SRSD strategy include self-instruction, self-monitoring, self-evaluation, and goal setting. The SRSD strategy has six components: (a) develop background knowledge; (b) discuss it; (c) model it; (d) memorize it; (e) support it; (f) perform it.

There are numerous specific writing strategies using SRSD for specific writing genres including narrative writing, persuasive writing, expository writing, story writing, and report writing. These strategies have been implemented with average writers, poor writers, gifted, and students with and without disabilities in grades second through eighth (e.g. Albertson & Billingsley, 1997; De La Paz, 1999; Saddler, Moran, Graham, & Harris, 2002). Numerous researchers have been studying writing instruction to improve poor writer's ability since the 1980's (e.g. see Graham & Harris, 1989, 2003; Graham, Harris, MacArthur, & Schwartz, 1991; Graham & Perin, 2007; Roger & Graham, 2008). The research has focused on planning, paragraph writing, organizing, mnemonics, peer editing, revising strategies, attributions, and self-regulation.

The first part of the mnemonic strategy for SRSD is POW (P = Pick my idea; O = Organize my notes; W = Write and say more). This POW mnemonic guides students through the writing process and is paired with a genre specific mnemonic strategy that helped student organize their notes with basic components of for the genre, like TREE (T = Topic sentences; R = Reasons three or more; E = Ending; E = Examine) for persuasive writing. The TREE mnemonic guides students through the elements of a persuasive essay. There are also other genre specific mnemonics for story writing, such as WWW, What = 2, How = 2 (W = Who is the main character?; W = When does the story take place?; W = Where does the story take place?; What = 2 = What does the main character do or want to do; How = 2 = How does the main character feel and how do the other characters feel?) (Harris, Graham, Mason, & Friedlander, 2008).

Numerous SRSD studies have been conducted. Researchers synthesized writing literature. Meta-analyses have been conducted on strategy instruction for teaching writing (Graham, 2006), for writing instruction for adolescents (Graham & Perin, 2007), and for written expression for students with LD (Gersten & Baker, 2001). Rogers and Graham (2008) synthesized single subject design studies for writing instruction. Finally two meta-analyses examine the SRSD studies with a focus on students at-risk (Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009) or who had LD (Graham & Harris, 2003).

A recent strategy instruction meta-analysis found 15 studies published in the 1980's, 34 studies published in the 1990's, and 13 studies published from 2000-2005 (Graham, 2006). The increase focus on writing research could be due to No Child Left Behind (NCLB) Act of 2001; or due to the increased number of students with learning disabilities, who increased by 1.5% of the total population (6.5 million) of individuals with disabilities (U.S. Department of Education, 2006). This focused teachers' teaching ability to meet student needs in subject areas such as, English, mathematics, and writing. NCLB includes the performance of students with disabilities on the standardized state assessment. The category of students with disabilities is composed partly by individuals with specific learning disabilities in areas such as reading, writing, English, mathematics, science, and other areas. This category Schools must examine their ability to meet the academic needs of students with disabilities since students with disabilities performance on high stakes tests must be reported.

Graham's (2006) meta-analysis found 39 writing studies. Nineteen were single-subject research and 20 were group comparison studies. A majority of the studies focused on students with LD or students who were struggling writers. The results indicated that studies that used strategy instruction produced the largest effects on the writing performance with an effect size (ES) of 1.32 compared with overall writing studies effectiveness of an ES of 1.15. Further investigation found that SRSD was used 45% of the time in the group comparison studies and 68% of the time in single subject design studies. These SRSD studies that used the group comparison design yielded nearly double the average ES than studies that used a different type of writing strategy instruction. Quality and elements of an essay for all students had moderate to large ES, .80 and .60 correspondingly. However, they also found that the length of essays in strategy instruction was greatly variable depending on the studies population. Length of essay with students with LD yielded the largest effect (.97), while poor and average students obtained a moderate to large effect (.54 and .78), however, there was little impact on good writers (-.002).

Graham and Perin (2007) meta-analysis on writing for adolescents found that 19% of the writing studies used strategy instruction to and 40% of those 20 writing studies used the SRSD strategy. The type of writing instruction that was found to be most effective was strategy instruction and summarization (.82), followed by: (a) peer assistance (.75); (b) setting product goals (.70); (c) word processing (.55); (d) sentence combining (.50); and (e) inquiry, prewriting activities, and process writing approach (.32) were least effective.

Gersten and Baker's (2001) meta-analysis examined studies that used an experimental or quasi-experimental design that taught expressive writing to students with LD. They identified 13 studies that met their inclusion criteria. The students in these studies were in grades one through nine who were taught expository or narrative writing. These writing intervention studies produced moderate to large effects on the written performance of students with LD. The effective interventions had several similar components that Gersten and Baker categorized into three important components, which included explicit instruction in: (a) the writing process, (b) knowledge of critical components and text structure for a given writing genre, (c) guided feedback from either teachers or peers.

Rogers and Graham's (2008) meta-analysis examined the literature on single subject writing interventions. They found 25 out of 75 (one-third) of the studies examined teaching writing strategies for planning and drafting and also used the SRSD writing model. The participants in those 25 studies that used SRSD were students in second to eighth grade who were struggling writers, but also included typical achieving writers. Struggling writers included students with disabilities, mostly with LD, also participants with attention deficit hyperactive disorder (ADHD) (e.g., De La Paz, 2001), EBD (e.g., Mason & Shriner, 2008), and mild mental retardation (e.g., Guzel-Ozmen, 2006). The results yielded positive effects on students' writing skills across multiple measures. The number of persuasive essay parts or elements in written products had mean percentage of nonoverlapping data (PND) of 96% for post-intervention, 90% PND for, and 85% PND for generalization. Findings revealed 10 of the 25 SRSD studies

assessed participants' productivity or length of written products by counting the number of words written, which was effective with 91% PND for post-intervention, and for the seven studies that also assessed maintenance was 100% PND. Although fewer studies assessed participants' written products for holistic quality, the obtained 99% PND for post-intervention and 86% PND for maintenance indicated these were successful.

The final two meta-analyses examined SRSD studies that included students who were at-risk or LD or who had LD (Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009; Graham & Harris, 2003). These studies identified 26 to 49 studies that used SRSD, but synthesis employed different inclusion criteria. Baker et al. (2009) included 21 studies, while Graham and Harris (2003) included 18 studies.

Graham and Harris (2003) analyses found that 72% of the studies involved students with LD in grades four through eight across different writing genres. The even studies that taught story writing included two group comparisons studies with five single subject studies. Graham and Harris (2003) reported a significant positive effect ( $ES = 1.47-3.52$  or 71-100% PND) on written story length, elements, grammar, and quality of writing. There were six studies that taught opinion essays that included two group comparison studies and four single subject studies. Graham and Harris (2003) reported more variable results from a low to high effect ( $ES = .32-5.18$  and 70-100% PND) on written essay length, elements, coherence, and quality.

Overall findings for students in general showed large effects: (a) quality ( $ES = 1.47$  or 97% PND), (b) elements ( $ES = 1.97$  or 92% PND), (c) story grammar ( $ES = 3.52$  or 100% PND), and (d) length ( $ES = 2.07$  or 82% PND). Students with LD showed

larger effects than for students in general: (a) quality (ES = 1.14 or 89% PND), (b) elements (ES = 2.15 or 93% PND), (c) story grammar (ES = 3.52 or 100% PND), and (d) length (ES = 1.86 or 89% PND). Students who were identified as poor writers also obtained larger effects than for students in general: (a) quality (ES = 1.67 or 100% PND), (b) elements (ES = 1.42 or 100% PND), and (c) length (ES = 2.02 or 100% PND). Baker's et al. (2009) meta-analysis found similar positive effects for students with LD or at-risk for LD in group comparison studies. In conclusion, both Baker's et al. (2009) and Graham and Harris (2003) meta-analysis found SRSD instruction to be a very effective method of writing instruction to improve the writing performance of all students including those with, at-risk for, and without disabilities.

**SRSD with students with EBD or at-risk for EBD.** In the SRSD literature there are few articles that use this specific strategy for writing with students who are at-risk for or identified with EBD. There are a number of studies that have used the SRSD strategy for students with or at risk for EBD but they tend to be elementary age students. The writing genre varied between story writing and persuasive essay writing. All of these intervention studies have found positive effects for students who are at-risk for EBD or students with EBD writing abilities.

In one of the studies that used SRSD strategy for persuasive writing for students with or at-risk for EBD was Mason and Shriner (2008). The six students ranged from second to fifth grade and included 4 students who were identified with EBD and 2 students who were at-risk for EBD. There were five males and one female (four Caucasians and two African Americans) who were between the ages of eight to twelve years old. The



intervention was taught one-on-one by two advanced doctoral students that lasted 11 to 13, 30-minute sessions. Mason and Shriner (2008) found significant improvement for the number of persuasive essay parts for all the participants (100% PND) at post intervention and maintenance.

Another SRSD study that used the SRSD strategy for story writing for student with or at-risk for EBD was Lane et al. (2008). In this study there were six students who were at risk for EBD. The participants were five males and one female, between the ages of seven and a half to eight and half years old. Four were Caucasians and two were African Americans. The intervention was taught one-on-one by three trained graduate students that delivered instruction three to four times a week for 30-minutes for 10 to 15 sessions over the course of three to six weeks. Lane et al. (2008) found that all participants increased on the number of persuasive essays elements to be 100% PND at post-intervention and other data for quality and total words was report with positive trends, but PNDs were not provided.

Thus far in the SRSD literature there are even fewer studies that target the middle school student population with EBD. Only four studies have investigated the effects of SRSD on the middle school students with EBD. These four studies have shed a light on the impact of SRSD for students with EBD, but it is just a start. The four studies include Mastropieri et al. (2010), Mastropieri et al. (2009), Mastropieri et al. (in press), and Mason, Kubina, Valasa, and Cramer (2010) and provide some preliminary positive benefits for SRSD implementation for persuasive writing with middle school students with EBD.

One study that used SRSD with middle school students with EBD was Mastropieri et al. (2010). The purpose of this study was to examine the implementation of SRSD instruction for multiple paragraph persuasive writing with a class of ten eighth grade students with severe EBD. A design experiment methodology was used to make modifications to the instruction to address the learning needs of the students to master the material. The study took place at a public day school for middle school students with EBD. The participants included ten students (eight males and two females) that included four African American students, four Caucasian students, one Multiracial student and one Hispanic student. In addition, to the students having EBD a majority of students also had co-morbidity diagnoses, which included ADHD, LD, depression, autism, anxiety, sleep disorders, other health impairments (OHI), and speech and language disabilities. The baseline assessments for these students' essays performance were very low with little variance on the writing measures. Performance on the Woodcock Johnson III Writing Fluency subtest (WJF) was low but variable and ranged from third to eleventh grade abilities with a mean of 6.5.

The SRSD intervention was delivered to the students, first, by the teacher, which then transitioned to two trained graduate research assistance. The intervention was provided four days a week for 55, 30-minute sessions over a four month period. The impact of this SRSD intervention was examined by pretests, posttests, and maintenance tests on persuasive essays and the WJF. Maintenance testing was administered 11.5 weeks after the conclusion of post-testing. The study found there was a moderate effect size (0.67) on WJF from pre- to post-intervention. The results for the persuasive essay

measurements from pretest to posttest and maintenance found increased performance on essay; (a) in the number of words, (b) number of parts, (c) number of paragraphs, (d) number of transition words, and (e) holistic scoring. These persuasive essay measurements had large effects from pretest to posttest (ES ranged from 1.28-2.46), and from pretest to maintenance (ES ranged from .66-2.11). Mastropieri et al. (2010) found on-task behavior of students with EBD was somewhat less than adequate ( $M = 86\%$ ), but missed instructional time was substantial ( $M = 32\%$ ). Missed instructional time was due to time-outs, crisis recourse center use, or absences. Some of the modifications that occurred during the study included adaptations of instruction, materials, and activities. For example initial instruction was delivered by the student's regular teacher, then after five weeks transitions to the graduate student researchers delivering instruction to the whole class. Then at approximately twelve weeks into the intervention instruction the whole group was divided into two groups based on academic skills and behavior to address the individual academic needs of the students.

Mastropieri et al. (2010) reported challenges during the intervention when working with students with serious EBD. Students with anxieties were resistant to teachers and instruction throughout the study. These students were anxious about participating in instruction, working with teachers, and teachers seeing their work. For example, one student was very aware of the proximity of teachers and become alarmed if they got too close. Another student would react to the teachers' proximity by closing his laptop computer to prevent the teacher from seeing his work. While students with moods disorders posed their own instructional challenges. These students found it hard to

concentrate in order to have a productive instructional day. One student was very aware of the effect of her mood and when asked she stated you know it depends on how I feel behavior-wise. While another student found it hard to focus on his work because he had a lot on his mind from previous classes. Some students also expressed frustrations over legibility of handwriting. When their papers appeared messy, students would tear them up and start over because they were not pleased with their work. Mastropieri et al. (2010) study found the SRSD intervention for middle schools students with serious EBD appeared to be promising, but the design study was limited due to the lack of a control or comparison group.

A second study that used SRSD for middle school students with EBD was Mastropieri et al. (2009). This study replicated and extended the Mastropieri et al. (2010) study by utilizing a more rigorous multiple-baseline design and added a fluency component as a second instructional phase. The first phase of instruction taught students how to write a multiple paragraph essays using the SRSD strategy for persuasive writing, POW+TREE. The second phase of instruction taught students to plan and write one paragraph essays in 10-minute referred to fluency. Intervention was delivered by four trained graduate students with students in small groups of three to four students in a classroom. Instruction was delivered four days a week for 30- minutes a session for a total of fifty-five sessions over more than four months. The participants in this study were twelve eighth grade students with serious EBD taught in a public day school for students with serious EBD, eleven males and one female, with six African Americans, four Hispanics, and two Caucasians, with an average age of 13.9 years old. There were

three students who dropped out of the study for various reasons, including expulsion or frequent in-school suspension.

The dependent measures used to investigate this study were the same dependent measures in Mastropieri et al. (2010). The writing prompts were administered three to five times at baseline, five times after the first phase of instruction, five times after the second phase of instruction, and one at maintenance and generalization. Essays were scored for number of words, number of parts, number of paragraphs, number of transition words, and quality. Maintenance and generalization testing was administered three months after the conclusion of the second post-test phase and was administered only for the second phase of fluency instruction. The WJF was administered once at baseline and again after the second phase of instruction along with the strategy interviews. The reliability of scoring for the writing prompts was 98%. The fidelity of instruction was examined and found to be high with a mean of 98%.

The Mastropieri et al. (2009) study included findings for on-task behavior, WJF, and the writing prompts. The results of the on-task behavior were less than desirable during the time students spent in the classroom ( $M = 72\%$ ), which is lower than what the researchers found the year before in Mastropieri et al. (2010). The results of the WJF showed students made statistically significant gains (pre  $M = 75.8$  to post  $M = 84.8$ ),  $t(11) = 3.55$ ,  $p = .005$ , with resulted in a large effect size of .81. The results of the writing prompt measures found all students gained significantly on their ability to write persuasive essays from baseline to both posttest phases of instruction, maintenance and generalization phases. More specifically, results revealed positive impact on students'

essays on quality and parts, which was supported by 100% PNDs. Further essay scores showed large and consistent growth from baseline to post-testing after the first phase of instruction on number of words, number of parts, number of paragraphs, number of transition words, and quality. The means and effect sizes (ES) across all students were large ( $p < .01$ ;  $ES = 1.83$ ). Fluency results yielded improvements for all students from baseline. The fluency data found three students improved on the number of words from the first phase of post-testing to post-fluency testing phase. Furthermore, all students showed growth from the first post-testing phase to post-fluency testing on quality and several students' essay quality was the highest at post-fluency testing. Maintenance and generalization results demonstrated that three out of the four instructional groups maintained a higher performance than their baseline performance on number of parts and quality. Maintenance performance was statistically significant from baseline on all essay measures but generalization was only significant from baseline on number of words and number transition words.

In conclusion the Mastropieri et al. (2009) study had three limitations. First, it is unknown the degree to which those samples of students with EBD, their severity of disability, and the presence of co-morbidity of disability areas affected the study results. Second, since there was no explicit generalization instruction embedded within the strategy. Finally, fluency instruction was implemented following a lengthy multi-paragraph intervention, so it is unknown whether or not students can begin to write more fluently after shorter initial instructional period.

The third SRSD study examining persuasive writing with middle school students with EBD was Mason, Kubina, Valasa, and Cramer (2010). The purpose of this study was to extend the literature on SRSD instruction by adding additional scaffolding and guided practice, and testing the effects for students with EBD for persuasive quick writing (QW), which is the same as fluency in the Mastropieri et al. (2010; 2009) studies, that was used with middle school students with LD and ADHD (Mason, Kubina, & Taft, 2011). The participants were five seventh and eighth grade Caucasian students with EBD including four males and one female. Students ranged in age from 12.1 to 14.5 years old and attended an alternative school. The intervention was taught one-to-one in a hallway outside of the classroom by a trained graduate student. Instructional sessions were 30-minutes for five sessions and three sessions were ten minutes that occurred over a course of two to three weeks. This study started with six students, but one left before the end of the study to go back to his base school.

This study (Mason et al., 2010) like the Mastropieri et al. (2009) study used the similar methodology, a multiple baseline design. To measure the outcome of this study the same dependent measures were used that were in Mastropieri et al. (2009). The intervention used the WJF at pretest and post-testing and writing prompts, with five to nine at baseline, five at posttest, and one at maintenance. Maintenance testing accorded two weeks after the conclusion of post-testing. The writing prompts were scored for quality, parts, and word count. The reliability for the writing prompts was broken down by measures with holistic quality reliability being 82% and parts reliability being 73%,

there was no reliability done for word count since Microsoft office word count was used. Fidelity of the SRSD intervention was reported at 100%.

The results of the Mason et al. (2010) study examine the writing prompts of the students on quality parts, and word count. First the quality of students' essays level and trend improved and was stabilized above baseline level, which is supported by an 84% PND at post-testing, and 60% PND at maintenance. Second, the essay parts showed a more stable performance across the phases with only two students PNDs were above a 0% with a 20% and 100% PND, and maintenance revealed 0% PND. These results showed initially students were inconsistent in the number of parts but became more stable at the high end of the baseline performance at post-testing, and maintenance testing. Finally with word count showed that the overall results appearance was more stable at post-testing, but only one student obtained a positive PND at post-testing with a 40% PND.

In conclusion, Mason et al. (2010) study results are not as promising as Mastropieri and colleagues studies (Mastropieri et al, 2010; 2009). The gains of students from the intervention were not as great in Mason et al. (2010) as in Mastropieri et al. (2009), but this could be from the intervention differences in that Mastropieri et al. (2009) started teaching the students a multiple paragraph essay before going to the one paragraph for fluency. Mason et al. (2010) had a few limitations. The first limitation is the sample used for the study was not ethnically or racially diverse population in that the study contained all Caucasian students, unlike the Mastropieri and colleagues (Mastropieri et al., 2010; 2009). Therefore it is unknown if the results could have been



different for the Mason et al. (2010) study if a more diverse population was included. Second, Mason et al. (2010) taught a student one-to-one, which is very different from a classroom setting (Mastropieri et al., 2010) or small group instruction (Mastropieri et al., 2009). Therefore this intervention and the effects of SRSD for quick writing should be tested in a small group instructional format because the one-on-one instructional format cannot be generalized to a small group instructional format. The results of Mason et al. (2010) could have been different if taught in small groups. Third, the differences between the Mastropieri and colleagues studies (Mastropieri et al, 2010; 2009) and Mason et al. (2010) could be accounted for in the intervention differences in duration with Mason et al. (2010) contained eight sessions (of which 3 were only 10-minutes in length) and the Mastropieri and colleagues studies (Mastropieri et al., 2010; 2009) contained 50-55 sessions in 30-minute sessions.

Mastropieri et al. (in press), which replicated and extended Mastropieri et al. (2009) study by using a inclusive middles school setting for students with EBD and inserting a counter argument component to the SRSD instruction. The participants in this study were 15 middle school male students with EBD. Three of those students were dropped from the study for various reasons. Therefore the final sample was 12 middle school male students with EBD and approximately half had other co-morbid conditions that included eight Caucasian students, three African American students, and one Asian American student that had a mean age of 12.4 years old. The intervention in this study included the same study design and the same two phases of instruction in the same order as Mastropieri et al. (2009). This study was taught in five small groups of two to three

students who were instructed by trained graduate students for a total of 18 sessions that were taught five days a week for 45-minutes per session over the course of three months. Furthermore, the first phase of instruction for multiple paragraph instruction took approximately 15 days of instruction and the second phase of instruction for writing one paragraph essay fluently took approximately three days of instruction.

The dependent measures in this study were the same as in the previous studies (Mason et al., 2010; Mastropieri et al., 2010; 2009). The results of the WJF found students had positive gains (pretest mean of 86.87 and posttest mean of 103.33) that resulted in a statistically significant difference between the two phases according to the Wilcoxon matched-pairs, signed rank tests. The essay prompts were administered to students at baseline, at post-testing after phase one of instruction, at post-testing after phase two of instruction, maintenance, and generalization. Maintenance and generalization testing was conducted for both instructional phases and was administered two weeks after the second post-testing phase concluded. The essay prompts were measured for words, sentences, paragraphs, transition words, essay elements, and quality scoring. At the first post-testing phase compared to baseline students made statistically significant gains on all the essay measures, which were sustained at both maintenance and generalization testing phases. At the second post-testing phase compared to baseline students again made significant gains on all essay measures, except for paragraphs, which was due to fluency instruction taught students to write one paragraph essays. The results from the second post-testing phase were sustained at both maintenance and generalization testing.

## **The Present Study**

The purpose of this study is to replicate and extend the literature in important ways. First, the research seeks to replicate the effects of SRSD with middle school student with EBD for persuasive writing. Second, this study extended the research base by examining the effects of the order of instructional phases for middle school students with EBD, by first teaching students to write a one paragraph essay followed by a second phase to teach student to write a multiple paragraph essay. This will provide additional information about fluency writing. Finally, this study examined the students' self-efficacy beliefs about their persuasive writing abilities and the effects of SRSD for persuasive writing have on the students' self-efficacy beliefs.

## **Research Questions**

- 1) Does SRSD for persuasive writing, using POW+TREE, improve the persuasive writing abilities (as measured by length, content, and quality) of students and fluency abilities (including WJF) after instruction (instructional phase 1) and after instruction in expanding essay length (phase 2), and are they able to maintain and generalize learning?
- 2) Are students with EBD able to accurately conduct all steps of the POW+TREE strategy? Are students with EBD maintaining good time on task behavior? Does learning how to write persuasive essays using SRSD improve students' self-efficacy?
- 3) Do students find the POW+TREE strategies easy to implement, useful and enjoyable?

### 3. METHODS

#### **Research Design**

This study used a multiprobe, multiple baseline design across participants to investigate the effects of SRSD for persuasive writing for one paragraph and multiple paragraph instruction across three small groups (Kennedy, 2005). In addition, to the multiprobe, multiple baseline design this study also employed Horner et al. (2005) quality indicators for single subject research. Therefore, the quality indicators that were followed for this study included description of the participants and settings, dependent variables, independent variables, baseline, experimental control or internal validity, external validity, and social validity (Horner et al., 2005).

The participants in this study were seven middle school students with EBD that attended an inclusive middle school program for students with learning and emotional disabilities. Students were stratified by baseline writing measures performance into three small groups ( $n = 2-3$ ), which resulted in three tiers of replication of instructional procedures across the small groups. The small groups were staggered from the lowest performing group to the highest performing group into the first intervention phase, when the small group achieved a stable baseline conforming to single subject methodology (Kennedy, 2005). This created three staggered replications of the instructional

procedures, since each group received separate instructions by a trained graduate instructor.

At baseline students were administered at least three timed essay prompts, one untimed essay prompt, the Woodcock Johnson III fluency subtest (WJF), a strategy probe, and self-efficacy measure over a period of two to four weeks. The baseline phase consisted of typical remediation instruction for the seventh and eighth grade students, which consisted of basic skills, reading, computer time, and opportunities for both students and teachers to meet on an as needed basis. Remediation instruction did not include explicated English instruction on persuasive essay writing.

During the intervention instruction, students received at least three essay probes. In the first intervention phase, students were taught the SRSD strategy and how to use the strategy fluently. To move past the first intervention phase students had to reach mastery performance. Mastery performance was attained when the student wrote two persuasive essays independently containing 11 persuasive essay parts. After mastery performance was obtained students were administered three posttest timed essay prompts. The second instructional phase to write multiple paragraph essays followed. Again, once students reached mastery performance by writing two persuasive essays with all the required components, they received three posttest instructional untimed essay prompts, the WJF, strategy probe, self-efficacy measure, and a social validity interview was conducted on the students' perceptions of the usefulness of the strategy and their perception of the strategy. Then students were provided a one-day lesson on generalization. Approximately five weeks after the conclusion of the multiple-paragraph post-testing

both maintenance and generalization prompts were administered for both phases of instruction, along with the self-efficacy measure and strategy probe. One student declined to participate in the maintenance and generalization testing.

### **Setting**

The study took place at a middle school located in a metropolitan school district on the east coast of the United States. At this particular middle school, there was a population of approximately 950 students with 15.8% of those students who received special education services. Of those 950 students enrolled in school, approximately 51.4% were male and 48.6% female. The diversity of the school was 32.1% Caucasian, 18.3% African American, 28.4% Hispanic, 14.2% Asian or Pacific Islander, and 7.0% other racial or ethnic groups. Furthermore, 36.2% of the school's population received free or reduced lunch and 13.8% of the school's population were English for speakers of other languages.

### **Participants**

**Students.** Students who participated in this study all met the following criteria: (a) were identified as ED or received ED services, and (b) had difficulty with written expression. These students ranged in emotional and/or behavioral issues from student to student, and included co-morbidity with anxiety, depression, bipolar disorder, attention deficit hyperactivity disorder, and mood disorders.

The school's special education administrator and one of the special education English teachers initially identified twelve students for this study. Only nine of those students volunteered to participate in the study. The nine students who volunteered to

participant in the study included seven males and two females. Furthermore, the participants were five Caucasians, three African Americans, and one Hispanic with an average age of 13.4 years old ( $SD = 0.37$ ). Three students dropped out of the study at various points. One student was dropped from the study after the first day of instruction due to her leaving the school for mental health reasons. Another student was dropped from the study due to being expelled from the school, but he completed half of the study that included the first phase of instruction for fluency instruction and post-fluency testing. A third student was dropped from the study due to his chronic absenteeism from school. In addition, a fourth student transferred to a different school near the end of the study, but finished the multiple paragraph instruction, multiple paragraph post-testing, maintenance and generalization testing at his new school. Table 1 shows demographic data on the sample. Therefore, the final student sample included seven middle school students with an average age of 13.5 years old ( $SD = 0.31$ ) (six males and one female) who were racially and ethnically diverse (four Caucasians, two African Americans, and one Hispanic).

**Trained graduate instructors.** The trained graduate instructors included three individuals from the local university. The school collaborated with the trained graduate instructors throughout the duration of the study. There were three trained graduate instructors that were all advanced graduate students. They all were Caucasian females who had an average of eight years of teaching experience with individuals with disabilities, and were an average age of 37 year old. All instructors had previous

Table 1. Student Demographic Characteristics

Name	Gender	Ethnicity	Age	Grade Level	Special Education Categories	Behavioral Goals	Writing Goal	Test Scores
Kevin	Male	Caucasian	13.86	8	ED; OHI	Social skills	Multi-paragraph essay	TOWL 3-79; WISC IV 98; VCI 110; PRI 108; WMI 8; PSI 75;
Cassandra	Female	African American	13.48	7	ED; SLD	Coping strategies	Sentence structure, punctuation, grammar	WJ BWL 67; SBIS V 93; VR 82; QR 86;
Jacob	Male	Caucasian	13.10	7	AUT; ED; SLD; OHI	Social skills; Coping skills	Multi-paragraph essay	WISC IV 105; VCI 112; PRI 125; WMI 88; PSI 75;
Caleb	Male	Caucasian	13.42	8	ED; OHI;		Multi-paragraph essay w/min. 5 paragraphs	WISC IV 103; VCI 128; PRI 135; WMI 102; PS 115; WJ BWL 103;
Miguel	Male	Hispanic	13.42	7	ED; ESOL; OHI;	Task avoidance	Multi-paragraph essay w/at least 3 paragraphs	WJ BWL 87; SBIS V 98;
Jamal	Male	African American	14.04	7	ED	Social skills, self regulation	Multi-paragraph essay	WISC IV 97; VCI 89; PRI 106; WMI 88; PSI 73;
Garrett	Male	Caucasian	13.47	7	ED; OHI	Behavioral improvement in class; self-awareness and self-advocacy;	Multi-paragraph essay	SBIS V 104; VR 115; QR 112; STM 105;

*Note.* ED = emotional disabilities; SLD = specific learning disabilities; OHI = other health impairments; AUT = autism spectrum disorder; ESOL= English for Speakers of Other Languages; TOWL- 3 = Test of Oral Written Language (Hammill & Larsen, 1996); WISC IV = Wechsler Intelligence Scale for Children-fourth edition (Wechsler, 2003) full scale IQ; VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; WMI = Working Memory Index; PSI = Processing Speed Index; WJ BWL = Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001) Broad Written Language; SBIS V= Stanford-Binet Intelligence Scale, fifth edition (Roid, 2003) full scale IQ; VR = Verbal Reasoning; QR = Quantitative Reasoning; STM = Short-Term Memory;



extensive training in using the SRSD writing model for persuasive writing and with over 60 days of instruction implementing the SRSD strategy with students with EBD. Two of the instructors had two years of experience and the third had three years of experience implementing the SRSD strategy intervention with students with EBD.

## **Materials**

A majority of the materials for this study were based on the SRSD model and were used in previous studies (Mason et al., 2010; Mastropieri et al., in press; Mastropieri et al., 2009), but materials were tailored to meet the demands of the current population and study. Initial materials were based on Mason and Shriner's (2008) study, which taught persuasive writing using the SRSD model to six second through fifth grade students with EBD or at risk for EBD. The SRSD model of instruction includes inexplicit self-regulation instruction through goal setting, self-instruction, and self-monitoring (Graham & Harris, 2003). In this study, persuasive essay writing was taught using the mnemonic POW+TREE in combination with the SRSD model of instruction (Harris, Graham, Mason, & Friedlander, 2008). The mnemonic POW+TREE helps the students remember to (P) pick your topic, (O) organize your notes, (W) write and say more, and (T) topic sentences, (R) reasons three or more with at least one counter reason, (E) explanations for my reasons, and (E) ending and examine.

**Student materials.** The students' materials included individual student folders that were used to store all the materials from the study for an individual student. Students' folders include: (a) a learning contract; (b) a POW+TREE

mnemonic chart (see Appendix A); (c) graphic organizer (see Appendix B); (d) a transition word chart (see Appendix C); (e) a self-statement sheet (see Appendix D) that contained positive statements the student could say to themselves while they work; (f) a daily self-monitoring checklist called a record sheet (see Appendix E); and (g) essay prompts. Completed essays were also stored in their folders.

**Training materials and procedures.** The trained instructors were provided with a binder of materials for the entire study that included lesson plans (see sample lesson plan for fluency instruction in Appendix F and sample lesson plan for multiple paragraph instruction in Appendix G), all instructional materials for the lesson plans, pencils, loose leaf paper, and erasers. To refresh the trained graduate instructors with the SRSD components the SRSD model was reviewed. The SRSD model included the six instructional stages: (a) develop background knowledge, (b) discuss it, (c) model it, (d) memorize it, (e) support it, and (f) perform it. Then the lessons were reviewed entirely with the instructors and instructors were refreshed on the lessons a second time before they taught that lesson. Each of the graduate instructors delivered a lesson while the researcher conducted fidelity. The training of the graduate instructions was approximately five hours over the course of the study. The researcher and trained instructors spoke regularly in person throughout the study about how the lessons were progressing, how the students were progressing, behavioral issues, and questions that came up during the study.

## **Data Sources and Scoring Procedures**

**Woodcock Johnson writing fluency subtest (WJF).** All participants were given the WJF (Woodcock, McGrew, & Mather, 2001b), which was administered according to the directions in the examiner's manual (Woodcock, McGrew, & Mather, 2001b) for the Woodcock Johnson III. The WJF was administered to all the participants at baseline and at post-multiple paragraph testing. This assessment looks at how many simple sentences a student could formulate and write in seven minutes about a picture using the three words given without changing any of the words provided in writing a complete sentence. For this assessment, everyone completes the four examples before beginning. An example was a picture of a ball with the three words "this", "ball", "big" and the student has to write a simple complete sentence that includes the three words without changing them in anyway. Therefore, their answer looked similar to "This ball is big." not "This ball is the biggest."

After the examples were completed, students were given seven minutes to complete as many of the 40 items as possible. Each item of the test items was awarded either one point or zero points. To earn one point the student's response had to use the three words given in a complete sentence without changing any of the words. Zero points were awarded to items that did not use all three words in a complete sentence. For example, if a student changed a word in some way like ring to rang the response would not be considered a complete sentence, since the student changed one of the provided words. The raw scores were calculated by adding up the points earned for a composite score. Then the raw scores were converted to standard

scores by using the software provided with the testing materials. The reliability of this test overall was reported to be 0.88 (Schrack, McGrew, & Woodcock, 2001).

**Essay prompts.** Students completed persuasive essays at multiple points throughout the study. Essay prompts were administered at baseline, at post-fluency instruction, at post-multiple-paragraph instruction, and maintenance and generalization phases. Each student completed at least three persuasive essays at each testing phase, except for maintenance and generalization where students received one maintenance and one generalization essay prompt for both fluency and multiple-paragraph.

For all persuasive essays, students were given a choice of two essay topics to choose from. There were two types of essay prompts administered to the students. The two types of essay prompts administered were timed and untimed essay prompts. For the timed essay prompts, students were given ten minutes to compose their essay. A time timer® was used to monitor the ten minute writing interval, which also allowed student to visually see how much time they had left while the red disk slowly disappeared as time elapsed. Once the teacher set the time timer® to ten minutes students were told they could begin and the clock was placed where all the students could see it. At the end of the ten minutes, students were told to stop writing and put down their pencils. Timed essay prompts were administered at baseline, post-fluency instruction, and maintenance and generalization for fluency instruction. Students received at least three timed essay prompts at each testing phase. For the untimed essay prompts, students were given the entire 40-minute class period to compose their

essay. Untimed essay prompts were administered at baseline, post-multiple-paragraph, and maintenance and generalization for multiple-paragraph instruction. Students received one untimed essay prompt at baseline, three at post-multiple paragraph instruction, and one at maintenance and generalization testing phases.

After each essay was written, regardless of condition, the instructor read through the student's essay to make sure all words were legible. If there were any illegible words, the instructor would write the word legibly next to the illegible word. The researcher then typed all the essays on the computer, without any corrections for spelling or grammar. All students' written essays were measured for: total number of words, number of sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score.

***Total words.*** The total number of words written for each essays was computed using Microsoft Word. Microsoft Word's review function was used to record the total number of words written for each essay.

***Sentences.*** The number of sentences was calculated by tallying the sentences in an essay that contained a subject and a verb with an ending punctuation mark.

***Paragraphs.*** The number of paragraphs was calculated by counting the number of three sentence clusters. A paragraph was considered a group of at least three sentences grouped together.

***Number of transition words.*** The number of transition words was calculated by counting the transition words used in the essay. The scorer read the essays looking for transition words that transitioned from one reason to another reason such as, my

first reason, another reason, my final reason, and so on. Each transition word was awarded one point and was totaled for the number of transition words.

***Number of persuasive essay parts.*** The number of persuasive essay parts was calculated by counting each essay component. This was done by awarding one point for each persuasive essay part in the essay. Each part was awarded one point for each of the following components: (a) topic sentence, (b) each reason, (c) each reason's explanation, (d) each counter reason, (e) each counter reason's explanation, (f) each refute of the counter reason, and (g) the ending.

***Holistic quality score.*** The holistic quality score for each essay was scored according to a holistic quality rubric. The holistic quality rubric scores ranged from 0-10 where a score of 10 was the highest possible score (see Appendix H). A score of 10 was defined by the essay containing a topic sentence, more than three reasons, with two counter reasons, at least three explanations, and an ending with the essay written in a logical order that strengthens the writer's argument. Whereas, a score of zero was defined by the essay containing no persuasive essay parts. For a score between one and five indicated the essay contained a topic, three reasons, and an ending were each part was awarded a point to obtain the essay's holistic quality score, but if the essay did not contain these essential essay components the additional components did not count, like explanations, counter reasons, or refute.

***Strategy probe.*** The strategy probe was administered at baseline, multiple times during the first phase of the intervention, post-multiple paragraph testing, and at maintenance testing. The strategy probe asked the students what the parts are in a

good persuasive essay. The instructor administered the strategy probe verbally one-on-one with the students and the students answered the question verbally while the instructor wrote down the student's response. Students' answers were scored for the components of the strategy that was taught, POW+TREE. Each persuasive essay part identified by the student was awarded a point, but if the essay part had multiple components extra points were awarded. For example, an extra point was awarded for reasons, if the student indicated there should be at least three reasons or more. There was a total of thirteen possible points for: pick your idea, organize your notes, write and say more, topic sentences, reasons three or more, one or more counter reasons, refute, explanations for reasons, ending, examine, and transition words.

**Self-efficacy measure for persuasive writing.** It has been shown that, the self-efficacy assessment has to correspond closely to a related outcome measure (Bandura, 2006; Pajares & Valiante, 2006). The literature was searched to no avail to find a validated self-efficacy measure for persuasive writing. Therefore, a self-efficacy measure was developed for persuasive essay writing specific about the SRSD strategy POW+TREE to meet the needs of this study. The self-efficacy measure was developed using recommendations from Bandura (2006), and Pajares and Valiante (2006). The developed self-efficacy measure for persuasive writing asked student how confident they were at completing specific aspects of the persuasive writing process and executing specific components of a persuasive essay. Students responded after seeing a persuasive writing essay prompt.

The self-efficacy measure was completed by the students in conjunction with an untimed essay prompt that was administered at baseline, post-multiple paragraph, and maintenance testing phases. The instructor read aloud the two writing prompt choices on the self-efficacy survey to the students and they had to indicate which essay prompt they were going to write on after they answered the 13 multiple-choice questions on the self-efficacy survey by circling the essay prompt. The thirteen questions on the self-efficacy survey asked the students how confident they were that they could complete specific aspects of the persuasive essay. These questions were answered using a five point Likert scale. The five point Likert scale ranged from 0% confident to 100% confident (See Appendix I).

**Student interviews.** Student interviews were completed one-on-one with the instructor after all post-testing was completed. Students were interviewed about their knowledge of the strategy, their perception of the instruction, and if they gained any meaningful knowledge from the intervention. The interview questions also asked the student if they had used the strategy in any of their other classes. These interviews were recorded and lasted no more than ten minutes. To see the full list of interview question see Appendix J.

**On-task behavior coding.** Time sampling procedures were used to record students' on and off-task behaviors during the intervention instruction in 30-second intervals for durations of 15-minutes for approximately a third of the session similarly to Mastropieri et al. (2009). Students' on task behavior was operationally defined according to Mastropieri et al. (2009) as "(a) in a designated area of the room, (b)



engaged with appropriate materials, (c) reading/writing to the writing prompt, (d) asking relevant questions(s), and (e) may appear in thought by intermittently and quietly looking away from materials (engaged only with self not with others) (p. 26).”

**Fidelity of treatment.** Fidelity of treatment was measured to make sure the SRSD instruction was delivered as it was intended to be delivered. The SRSD strategy was delivered by experienced instructors that have taught the SRSD strategy to middle school students with EBD. Observers were trained to identify the parts of the lessons and check them off lesson plans as they were completed with space for open ended items and anecdotal notes from the videotaped lessons that were taught. This was completed for approximately 33% of the lessons.

**Reliability.** Reliability was measured to insure students’ essay scores had inter-rater reliability. Reliability gives the study more validity by confirming the operational definitions of: the number of written words, number of sentences, number of paragraphs, number of transition words, number of elements, and holistic quality scores. Reliability was completed for 42.5% of the baseline, post-fluency, post-multiple paragraph, and maintenance and generalization testing phases. The scorers were trained to 90% reliability on scores. Scorers met to assess inter-rater reliability and discussed disagreements until they were resolved. Scorers’ inter-rater reliability was 94%.

## **Procedures**

Before the study began human subject review board approval (HSRB) at the university and the district levels was obtained. Then permissions from the school

administration, students, and parents were obtained followed by the start of baseline in all groups. Baseline phase was staggered over two to four weeks for the three groups of students. Students were grouped into small groups based on writing ability indicated by the initial baseline assessments. Baseline was followed by instructional phases and testing. Once the first group obtained a stable baseline, they began the first instructional phase for the fluency. These procedures were replicated and staggered over time for each of the remaining groups. After students mastered the fluency material by writing two essays independently with all the persuasive essay components they were post-tested. During post-testing students received three timed essay prompts. Immediately, after post-testing students received the second phase of instruction for multiple paragraph essays and when students mastered this material they were post-tested. During post-testing students received three untimed essay prompts along with the WJF, the self-efficacy survey, strategy probe, and student interview. Students were then provided a one-day lesson on generalization. Then maintenance and generalization assessments for each phase were administered. After post-multiple paragraph testing students received both maintenance and generalization testing which included the strategy probe, maintenance and generalization essay prompts for each of the instructional phases, and the self-efficacy survey.

The intervention was implemented during a 40-minute period that the school had reserved for remediation. The entire intervention was completed in a mean of 44.50 ( $SD = 4.72$ ) sessions. The total instructional days ranged from 26 to 39

sessions with students receiving an average of 33 instructional sessions. Instruction was initially delivered three days a week for the months of November and December and then was delivered four days a week for the months of January and February. Students were taught in two phases of instruction. Phase one taught SRSD instruction with fluency instruction to plan and write an essay in ten minutes. Phase two taught students to write multiple paragraph essays. Phase one of instruction resulted in a mean of 26.71 ( $SD = 3.35$ ) sessions of instruction that ranged from 22-32 days while phase two resulted in a mean of 6.67 ( $SD = 1.51$ ) days of instruction that ranging from 4-8 days. An additional one-day lesson on generalization was delivered. Then approximately five weeks after the completion of post-multiple paragraph testing for phase two, students were administered an unannounced maintenance and generalization essay prompts for both fluency instruction and multiple paragraph instruction.

**Baseline.** During baseline students were administered three to five timed essay prompts with each student receiving at least three timed and one untimed essay prompt where students wrote a response to one of the two writing prompts. In addition, students also received the WJF, strategy probe, and self-efficacy survey at baseline.

**First instruction phase: fluency instruction.** This first phase of fluency instruction consisted of five modified SRSD lessons for persuasive writing used for the mnemonic, POW+TREE that was developed by Harris, Graham, Mason, and Friedlander (2008). The SRSD model is composed of six steps that were completed

through the lessons and were employed in this study. The first step was to develop background knowledge and assess students' background knowledge, by asking them what is a persuasive essay. The second step was to discuss it; this is where the instructors lead a discussion on the purpose, benefits, and how and when persuasive writing could be used. During step three instructors modeled planning and writing a persuasive essay while thinking aloud. During the fourth step instructors checked to make sure students had memorized the mnemonic by using the strategy probe. The fifth step was to support it, which allowed the students to work on their own essays while being guided through the writing process by the instructors. During the final sixth step students planned and wrote persuasive essays independently.

***Lesson 1.*** The first lesson was to develop background knowledge which included skills and vocabulary needed to understand what persuasive writing is and the use of the POW+TREE strategy. The instructor introduced the first half of the strategy, POW (stands for pick your ideas, organize your notes, and write and say more), with a discussion of each letter, what it stood for, and why these parts are important to persuasive writing. The instructor checked for student learning throughout the lesson to make sure students were paying attention and could recall POW. Then the class discussed the components of a good persuasive essay followed by the introduction of the second half of the strategy, TREE (topic sentence, reasons three or more, explanations, and ending and examine), with a discussion of each letter, what it stood for, and how this could help them remember to plan and write persuasive essay containing at least eleven parts. Instructors led a group discussion of

examples for each part and read aloud a one paragraph persuasive essay example. The students then helped identify the persuasive essay parts in the example essay and the instructor wrote the persuasive essay parts in the appropriate place in the graphic organizer. Essay examples were used until the students were able to identify persuasive essay parts accurately. In each subsequent lesson, time was spent with the students rehearsing the POW+TREE strategy.

***Lesson 2.*** The focus of this lesson was to discuss it. To review, the instructors went over strategy POW+TREE and its importance. Then the group practiced again another essay example that was read aloud and students identified the essay components while the instructor filled them in on the graphic organizer. Then the group discussed the importance of using the strategy whenever and wherever they may need the strategy. The students were explicitly told that they will be asked about times they have used the strategy in other classes. An additional, essay example was read aloud to the students for them to independently identify the persuasive essay parts in and filled out the graphic organizer. The instructor went over what each student had filled in on their graphic organizers introduced a transition word discussion.

***Lesson 3.*** The focus of this lesson was to model it. The instructors modeled how to use the strategy to write a one paragraph persuasive essay from beginning (planning) to end (written product with the written product examined), and introduced the use of self-statements. Modeling the strategy showed the students: how to complete an essay using POW+TREE, how to plan their essay, then write essays, the

use of transition words, and how to use self-statements. The instructors used think-aloud during the writing process. They used self-statements or self-talk to guide themselves through the process to encourage themselves. At the end of writing the essay the instructor had the students examine the essay to identify all the parts. It was important for the instructor to make sure students were actively watching and participating through the modeling the process.

After the essay was composed and examined, the instructor discussed with the students the self-statements that were used while the instructor went through the writing process. Students were required to write self-statements on their own self-statement sheet that they could use when they planned, wrote, and examined their own essays. Then the instructor introduced goal setting and the self-monitoring sheet. Goal setting was also introduced to the students. Teachers indicated goals for persuasive essays meant including include all eleven parts and the use transition words. The self-monitoring sheet was introduced so students could record their progress, such as the number of parts, and transition words they used in each essay. Each component on the self-monitoring sheet was discussed and students completed a self-monitoring sheet with teachers.

***Lesson 3b.*** The focus of this lesson was to memorize it. During each lesson the instructor reviewed the mnemonic strategy with the students. At this time the instructor took a moment with each student to have them tell them the strategy and what it stood for. This insured that all the students had memorized the strategy. If a

student struggled to recall a specific letter of the strategy and what it meant, the instructor provided additional practice for the student to memorize the strategy.

***Lesson 4.*** The focus of this lesson was to support it, also known as guided practice. The instructor started this lesson guiding the students through planning and writing an essay, but transitioned such that students directed the writing process with instructor support as needed. The class collaborated through the planning stage, but each student filled out their own graphic organizer with their own reasons and explanations. Once students understood what they were doing the teacher let them work at their own pace, but helped students as they transitioned from planning to writing their essay. During this time the instructor provided support to the students as needed and checked on students' writing progress, essay structure, provided feedback and guidance as needed. After students completed their essay, they then examined it, completed their self-monitoring sheet, and saw if they met their goal of having all eleven parts. Then everyone came back together as a group and discussed how the strategy helped them with their essay.

This lesson was repeated to assist students in mastering the strategy and required little or no assistance planning and composing. During this time the instructor provided differentiated instruction to students who needed different amounts of assistance to master the strategy. This allowed students to work at their own pace. Once the student mastered the use of the strategy with the use of all support materials (i.e., self-statements, graphic organizer, and other reminders) was faded to ensure independent strategy use. The graphic organizer used to organize

student's notes was replaced by a blank piece of paper for planning. Students were told they could create and fill in their own graphic organizer. Students continued to record their progress on their self-monitoring sheet for each essay and set their own goals for the next essay.

***Lesson 5.*** The focus of this lesson was independent performance. At this point participants were able to use the strategy to write essays without the use of any materials. The criterion to pass this stage was that participants were able to write two essays independently which contained all eleven persuasive essay parts.

***Lesson 6.*** The focus of this lesson was to model how to plan and write a one paragraph persuasive essay containing all eleven persuasive essay parts in ten minutes. The instructor modeled the writing process similar to lesson 3, but the students had ten minutes to complete the planning and writing. Students were taught how to modify the planning process.

***Lesson 7.*** The focus of this lesson was to demonstrate independent performance. Students were asked to use the strategy to plan and write persuasive essays without using any formal materials, in ten minutes. The students practiced planning and writing a persuasive essay in ten minutes. Then the instructor and students talked about how they did, filled out the self-monitoring sheet, and set a goal for next time. The instructor provided students with additional feedback as needed. The criterion performance was writing two essays independently containing all eleven persuasive essay parts in ten minutes.



**Post-fluency testing.** During post-fluency testing students were administered three timed essay prompts where students wrote a response to one of the two writing prompts. Post-fluency writing prompts were administered on separate sessions.

**Second instruction phase: multiple-paragraph instruction.** This phase of the intervention expanded the students' knowledge of writing a one paragraph persuasive essay they obtained during phase one of the intervention. During this phase of the instruction students were taught to elaborate their one paragraph persuasive essays into multiple paragraph persuasive essays containing all the same components.

***Lesson 1.*** This lesson was focused on modeling how to use the strategy to plan and write a persuasive essay with multiple paragraphs. This lesson replicated lesson 3 that was completed during the first phase of instruction for fluency writing. During this lesson, students learned how to adjust one paragraph essays into a multiple paragraph structure by elaborating on their explanations. They were also taught how to create multiple paragraphs from their reasons. Instructors modeled the similarities and differences between the one paragraph essay and the multiple paragraph essay. The instructors introduced each paragraph and its content. Students were taught that a finished multiple paragraph essay was composed of at least six paragraphs containing the eleven persuasive essay parts.

***Lesson 2.*** This lesson focused on guided practice. The students wrote their own multiple paragraph persuasive essays and were guided through the writing process of organization with multiple paragraphs. The students were quickly able to

see the multiple paragraph essay pattern and the differences between the one and multiple paragraph essay structures.

**Lesson 3.** This lesson was focused on independent performance. At this point students were able to use the strategy to plan and write multiple paragraph essays independently without support materials, but the instructor provided students with feedback as needed. The criterion to pass this stage was that students were able to write two multiple paragraph persuasive essays independently containing all eleven persuasive essay parts.

**Post-multiple paragraph testing.** During post-multiple paragraph testing students were administered three untimed essay prompts in response to one of the two writing prompts. In addition, the WJF, strategy probe, self-efficacy survey, and student interview at post-multiple paragraph testing were administered.

**Generalization lesson.** This lesson focused on a discussion and example of other places the strategy can be used. The instructor lead the group in writing an essay together on a topic that is related to another class like science or social studies, but the students helped by coming up with what will go in the essay as well how to put their ideas into sentences for the essay.

**Maintenance and generalization testing.** During maintenance and generalization testing students were administered a fluency maintenance and generalization timed essay prompts. Students also received multiple paragraph maintenance and generalization untimed essay prompts in response to one of the two

writing prompts. In addition, students also received the strategy probe, and self-efficacy survey at maintenance and generalization testing.

## 4. RESULTS

### **Data Analyses**

The data analyses for the results were twofold. First, overall results were reported by means and standard deviations. Means and standard deviations were reported for on-task behavior, WJF, overall essay results by essay measure, student's overall essay results by essay measure, and self-efficacy measure. Then nonparametric tests were used to analyze the overall results including: (a) the WJF pre- and post-test assessments; (b) overall writing performance by essay measure type across phases; (c) the self-efficacy measure pre-, post-, and maintenance assessments. Second, visual analyses for traditional multiple probe, multiple baseline analysis procedures were used on the individual writing performance of students to examine the data level, stability, variability, and trends (Kennedy, 2005). In addition, to the visual analyses, the percent of nonoverlapping data points (PND) (Scruggs, Mastropieri, & Castro, 1987) were reported between baseline and corresponding posttest phases, and maintenance and generalization phases were calculated to determine the PND outcome.

### **Treatment Fidelity**

Instructors with SRSD instructional expertise and experience working with students with EBD implemented this intervention. Instructor training included teacher practice using the lesson plans and appropriate materials to teach each lesson. Detailed lesson plans with sample scripts were provided to the instructors for the

entire intervention. Over 90% of the lessons were videotaped. One third of the lessons were observed for treatment integrity by two observers. In addition, one observer viewed all remaining videotaped lessons. Observational guides containing open-ended items, checklists, and spaces for anecdotal field notes were completed while lessons were observed. All observed lessons and notes were compared with the lesson plans to identify whether all instructional components were presented and implemented as intended to assess fidelity. The fidelity checks during intervention indicated that the instruction had been delivered with high degree of fidelity ( $M = 97$ ; range 94-100%). Since instructional time was approximately 40 minutes in length, it was noted that teachers were able to cover more content in depth daily than in previous investigations conducted in 30 minute classes (e.g. Mastropieri, et al., 2009), but less than previous investigations conducted in 45 minute classes (e.g. Mastropieri, et al., in press).

### **On-Task Behavior**

Students were observed throughout the investigation for on-task behavior. Observations were recorded across the three instructional groups, 27 days of observation were implemented, and reliability was assessed. Observations were implemented 93.48% of the instructional days, and reliability of observation was assessed 36.05% of the time. Overall, it was noted that the percentage of on-task behavior was moderate for students with ED ( $M = 67.94\%$ ;  $SD = 15.31\%$ ; range = 21.67-96.67%) during the time students spent in this intervention. Reliability of the observations yielded 96.20% ( $SD = 3.98\%$ ) agreement.

## **Standardized Test**

Although students made descriptive gains on standardized scores of the fluency subtest of the WJF, with a pretest mean of 73.57 ( $SD = 8.52$ ) and posttest means of 86.29 ( $SD = 12.38$ ), these differences were not statistically significant according to the Wilcoxon matched-pairs, signed rank test,  $p = .06$ . However, the effect size was 1.22, which is considered high.

## **Writing Performance**

Results for the writing performance measures are organized by testing phase then by overall results and individual student results.

**Baseline.** During baseline students were administered both timed and untimed essays. At baseline, students were provided with pencils and the prompt sheet that contained two essay prompts for the student to choose from with line for the student to write. Students were also provided an additional sheet of loose leaf paper in case they wrote more or if they needed space to plan their essay. For both timed and untimed essays tasks students were asked to write an essay response to one of the following two prompts which were read aloud twice. Directions for the timed fluency essays added that students would only have ten minutes to complete the task and not to start until instructed. A time timer® was used to monitor the ten minute writing interval, which also allowed student to visually see how much time they had left while the red disk slowly disappeared as time elapsed. Once the teacher set the time timer® to ten minutes students were told they could begin and the clock was placed where all the students could see it. At the end of the ten minutes, students

were told to stop, put down their pencils; papers were collected, and read for legibility. If there were any words that were not legible students were asked for clarifications.

During the untimed essays tests, students were directed as they were for the timed essay prompts, but were informed they had as much time as they needed within the period to write their essay. The sequence of days for both timed and untimed essays was a minimum of three timed essays followed by one untimed essay followed by additional timed essays for students in groups two and three. Essays were administered over five to nine days. Students were given at least three, but no more than five untimed and one timed essay prompts during baseline.

***Baseline overall timed essays.*** Overall baseline performance on timed persuasive essays was low and varied, as can be seen in Table 2. On the persuasive essays, students wrote between zero and 111 words with a mean of 45.04 ( $SD = 26.20$ ) words. Students wrote a mean of 2.76 ( $SD = 2.01$ ) sentences, with a mean of 0.40 ( $SD = 0.50$ ) paragraphs. Students included a mean of 0.80 ( $SD = 0.58$ ) transition words. The mean number of persuasive essay parts was 3.32 ( $SD = 1.11$ ) parts. Students' overall mean holistic quality score was 2.40 ( $SD = 0.76$ ).

One student, Jacob, refused to write during baseline, he was provided another opportunity on a subsequent day. Data were presented using that additional opportunity. When that make-up score was included the baseline range was 14 to 111 words with a mean of 46.84 ( $SD = 25.05$ ). Students wrote a mean of 2.76 ( $SD = 2.01$ )

*Table 2.* Fluency results for the overall student essay results of timed essays by phase: baseline, post-fluency, maintenance, and generalization.

	Baseline <sup>b</sup> ( <i>n</i> = 7)		Post-Fluency ( <i>n</i> = 7)		Maintenance ( <i>n</i> = 5)		Generalization ( <i>n</i> = 5)	
	M	SD	M	SD	M	SD	M	SD
Total Words	45.04	26.20	88.29 <sup>a</sup> <i>ES</i> = 2.48	17.42	99.20 <sup>a</sup> <i>ES</i> = 1.58	34.27	88.60 <i>ES</i> = 1.18	37.07
Total Sentences	2.65	2.04	10.38 <sup>a</sup> <i>ES</i> > 3 <sup>c</sup>	1.24	10.40 <sup>a</sup> <i>ES</i> = 1.47	5.27	9.80 <sup>a</sup> <i>ES</i> = 1.20	5.98
Total Paragraphs	0.38	0.50	1.00 <sup>a</sup> <i>ES</i> = 0.62	0.00	2.00 <i>ES</i> = 1.62	1.00	1.80 <i>ES</i> = 1.09	1.30
Total Transition Words	0.77	0.59	5.62 <sup>a</sup> <i>ES</i> > 3 <sup>c</sup>	0.74	6.00 <sup>a</sup> <i>ES</i> > 3 <sup>c</sup>	1.73	5.04 <sup>a</sup> <i>ES</i> = 1.86	2.30
Total Parts	3.19	1.26	10.57 <sup>a</sup> <i>ES</i> > 3 <sup>c</sup>	0.93	8.00 <sup>a</sup> <i>ES</i> = 1.89	2.55	7.60 <sup>a</sup> <i>ES</i> = 1.58	2.79
Total Holistic Score	2.31	0.88	7.90 <sup>a</sup> <i>ES</i> = 2.84	1.97	6.00 <sup>a</sup> <i>ES</i> = 1.35	2.74	6.20 <sup>a</sup> <i>ES</i> = 1.40	2.78

*Note.* Effect sizes were calculated using all relevant post measures standards deviations due to apparent floor effect in baseline measure.

<sup>a</sup> Significantly greater than baseline,  $p < .05$ , according to the Wilcoxon matched-pairs, signed rank test

<sup>b</sup> Means and Standard Deviations without including students' refusal to write scores: total number of words 46.84 (25.05), total sentences 2.76 (2.01), total paragraphs 0.40 (0.50), total transition words 0.80 (0.58), total parts 3.32 (1.11), total holistic score 2.40 (0.76).

<sup>c</sup> When effect sizes are larger than 3 they are reported as greater than 3.



sentences, with a mean of 0.40 ( $SD = 0.50$ ) paragraphs. Students included a mean of 0.80 ( $SD = 0.58$ ) transition words. The mean number of persuasive essay parts was 3.32 ( $SD = 1.11$ ) parts. Students' overall mean holistic quality score was 2.40 ( $SD = 0.76$ ).

***Baseline overall untimed essays.*** Overall, the baseline untimed essay followed similar patterns to the baseline timed essays (see Table 3). On the untimed baseline persuasive essays, students wrote a mean of 35.86 ( $SD = 20.23$ ) words. Students wrote a mean of 1.71 ( $SD = 1.11$ ) sentences, with a mean of 0.29 ( $SD = 0.49$ ) paragraphs. Students included a mean of 0.71 ( $SD = 0.49$ ) transition word. The number of persuasive essay parts included a mean of 3.14 ( $SD = 1.46$ ) parts. Students' overall mean holistic quality score was 2.57 ( $SD = 0.98$ ).

***Individual baseline essays performance.*** Overall baseline performances on timed persuasive essays were low and varied, as can be seen in Figures 1, 2, and 3. Students completed between three and five timed essays at baseline.

***Garrett.*** Garrett was in group one and completed three timed baseline essays before beginning instruction. His performance was consistently low overall and he wrote a mean of 19.67 ( $SD = 3.78$ ) words. His essays contained a mean of 0.33 ( $SD = 0.58$ ) sentences, with no paragraphs. This student included no transition words. The number of persuasive essay parts included a mean of 2.33 ( $SD = 0.58$ ). The student's overall mean holistic quality score was 2.00 ( $SD = 0.00$ ).

*Table 3.* Multiple paragraph results for the overall student essay results of untimed essay by phase: baseline, post-multi-paragraph, maintenance, and generalization.

	Baseline ( <i>n</i> = 7)		Post-Multi-Paragraph ( <i>n</i> = 6)		Maintenance ( <i>n</i> = 5)		Generalization ( <i>n</i> = 5)	
	M	<i>SD</i>	M	<i>SD</i>	M	<i>SD</i>	M	<i>SD</i>
Total Words	35.86	20.23	139.28 <sup>a</sup> <i>ES</i> = 2.62	39.45	137.20 <sup>a</sup> <i>ES</i> = 2.48	40.81	136.00 <sup>a</sup> <i>ES</i> = 2.99	33.49
Total Sentences	1.71	1.11	15.89 <sup>a</sup> <i>ES</i> = 2.90	4.89	16.20 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	4.26	17.20 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	4.32
Total Paragraphs	0.29	0.49	4.22 <sup>a</sup> <i>ES</i> = 1.90	2.07	4.00 <sup>a</sup> <i>ES</i> = 2.63	1.41	4.40 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	0.89
Total Transition Words	0.71	0.49	8.11 <sup>a</sup> <i>ES</i> = 2.84	2.61	8.40 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.52	8.60 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.52
Total Parts	3.14	1.46	10.28 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.96	9.60 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.52	9.40 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.34
Total Holistic Score	2.57	0.98	8.17 <sup>a</sup> <i>ES</i> = 2.92	1.92	8.40 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	0.89	8.20 <sup>a</sup> <i>ES</i> > 3 <sup>b</sup>	1.30

*Note.* Effect sizes were calculated using all relevant post measures standards deviations due to apparent floor effect in baseline measure.

<sup>a</sup> Significantly greater than baseline,  $p < .05$ , according to the Wilcoxon matched-pairs, signed rank test.

<sup>b</sup> When effect sizes are larger than 3 they are reported as greater than 3.

Figure 1. Total number of words by participant by phase: baseline, instruction, post-fluency, and post-multiple paragraph, maintenance, and generalization.

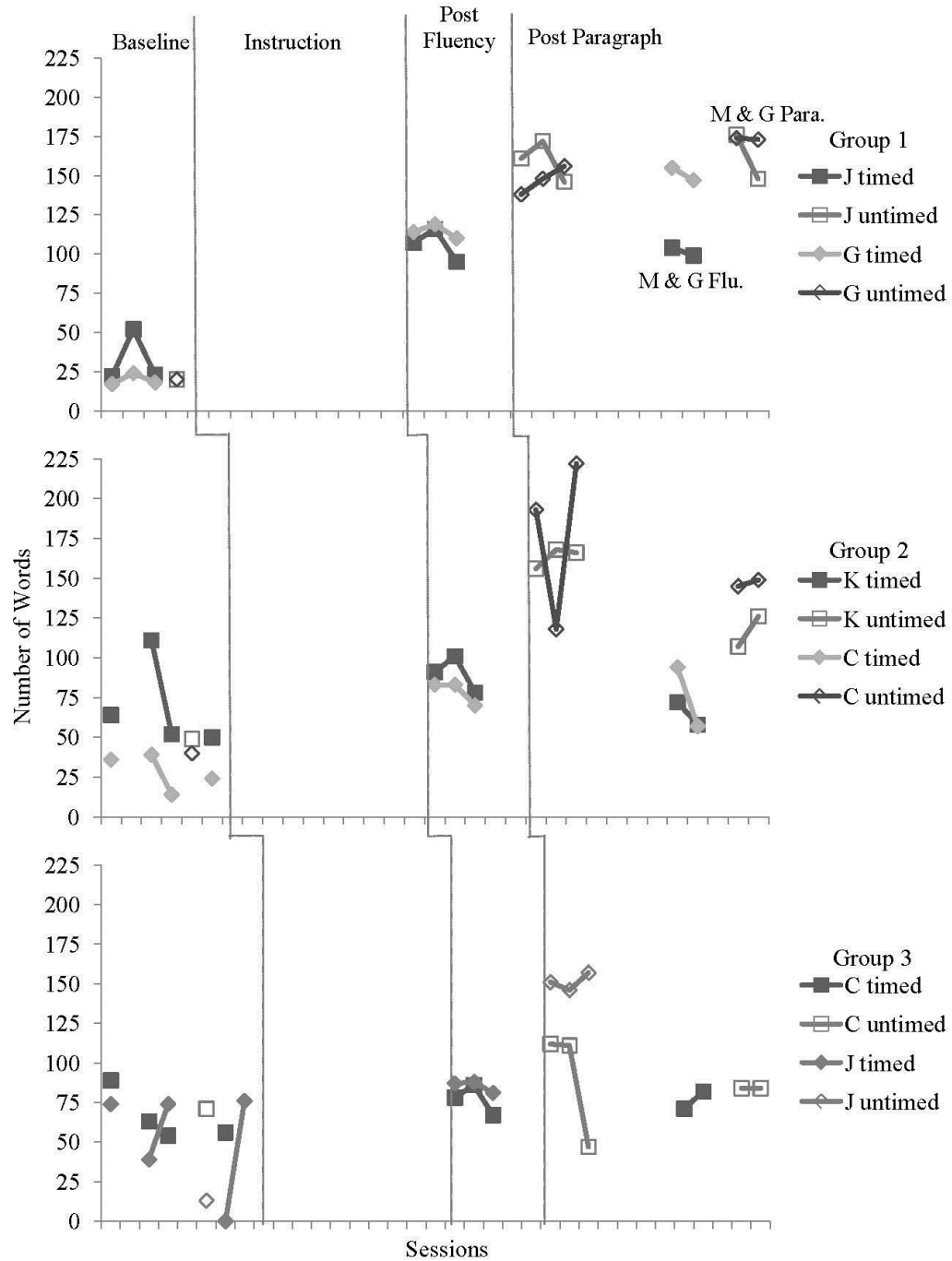


Figure 2. Total number of persuasive essay parts by participant by phase: baseline, training, post-fluency, and post-multiple paragraph, maintenance, and generalization.

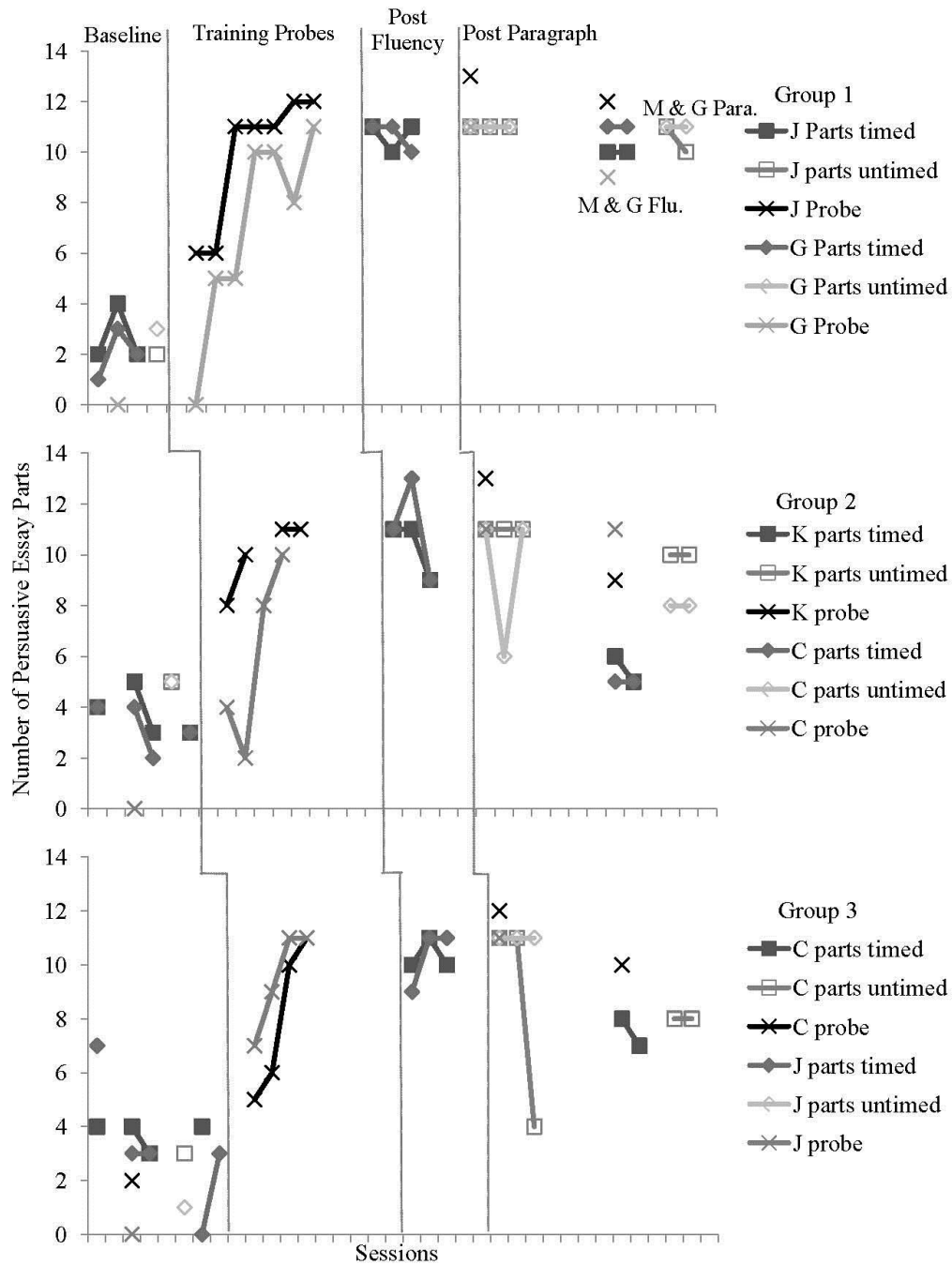
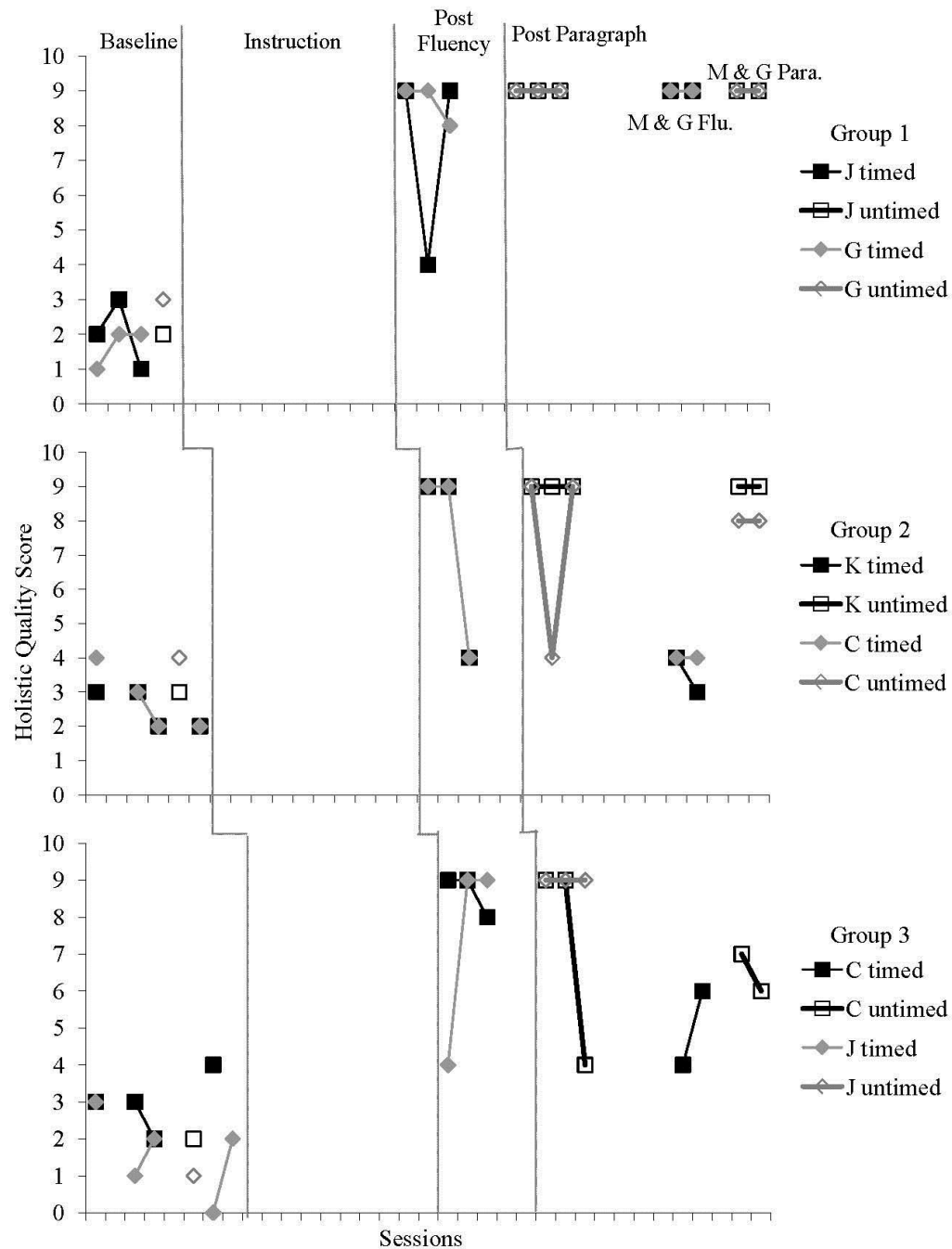


Figure 3. Holistic quality score by participant by phase: baseline, instruction, post-fluency, and post-multiple paragraph, maintenance, and generalization.



On the untimed baseline prompt Garrett wrote 20 words consisting of one long run on sentence without any punctuation. His product contained no sentences, no paragraphs, and no transition words, but he did include three persuasive essay parts and a score of three for holistic quality because he included a topic and two reasons.

*Jamal.* Jamal was in group one and completed three timed baseline essays. His performance was consistent with a slight increase on a few measures on the second essay prompt. He wrote a mean of 32.33 ( $SD = 17.04$ ) words that ranged from 22 to 52. His essays contained a mean of 1.67 ( $SD = 0.58$ ) sentences, with a mean of no paragraphs. This student included a mean of 0.67 ( $SD = 0.58$ ) transition words. The number of persuasive essay parts included a mean of 2.67 ( $SD = 1.16$ ) that ranged between two and four persuasive essay parts. The student's overall mean holistic quality score was 2.00 ( $SD = 1.00$ ) that ranged between one and three.

On the untimed baseline essay, Jamal wrote 20 words. His essays contained one sentence with no paragraphs. This student included one transition word with two persuasive essay parts. The student's overall holistic quality score was two.

*Miguel.* Miguel was in group one and completed three timed baseline essays. His performance was consistent on most measures, but varied a little with the total number of words written. Overall, he wrote a mean of 33.33 ( $SD = 11.50$ ) words that ranged from 22 to 45. His essays contained a mean of 1.67 ( $SD = 0.58$ ) sentences, with a mean of no paragraphs. This student included a mean of 1.00 ( $SD = 0.00$ )

transition word. The number of persuasive essay parts included a mean of 3.33 ( $SD = 0.58$ ). The student's overall mean holistic quality score was 2.67 ( $SD = 0.58$ ).

On the untimed baseline essay, Miguel wrote 38 words. His essays contained two sentences, with no paragraphs, and one transition word. The essay contained one transition word with three persuasive essay parts for an overall holistic quality score of three.

*Cassandra.* Cassandra who was in group two completed four timed baseline essays. Her performance was somewhat consistent. Overall, she wrote a mean of 28.25 ( $SD = 11.50$ ) words that ranged from 14 to 39. Her essays contained a mean of 1.00 ( $SD = 0.82$ ) sentence, with no paragraphs. This student included a mean of 1.00 ( $SD = 0.00$ ) transition words. The number of persuasive essay parts included a mean of 3.25 ( $SD = 0.96$ ) that ranged between two and four persuasive essay parts. Cassandra's overall mean holistic quality score was 2.75 ( $SD = 0.96$ ), which ranged between two and four.

On the untimed baseline essay, Cassandra wrote 40 words. Her essay contained three sentences, with one paragraph. She included one transition word with five persuasive essay parts. Her overall holistic quality score was four.

*Kevin.* Kevin who was in group two completed four timed baseline essays. His performance was inconsistent across the total number of words and sentences, but was more stable across other essay measures. He wrote a mean of 69.25 ( $SD = 28.51$ ) words with a range of 50 to 111 words. His essays contained a mean of 4.00 ( $SD = 1.41$ ) sentences that ranged from three to six sentences, with a mean of 0.75 ( $SD =$

0.50) paragraphs. This student included a mean of 1.25 ( $SD = 0.50$ ) transition words and a mean of 3.75 ( $SD = 0.96$ ) persuasive essay parts that ranged between three and five parts. Kevin's overall mean holistic quality score was consistent, 2.50 ( $SD = 0.58$ ).

On the untimed baseline essay, Kevin wrote 49 words. His essay contained three sentences, with one paragraph. His essay included one transition word with five persuasive essay parts. His overall holistic quality score was three.

*Caleb.* Caleb who was in group three completed four timed baseline essays. His performance varied a little in words, sentences, and transition words, but was more consistent across other essay measures. For example, he wrote a mean of 65.50 ( $SD = 16.14$ ) words, but his scores ranged from 54 to 89. His essays contained a mean of 4.50 ( $SD = 1.00$ ) sentences that ranged from four to six, with a mean of 1.00 ( $SD = 0.00$ ) paragraph. This student included a mean of 1.00 ( $SD = 0.82$ ) transition words, but ranged from zero to two. The number of persuasive essay parts included a mean of 3.50 ( $SD = 0.58$ ). Caleb's overall mean holistic quality score was 2.75 ( $SD = 0.96$ ) that ranged from two to four.

On the untimed baseline essay, Caleb wrote 71 words. His essay contained two sentences, with no paragraphs. His essay included one transition word with three persuasive essay parts. His overall holistic quality score was two.

*Jacob.* Jacob who was in group three completed five timed baseline essays. His performance was inconsistent with a negative trend across a majority of the essay measures. One of the baseline days, he refused to write. Therefore a make-up



opportunity essay was given to him another day. The results presented first included all the results with the make-up essay followed by the results including the essay that was not attempted. Overall, he wrote a mean of 65.75 ( $SD = 17.86$ ) words that ranged from 39 to 76 or with the refused essay a mean of 52.60 ( $SD = 33.22$ ) words that ranged from zero to 76. His essays contained a mean of 5.00 ( $SD = 1.63$ ) sentences that ranged between three and seven sentences and with the refused essay sentences mean of 4.00 ( $SD = 2.65$ ) with a range of zero to seven. He had a mean of 0.75 ( $SD = 0.50$ ) paragraphs that ranged between zero and one, while including the refused essay paragraph mean was .60 ( $SD = 0.55$ ) that also ranged from zero to one. Jacob was more consistent with his use of transition words that ranged zero to one with a mean of ( $SD = 0.58$ ) transition words while with the refused essay mean was .40 ( $SD = 0.55$ ). The number of persuasive essay parts included a mean of 4.00 ( $SD = 2.00$ ) that ranged from three to seven while with the refused essay was a mean of 3.20 ( $SD = 2.49$ ). Jacob's seven part essay was a compare and contrast essay not a persuasive essay, but still included numerous reasons. The student's overall mean holistic quality score was 2.00 ( $SD = 0.82$ ) that ranged from one to three, while with the refused essay had a mean of 1.60 ( $SD = 1.14$ ) that ranged from zero to three. The Figures 1, 2, and 3 include Jacob's results with both the refused essay and the make-up essay results, which theses same results were used to calculate the overall findings (see Table 2 and 3).

On the untimed baseline essay, Jacob wrote 13 words. His essay contained one sentence, with no paragraphs. His essay included one transition word with one persuasive essay parts. Jacob's overall holistic quality score was a one.

**Strategy probes baseline, intervention, and post-intervention performance.** Students were asked to name the parts of a good persuasive essay (pick your idea, organize my notes, write and say more, topic sentence, reasons three or more with at least one counter reason, explanations, ending and examine) to assess their knowledge of the components for a good persuasive essay. Students were asked this question once during baseline, approximately every three day or once a week during the intervention, once during post-testing, and once again at maintenance testing. The instructor administered the strategy probe with each student individually. The instructor would ask the student to name the parts that are in a persuasive essay. Then the student would verbally respond to the question and the instructor would write down the student's response. This was done to assess what students knew about essay components at the beginning and throughout the intervention.

During baseline, all students' knowledge was limited to an average of less than one part. The baseline scores ranged from 0 to 2 with a mean of 0.29 ( $SD = 0.76$ ). The data are displayed graphically in Figures 2. As illustrated in the figures, during intervention, all students increasingly gained in knowledge of the components to a good persuasive essay. Students recalled an additional component not part of the mnemonic after the third probe (i.e., transition words, refute) but correctly recalled all persuasive essay components by the fourth probe. The students' knowledge of the

components of a persuasive essay maintained through post-testing. The post-test scores mean was 12.00 parts recalled ( $SD = 1.00$ ). By maintenance-testing students' performance decreased slightly. Students overall recalled a mean of 10.20 parts ( $SD = 1.30$ ).

**Post-intervention fluency.** Students were administered three post-fluency testing essays. Essays were administered using identical procedures as in baseline giving ten minutes for students to plan and write and were administered over three sessions.

**Overall.** Students' post-intervention fluency essay scores indicate large growth across all essays measures. For the number of words at post-fluency mean of 88.29 with a range from 58 to 119 from a baseline mean of 45.04 that ranged from zero to 111. Then for the number of sentences at post-fluency mean was 10.38 with a range of eight to 13 where baseline mean was zero to seven that ranged from zero to seven. In the number of paragraphs, post-fluency mean was 1.00 without any variability from a baseline mean of 0.38 that ranged from zero to one. The numbers of paragraphs at post-fluency testing showed a small level change with no variability were all students at post-fluency testing wrote one paragraph essay, where at baseline testing students wrote between zero and one paragraph. However, students were taught during instruction to write a one paragraph response, so their performance indicated a few students master this criterion at baseline while others were able to master what they were taught.

Furthermore, there was also large growth across the number of transition words, persuasive essay parts, and holistic quality. Post-fluency transition words mean was 5.62 with a range of four to seven; baseline mean was 0.77 that ranged of zero to two. The transition words variability was slightly more than baseline, but post-fluency testing showed a large level change with students increasing to a mean of 5.62 ( $SD = 0.74$ ) from a mean of 0.77 ( $SD = 0.59$ ) baseline. However, there was a small level change with slightly more variability in the number of transition words. The number of persuasive essay parts showed growth at post-fluency with a mean of 10.57 with a range of nine to 13 from a baseline mean of 3.19 that ranged from zero to seven. Lastly, the holistic quality scores at post-fluency had a mean of 7.90 with a range of four to nine from baseline scores from a baseline mean of 2.31 that ranged from zero to four, as indicated by large level changes with less variability.

The mean scores across all students essay measures for post-fluency testing were large and statistically significant from baseline scores, as indicated in Table 2 (all  $p$ 's < .05, according to Wilcoxon matched-pairs, signed rank tests, from baseline to post-intervention fluency testing on all essay prompt measures). These positive results are supported by the 100% PNDs for number of sentences, number of transition words, and persuasive essay parts, while 95.24% PND for holistic quality score, 71.43% PND for total number of words, and 57.14% PND for number of paragraphs at post-fluency performance on essay measures compared to baseline measures for all students (see Table 4).

*Table 4.* Percentage of Nonoverlapping data by participant by testing phase: post-fluency, post-multiple paragraph, fluency maintenance, fluency generalization, multiple paragraph maintenance, and multiple paragraph generalization compared to baseline.

Student	Testing Phase	Essay Measures					
		Words	Sentences	Paragraphs	Transition words	Essay Parts	Holistic Quality
Garrett	Post-Fluency	100%	100%	100%	100%	100%	100%
	Post-Paragraph	100%	100%	100%	100%	100%	100%
	Main. Fluency	100%	100%	100%	100%	100%	100%
	Gen. Fluency	100%	100%	100%	100%	100%	100%
	Main. Paragraph	100%	100%	100%	100%	100%	100%
	Gen. Paragraph	100%	100%	100%	100%	100%	100%
Jamal	Post-Fluency	100%	100%	100%	100%	100%	100%
	Post-Paragraph	100%	100%	100%	100%	100%	100%
	Main. Fluency	100%	100%	100%	100%	100%	100%
	Gen. Fluency	100%	100%	100%	100%	100%	100%
	Main. Paragraph	100%	100%	100%	100%	100%	100%
	Gen. Paragraph	100%	100%	100%	100%	100%	100%
Miguel <sup>a</sup>	Post-Fluency	100%	100%	100%	100%	100%	100%

## Cassandra

Post-Fluency	100%	100%	100%	100%	100%	66.7%
Post-Paragraph	100%	100%	100%	100%	100%	66.7%
Main. Fluency	100%	100%	100%	100%	100%	100%
Gen. Fluency	100%	100%	100%	100%	100%	100 %
Main. Paragraph	100%	100%	100%	100%	100%	100%
Gen. Paragraph	100%	100%	100%	100%	100%	100%

## Kevin

Post-Fluency	0%	100%	0% <sup>c</sup>	100%	100%	100%
Post-Paragraph	100%	100%	100%	100%	100%	100%
Main. Fluency	0%	0%	0% <sup>c</sup>	100%	100%	100%
Gen. Fluency	0%	0%	0% <sup>c</sup>	100%	0%	0%
Main. Paragraph	100%	100%	100%	100%	100%	100%
Gen. Paragraph	100%	100%	100%	100%	100%	100%

## Caleb

Post-Fluency	0%	100%	0% <sup>c</sup>	100%	100%	100%
Post-Paragraph	66.7%	100%	100%	100%	100	100%
Main. Fluency	0%	100%	0% <sup>c</sup>	100%	100%	0%
Gen. Fluency	0%	100%	0% <sup>c</sup>	100%	100%	100%
Main. Paragraph	100%	100%	100%	100%	100%	100%

Jacob <sup>b</sup>	Gen. Paragraph	100%	100%	100%	100%	100%	100%
	Post-Fluency	100%	100%	0% <sup>c</sup>	100%	100%	100%
Overall	Post-Paragraph	100%	100%	100%	100%	100%	100%
	Post-Fluency	71.43%	100%	57.14% <sup>c</sup>	100%	100%	95.24%
	Post-Paragraph	94.45%	100%	100%	100%	100%	94.45%
	Main. Fluency	60%	80%	60% <sup>c</sup>	100%	100%	80%
	Gen. Fluency	60%	80%	60% <sup>c</sup>	100%	80%	80%
	Main. Paragraph	100%	100%	100%	100%	100%	100%
	Gen. Paragraph	100%	100%	100%	100%	100%	100%

Note: <sup>a</sup> Miguel dropped out of the study after post-fluency testing and was not present for any further testing passed that point

<sup>b</sup> Jacob refused to participate in both maintenance and generalization testing

<sup>c</sup> Student could write one paragraph at baseline therefore at post-fluency growth was limited due to the student already meeting the criteria for one paragraph

***Individual post-fluency instruction essay performance.*** Overall, post-fluency performances on timed persuasive essays were higher and varied from baseline performance, as can be seen in Figures 1, 2, and 3. Students completed between three and five timed essays at baseline and three essays at post-fluency testing.

*Garrett.* Garrett completed three timed post-test fluency essays. A visual analysis of Garrett's performance at post-fluency testing compared to baseline demonstrated a large level change with less or equal variability across all essay measures. He had 100% PNDs from post-fluency measures to baseline measures on all essay measures. His performance was consistently higher at his post-fluency essays as demonstrated by a post-fluency mean of 114.33 ( $SD = 4.51$ ) words that ranged from 110 to 119 where his baseline mean was 19.67 that ranged from 17 to 24. The number of written sentences also grew at post-fluency with a mean of 10.00 ( $SD = 1.00$ ) sentences that ranged from nine to 11 where at baseline the mean was 0.33 that ranged between zero and one. Garrett also improved at post-fluency on the number of paragraph with a mean of one paragraph with no variability where at baseline he had a mean of zero without any variability.

Garrett's growth is further illustrated at post-fluency in the number of transition words that had a mean of 6.00 ( $SD = 0.00$ ). His mean at baseline was zero without any variability. He also had a consistent large step change that was illustrated in a level change with equal variability in the number of persuasive essay parts with a mean of 10.67 ( $SD = 0.58$ ) where the baseline mean was 2.33 ( $SD =$



0.57). Garrett also has a large step change with a level change with increased variability in the holistic quality score mean of 8.67 ( $SD = 0.58$ ) where baseline mean was 2.00 without any variability.

*Jamal.* Jamal completed three timed post-test fluency essays. A visual analysis of Jamal's performance at post-fluency testing compared to baseline demonstrated large step changes in level on all essay measures, with similar or a decrease in the variability, except for holistic quality. The holistic quality score at baseline was low with a mean of 2.00 with minimal variability that ranged from one to three. At post-fluency testing his quality score exhibited a large step change in level with a mean of 7.33, but demonstrated more variability with a range of four to nine. He had 100% PNDs from post-fluency measures to baseline measures on all essay measures.

The large level changes are illustrated in four essay measures: number of words, number of sentences, number of paragraphs, number of transition words, and the number of persuasive essay parts with a moderate (medium) level change in holistic quality score. The number of words are illustrated in a post-fluency mean of 106.00 ( $SD = 10.54$ ) words that ranged from 95 to 116 while at baseline he wrote a mean of 32.33 with a range of 22 to 52. Jamal also improved in number of sentences with mean of 10.67 ( $SD = 0.58$ ) at post-fluency testing that ranged from 10 and 11 while at baseline his mean was 1.67 with a range of one to two. He also had a growth in number of paragraphs with a mean of one paragraph without any variability at post-fluency testing while his baseline mean was zero without any variability. Jamal

also made positive gains in number of transition words where he had a mean of 5.33 ( $SD = 0.58$ ) at post-fluency testing that ranged from five and six where at baseline mean was 0.67 with a range of zero to one. Jamal also had a large step change in level in the number of persuasive essay parts with a mean of 10.67 ( $SD = 0.58$ ) while at baseline his mean was 2.67 that ranged between two and four. His holistic quality score showed a medium level change with greater variability at post-fluency testing with a mean 7.33 ( $SD = 2.89$ ) that ranged from four to nine while at baseline mean was 2.00 that ranged between one and three.

*Miguel.* Miguel completed three timed post-test fluency essays. A visual analysis of Miguel's performance at post-fluency testing compared to baseline demonstrated large step change in level changes across all essay measures with similar to less variability, except for transition words. Since he wrote one transition word at baseline with no variability, but increased at post-fluency testing by writing four to six transition words that has greater variability. He had 100% PNDs from post-fluency measures to baseline measures on all essay measures.

He wrote more words with less variability at post-fluency testing with a mean of 66.67 ( $SD = 8.51$ ) words that ranged from 58 to 75, were at baseline he wrote a mean of 33.33 word that ranged from 22 to 45. He also wrote more transition words with a mean of 5.33 ( $SD = 1.16$ ) compared to a mean of one at baseline. Miguel's essays were composed of a mean of 10.67 ( $SD = 0.58$ ) sentences that ranged between 10 and 11 in one paragraph. At baseline he wrote a mean of 1.67 sentences in zero paragraphs. Miguel wrote a mean of 10.67 ( $SD = 0.58$ ) persuasive essay parts were at

baseline he was only writing a mean of 3.33 persuasive essay parts. His overall mean holistic score increased to 8.67 ( $SD = 0.58$ ) from a mean of 2.67 at baseline. After post-fluency testing was completed Miguel withdrew from the study and subsequently did not complete any other testing phase after this point.

*Cassandra.* Cassandra completed three timed post-test fluency essays. A visual analysis of Cassandra's performance at post-fluency testing compared to baseline demonstrated a large step change in level on all essay measures, but the data demonstrated more variability. The variability at post-fluency testing is evident due to her low baseline performance that was consistently low. However, she had 100% PNDs from post-fluency measures to baseline measures on essay measures total number of words, number of sentences, number of paragraphs, number of transition words, and the number of persuasive essay parts. In addition, she had a 66.7 % PND for holistic quality.

Cassandra at post-fluency testing wrote a mean of 78.67 ( $SD = 7.51$ ) words that ranged from 70 to 83 while at baseline the mean 28.25 that ranged from 14 to 39. Her post-fluency essays had a mean of 11.00 ( $SD = 2.00$ ) sentences in one paragraph where at baseline she wrote a mean of one sentence that ranged from zero to two in zero paragraphs without any variability. At post-fluency Cassandra wrote a mean of 6.00 ( $SD = 1.00$ ) transition words where at baseline her mean was one transition word without any variability. Her number of post-fluency persuasive essay parts included a mean of 11.00 ( $SD = 2.00$ ) that ranged from nine to 13 that resulted in a mean holistic quality score of 7.33 ( $SD = 2.89$ ) that ranged from four to nine. Were at baseline she

wrote a mean of 3.25 persuasive essay parts with a mean holistic quality score of 2.75 with both measures that ranged from two to four.

*Kevin.* Kevin completed three timed post-test fluency essays. A visual analysis of Kevin's performance at post-fluency testing compared to baseline demonstrated a large step change in level with similar variability on the number of transition words. He wrote a mean of 5.33 ( $SD = 0.58$ ) transition words at post-fluency. He wrote a mean of 1.25 ( $SD = 0.50$ ) at baseline. There is a medium level change on three essays measures (number of sentences, number of persuasive essay parts, and holistic quality) with similar variability for the number of sentences and persuasive essay parts, but more variability for the holistic quality score than at baseline. At post-fluency he wrote a mean of 10 sentences with no variability while at baseline his essays contained a mean of four sentences that ranged from three to six. Kevin's post-fluency number of persuasive essay parts was 10.33 that ranged from nine to 11 that increased from baseline where he had a mean of 3.75 that ranged from three to five. Lastly, his post-fluency holistic quality score mean was 7.33 that ranged from four to nine from a baseline mean of 2.50 that ranged from two to three. There was a small level change for total number of words and number of paragraphs with less variability than baseline performance.

His small growth at post-fluency number of word mean was 90 with a range of 78 to 101 where at baseline his mean was 69.25 that ranged from 50 to 111. He also had small growth in number of paragraphs at post-fluency with a mean of one without any variability from a mean of 0.75 that ranged from zero to one. During

fluency instruction, Kevin was taught to write his persuasive essay responses in one paragraph, which he was inconsistent in doing at baseline and even though his number of words did not drastically change, his persuasive essay parts and holistic quality score grew. He had 100% PNDs from post-fluency measures to baseline measures on four essay measures: number of sentences, number of transition words, number of persuasive essay parts, and holistic quality score. He had 0% PNDs for number of words and number of paragraphs.

*Caleb.* Caleb completed three timed post-test fluency essays. A visual analysis of Caleb's performance at post-fluency testing compared to baseline demonstrated a large step change in level with similar variability on three measures. First, he had a large change in the number of transition words at post-fluency with a mean of six without any variability where at baseline he had a mean of one that ranged from zero to two transition words. Secondly, Caleb also had a large change in the number of persuasive essay parts at post-fluency with a mean of 10.33 that ranged from 10 to 11 from a baseline mean of 3.50 that ranged from three to four persuasive essay parts. His final large step change in level was for holistic quality score at post-fluency with a mean of 8.67 with a range of eight to nine from a baseline mean of 2.75 with a score that ranged from two to four.

There was a small step change in level with slightly more variability in the number of sentences at post-fluency with a mean of ten that ranged from eight to 13 from a baseline mean of 4.50 that ranged from four to six sentences. There was also a small step change in level with less variability in the number of words written at post-

fluency testing with a mean of 77 that ranged from 67 to 89, which is up from a baseline mean of 65.50 that ranged from 54 to 89 words written. Lastly, there was no change in Caleb's performance in the number of paragraphs written from post-fluency testing to baseline where he wrote one paragraph without any variability on both. Caleb's paragraphs performance meet criterion at baseline for post-fluency testing, where he was taught to write one paragraph response during fluency instruction. He had 100% PNDs from post-fluency measures to baseline measures on four of the essay measures (sentences, transition words, persuasive essay parts, and holistic quality score), and 0% PND for words and paragraphs.

*Jacob.* Jacob completed three timed post-test fluency essays. A visual analysis of Jacob's performance at post-fluency testing compared to baseline demonstrated small step changes in level on five essay measures (number of words, number of sentences, number of paragraphs, number of persuasive essay parts, and holistic quality) with variability being less than or equal to baseline. Jacob's small level change at post-fluency for number of words had a mean of 85.33 with a range of 81 to 88 where at baseline his mean was 52.60 with a range of zero to 76 words. He also had a small level change in number of sentences where post-fluency mean was 10.33 with a range of nine to 11 while his baseline mean was four with a range of zero to seven. Another, small level change was in the number of paragraphs were post-fluency performance mean was one without any variability where baseline mean was .60 that ranged from zero to one.

Jacob's number of persuasive essay parts also had a small level change at post-fluency with a mean of 10.33 that ranged from nine to 11 from a baseline mean of 3.20 that ranged zero to seven. Lastly, his holistic quality score had a small level change with a slight increase in variability at post-fluency mean of 7.33 that ranged from four to nine where his baseline mean was 1.60 that ranged from zero to three. During one of the post-fluency essays Jacob ran out of time to complete his essay, which impacted him most in the holistic quality score since he had not completed the ending which is a critical component in which he scored a four. There was a medium step change in level on the number of transition words with more variability than baseline, which is illustrated with a post-fluency mean of 5.33 ( $SD = 1.16$ ) while his baseline mean was 0.40 ( $SD = 0.55$ ). The lack of variability in the transition words at baseline was due to Jacob writing either zero or one transition word while at post-fluency testing he wrote between four and six transition words. He had 100% PND from post-fluency measures to baseline measures on a majority of the essay measures (words, sentences, transition words, persuasive essay parts, and holistic quality score), except for 0% PND for paragraphs.

**Post-intervention multiple paragraph.** Post-multiple paragraph testing, students were administered three essays. Essays were administered using identical procedures as in baseline giving students the forty-minute period to plan and write their essay response and prompts were administered over three sessions.

**Overall.** Students' post-intervention multiple paragraph essay scores indicate large growth across all essay prompt measures (number of words, number of

sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score) from baseline scores, as indicated by large step changes in level with greater variability. Students' overall number of words at post-multiple paragraph testing had a mean of 139.28 ( $SD = 39.45$ ) where at the untimed baseline overall students' mean was 35.86 ( $SD = 20.23$ ). They also wrote more sentences at post-multiple paragraph testing with a mean of 15.89 ( $SD = 4.89$ ) where at baseline they only wrote a mean of 1.71 ( $SD = 1.11$ ) sentences showing the large step change in level. Students' grew in the number of paragraphs they wrote. At baseline they wrote 0.29 ( $SD = 0.49$ ) while at post-multiple paragraph they wrote a mean of 4.22 ( $SD = 2.07$ ) paragraphs. The overall number of transition words show a large step change in level with baseline mean of 0.71 ( $SD = 0.49$ ) to a post-multiple paragraph mean of 8.11 ( $SD = 2.61$ ). Students' essays also showed a large step change in level in the number of persuasive essay parts in their essay from a baseline mean of 3.14 ( $SD = 1.46$ ) to a post-multiple paragraph mean of 10.28 ( $SD = 1.96$ ). Lastly, the students' overall holistic quality score had a large step change in level. Their baseline mean score was 2.57 ( $SD = 0.98$ ) to a post-multiple paragraph mean score of 8.17 ( $SD = 1.92$ ).

The mean scores across all students essay measures for post-intervention multiple paragraph essays were large and statistically significant from baseline scores, as indicated in Table 3 (all  $p$ 's < .05, according to Wilcoxon matched-pairs, signed rank tests, from baseline to post-intervention multiple paragraph essay on all essay prompt measures). These results are supported by the 100% PND for number



of sentences, number of paragraphs, number of transition words, and number of persuasive essay parts at post-multiple paragraph performance compared to overall baseline measures for all students, while there was a 94.45% PND for number of words and holistic quality score (see Table 4).

***Individual students post multiple paragraph instruction essay performance.***

Overall, post-multiple paragraph performances on untimed persuasive essays were higher and varied from baseline performance, as can be seen in Figures 1, 2, and 3. Students completed one untimed essay at baseline and three essays at post-multiple paragraph testing.

*Garrett.* Garrett completed three untimed post-test multiple paragraph essays. A visual analysis of Garrett's performance at post-multiple paragraph testing compared to baseline demonstrated large step change in level on across all essay measures with greater variability than baseline due to there being no variability with one untimed essay prompt at baseline. One of Garrett's large step change in level are illustrated by his post-multiple paragraph number of words that had a mean of 147.33 ( $SD = 9.02$ ) words from a baseline of 20 words. His large growth is further shown in his number of sentences and paragraphs where at post-multiple paragraph he had a mean of 17.67 ( $SD = 2.08$ ) sentences in a mean of 4.33 ( $SD = 1.53$ ) paragraphs from baseline of zero for both sentences (he did not use any punctuation) and paragraphs. He also had a large level change at post-multiple paragraph testing in the number of transition words with a mean of 10.00 ( $SD = 1.73$ ) transition words from zero at baseline. Another large step change in level was in the number of persuasive essay

parts where Garrett had a mean of 11.00 without any variability at post-multiple paragraph testing where at baseline his persuasive essay was three. Lastly, his essay overall holistic quality score improved with a large level change with a mean of 9.00 ( $SD = 0.00$ ) at post-multiple paragraph testing from a three at baseline. Garrett's large step changes in level were supported by 100% PNDs from post-multiple paragraph measures to baseline measures on number of words, number of sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score.

*Jamal.* Jamal completed three untimed post-multiple paragraph essays. A visual analysis of Jamal's performance at post-multiple paragraph testing compared to baseline demonstrated clear step changes in level on all essay measures, except there was a small level change in holistic quality score with variability being greater at post-multiple paragraph testing than baseline. The lack of variability at baseline is due to there being one untimed essay prompt at baseline.

Jamal's large level changes are illustrate in the number of words he wrote where at post-multiple paragraph testing he wrote a mean of 159.67 ( $SD = 13.05$ ) words where at baseline he wrote a 20 words. He also had a large level change in the number of transition word with a post-multiple paragraph mean of 9.00 ( $SD = 0.00$ ) transition words with no variability where at baseline he wrote one transition word. There was also a large level change in the structure of Jamal's essays at post-multiple paragraph testing with a mean of 18.67 ( $SD = 0.58$ ) sentences in a mean of six paragraphs where at baseline he had one sentence in no paragraphs. Furthermore,

there was a large growth through a large level change in the content of his essays at post- multiple paragraph testing with a mean of 11.00 ( $SD = 0.00$ ) persuasive essay parts that resulted in a mean holistic quality score of 9.00 ( $SD = 0.00$ ) where at baseline he had two persuasive essay parts for a holistic quality score of two. These results were supported by Jamal's 100% PNDs from post-multiple paragraph measures to baseline measures on all essay measures.

*Cassandra.* Cassandra completed three untimed post-multiple paragraph essays. A visual analysis of Cassandra's performance at post-multiple paragraph testing compared to baseline demonstrated real growth through large level changes on number of words, number of sentences, number of paragraphs, and number of transition words. The lack of variability at baseline is due to there being one untimed essay prompt at baseline. Cassandra's large level change is illustrated in her post-multiple paragraph number of words where she wrote a mean of 177.67 ( $SD = 53.67$ ) words where at baseline she wrote 40 words. There was large growth in the structure of Cassandra's essays. At post-multiple paragraph testing she wrote a mean of 17.33 ( $SD = 6.43$ ) sentences in a mean of 5.00 ( $SD = 1.73$ ) paragraphs. At baseline she wrote three sentences in one paragraph. The large growth is also displayed through the number of transition words at post-multiple paragraph testing with a mean of 10.00 ( $SD = 3.46$ ) transition words while at baseline she wrote one transition word.

In addition, there was a medium level change in the number of persuasive essay parts. At post-multiple paragraph testing she wrote a mean of 9.33 ( $SD = 2.89$ ) persuasive essay parts with a baseline of five persuasive essay parts. Lastly, there

was an increase in the holistic quality score at post-multiple paragraph testing with a mean score of 7.33 ( $SD = 2.89$ ) with a baseline holistic quality score of four. These results were supported by Cassandra's 100% PNDs from post-multiple paragraph measures to baseline measures on number of words, number of sentences, number of paragraphs, number of transition words, and number of persuasive essay parts. The only exception was she had 66.7 % PND for holistic quality score.

*Kevin.* Kevin completed three untimed post-multiple paragraph essays. A visual analysis of Kevin's performance at post-multiple paragraph testing compared to baseline demonstrated clear step changes in level on all essay measures with variability being greater at post-multiple paragraph testing than baseline. The lack of variability at baseline is due to there being one untimed essay prompt at baseline. The large level changes are illustrated in Figures 1, 2, and 3 and in the means and standard deviations for the measures with post-multiple paragraph compared to baseline scores.

Kevin wrote a mean of 163.33 ( $SD = 6.43$ ) words at post-multiple paragraph while he wrote 49 words at baseline. His essays structure at post-multiple paragraph were composed of a mean of 19.33 ( $SD = 0.58$ ) sentences with a mean of 6.00 ( $SD = 0.00$ ) paragraphs while at baseline his essays contained three sentences in one paragraph. In his large level change at post-multiple paragraph Kevin's number of transition words increased to a mean of 9.00 ( $SD = 0.00$ ) transition words from one transition word at baseline. His persuasive essays at post-multiple paragraph grew in strength with a large level change and he included a mean of 11.00 ( $SD = 0.00$ )

persuasive essay parts with an overall mean holistic quality score of 9.00 ( $SD = 0.00$ ). At baseline he included five persuasive essay parts with a holistic quality score of three. These results are support by Kevin's 100% PNDs on all essay measures from post-multiple paragraph to baseline.

*Caleb.* Caleb completed three untimed post-multiple paragraph essays. A visual analysis of Caleb's performance at post-multiple paragraph testing compared to baseline demonstrated positive trends with small level changes across all essay measures, except number of words that didn't included the level change. There was greater variability for Caleb at post-multiple paragraph testing than baseline due to one of his essay scores nearing or overlapping baseline due to him not completing the essay. There was also a lack of variability at baseline due to there being one untimed essay prompt at baseline. Caleb's growths are illustrated by the means and standard deviations from post-multiple paragraph testing compared to baseline testing.

Caleb wrote a mean of 90.00 ( $SD = 37.24$ ) words at post-multiple paragraph from 71 words at baseline. His essay structure at post-multiple paragraph essays included a mean of 11.33 ( $SD = 6.66$ ) sentences in a mean of 3.00 ( $SD = 1.73$ ) paragraphs. At baseline he had two sentences in zero paragraphs. Furthermore, he had a mean of 6.33 ( $SD = 3.79$ ) transition words at post-multiple paragraph testing to one at baseline. Caleb's persuasive essays at post-multiple paragraph grew with a small level changes. He included a mean of 8.67 ( $SD = 4.04$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ). At baseline he had three persuasive essay parts with a holistic quality score of two. These results were

supported by Caleb's 100% PNDs from post-multiple paragraph to baseline measures on the number of sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score. The only exception was his 66.7% PND for number of words.

*Jacob.* Jacob completed three untimed post-multiple paragraph essays. A visual analysis of Jacob's performance at post-multiple paragraph testing compared to baseline demonstrated large step change in level across all essay measures with greater variability than baseline due to there being no variability with one untimed essay prompt at baseline. Jacob's large level changes are illustrated by the means and standard deviations from post-multiple paragraph testing from baseline testing.

He wrote a mean of 151.33 ( $SD = 5.51$ ) words at post-multiple paragraph from 13 words at baseline. Jacob's essay structure at post-multiple paragraph essays included a mean of 19.00 ( $SD = 0.00$ ) sentences in a mean of 6.00 ( $SD = 0.00$ ) paragraphs. At baseline he included one sentence in zero paragraphs. In addition, he had a mean of 8.00 ( $SD = 0.00$ ) transition words at post-multiple paragraph to zero transition words at baseline. His persuasive essays at post-multiple paragraph had a large level change with a mean of 11.00 ( $SD = 0.00$ ) persuasive essay parts with an overall holistic quality score mean of 9.00 ( $SD = 0.00$ ) while at baseline he had one persuasive essay part with a holistic quality score of one. These results were supported by Jacob's 100% PNDs from post-multiple paragraph testing to baseline testing on all of the essay measures.

**Maintenance and Generalization.** Maintenance and generalization testing, students were administered four essays five weeks after the completion of post-multiple paragraph testing concluded. There was both maintenance and generalization prompts for each instructional phase. Essays were administered using identical procedures at baseline for the type of prompt. For the fluency prompts students were given timed essays when they had ten minutes to plan and write their essays while for the multiple paragraph prompts they were given untimed essays. The prompts were administered over four sessions starting with fluency testing for maintenance and generalization followed by multiple paragraph testing for maintenance and generalization.

***Fluency overall for maintenance and generalization.*** Students' fluency maintenance essay scores indicate large growth with greater variability across all essay measures. For the number of words at fluency maintenance mean of 99.20 ( $SD = 34.27$ ) from a baseline mean of 45.04 ( $SD = 26.20$ ). The structure of students' essays at fluency maintenance included a mean of 10.40 ( $SD = 5.27$ ) sentences in a mean of 2.00 ( $SD = 1.00$ ) paragraphs while at baseline the essays included a mean of 2.65 ( $SD = 2.04$ ) sentences in a mean of 0.38 ( $SD = 0.50$ ) paragraphs. In addition, there was a large growth in the number of transition words at fluency maintenance with a mean of 6.00 ( $SD = 1.73$ ) transition words while baseline had a mean of 0.77 ( $SD = 0.59$ ). Students' persuasive essays at fluency maintenance had a large level change. They included a mean of 8.00 ( $SD = 2.55$ ) persuasive essay parts with an overall holistic quality score mean of 6.00 ( $SD = 2.74$ ) while at baseline they had a

mean of 3.19 ( $SD = 1.26$ ) persuasive essay parts with a holistic quality mean score of 2.31 ( $SD = 0.88$ ).

The mean scores for number of words, number of sentences, number of transition words, number of persuasive essay parts, and holistic quality scores for students essay measures at fluency maintenance were statistically significant from baseline scores, as indicated in Table 2 (all  $p$ 's < .05, according to Wilcoxon matched-pairs, signed rank tests, from baseline to fluency maintenance testing). These positive results are supported by the 100% PNDs for number of transition words, and persuasive essay parts, while 80% PNDs for number of sentences, and holistic quality score, 60% PNDs for total number of words and number of paragraphs at fluency maintenance performance on essay measures compared to baseline measures for all students (see Table 4).

Students' fluency generalization essay scores indicate large growth with greater variability across all essays measures. These are illustrated by the means and standard deviations for the essay measures at fluency generalization compared to baseline. For the number of words at fluency generalization a mean of 88.60 ( $SD = 37.07$ ) from a baseline mean of 45.04 ( $SD = 26.20$ ). The overall students' structure for their essays at fluency generalization included a mean of 9.80 ( $SD = 5.98$ ) sentences in a mean of 1.80 ( $SD = 1.30$ ) paragraphs while at baseline the essays included a mean of 2.65 ( $SD = 2.04$ ) sentences in a mean of 0.38 ( $SD = 0.50$ ) paragraphs. There was large growth in the number of transition words at fluency generalization with a mean of 5.04 ( $SD = 2.30$ ) transition words while at baseline they



had a mean of 0.77 ( $SD = 0.59$ ) transition words. Students' persuasive essays at fluency generalization were much stronger because they included a mean of 7.60 ( $SD = 2.79$ ) persuasive essay parts with an overall holistic quality score mean of 6.20 ( $SD = 2.78$ ) while at baseline they had a mean of 3.19 ( $SD = 1.26$ ) persuasive essay parts with a holistic quality mean score of 2.31 ( $SD = 0.88$ ).

The mean scores for number of sentences, number of transition words, number of persuasive essay parts, and holistic quality scores overall at fluency generalization were statistically significant from baseline scores, as indicated in Table 2 (all  $p$ 's  $< .05$ , according to Wilcoxon matched-pairs, signed rank tests, from baseline to fluency maintenance testing). These positive results are supported by the 100% PND for number of transition words, while 80% PNDs for number of sentences, persuasive essay parts, and holistic quality score and 60% PNDs for total number of words, and number of paragraphs at fluency generalization performance compared to baseline measures for all students (see Table 4).

*Individual student fluency maintenance and generalization performance.*

Overall, fluency maintenance and generalization performances on timed persuasive essays were higher and varied from baseline performance, as can be seen in Figures 1, 2, and 3. Students completed between three and five timed essays at baseline and one essays each for fluency maintenance and generalization testing. Therefore, at fluency maintenance and generalization testing there was no variability in student performance.

*Garrett.* Garrett completed one timed fluency maintenance essay. A visual analysis of Garrett's performance at fluency maintenance testing compared to baseline demonstrated a large step change in level change across all essay measures with no variability. He had 100% PNDs from fluency maintenance to baseline on all essay measures. His performance also increased from post-fluency testing on all measures at fluency maintenance testing. Garrett's number of words at fluency maintenance testing was 155 compared to his post-fluency mean of 144.33 ( $SD = 4.51$ ). His essay structure at fluency maintenance was more robust with 19 sentences in three paragraphs while at post-fluency testing he included a mean of 10.00 ( $SD = 1.00$ ) sentences in a mean of one paragraph. He also used more transition words at fluency maintenance that included nine transition words while at post-fluency he used a mean of 6.00 ( $SD = 0.00$ ) transition words. Garrett's persuasive essays at fluency maintenance included 11 persuasive essay parts with an holistic quality score of nine from post-fluency mean of 10.67 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 8.67 ( $SD = 0.58$ ).

Garrett also complete one timed fluency generalization essay. A visual analysis of Garrett's performance at fluency generalization testing compared to baseline also demonstrated a large step change in level change across all essay measures with no variability. He had 100% PNDs from fluency generalization measures to baseline measures on all essay measures. Garrett's performance also increased from post-fluency testing on all measures at fluency generalization testing. For Garrett's number of words at fluency generalization testing was 147 compared to

his post-fluency mean of 144.33 ( $SD = 4.51$ ). His essay structure at fluency generalization was more substantial with 20 sentences in four paragraphs while at post-fluency testing he included a mean of 10.00 ( $SD = 1.00$ ) sentences in a mean of one paragraph. He included more transition words at fluency generalization which had nine transition words while at post-fluency testing he had a mean of 6.00 ( $SD = 0.00$ ). Garrett's persuasive essays were more robust at fluency generalization that included 11 persuasive essay parts with an holistic quality score of nine from post-fluency mean of 10.67 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 8.67 ( $SD = 0.58$ ).

*Jamal.* Jamal completed one timed fluency maintenance essay. A visual analysis of Jamal's performance at fluency maintenance testing compared to baseline demonstrated large level changes on all essay measures with no variability. He had 100% PNDs from fluency maintenance to baseline on all essay measures. His performance paralleled his post-fluency scores on most essay measures, except for holistic quality score that improved from post-fluency testing mean of 7.33 to a nine at fluency maintenance testing. His number of words at fluency maintenance testing was 104 compared to his post-fluency mean of 106.00 ( $SD = 10.54$ ). Jamal's essay structure at fluency maintenance had nine sentences in two paragraphs while at post-fluency testing he included a mean of 10.67 ( $SD = 0.58$ ) sentences in a mean of one paragraph. His transition words at fluency maintenance had six transition words while he had a mean of 5.33 ( $SD = 0.58$ ) at post-fluency testing. Jamal's persuasive essays at fluency maintenance included ten persuasive essay parts with an holistic

quality score of nine from post-fluency mean of 10.67 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

Jamal also completed one timed fluency generalization essay. A visual analysis of Jamal's performance at fluency generalization testing compared to baseline also demonstrated large level changes on all essay measures with no variability. He had 100% PNDs from fluency generalization to baseline on all essay measures. His performance also paralleled his post-fluency and fluency maintenance scores on most essay measures, except for holistic quality score improved from post-fluency testing that had a mean of 7.33 to a nine at fluency generalization testing. His number of words at fluency generalization testing had 99 words compared to his post-fluency mean of 106.00 ( $SD = 10.54$ ). Jamal's essay structure at fluency generalization included 10 sentences in one paragraph while at post-fluency testing he had a mean of 10.67 ( $SD = 0.58$ ) sentences in a mean of one paragraph. At fluency generalization he had six transition words with a mean of 5.33 ( $SD = 0.58$ ) transition words at post-fluency. Jamal's persuasive essays at fluency generalization included 10 persuasive essay parts with an holistic quality score of nine from post-fluency mean of 10.67 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

*Cassandra.* Cassandra completed one timed fluency maintenance essay. A visual analysis of Cassandra's performance at fluency maintenance testing compared to baseline demonstrated real growth through a large level change in the number of transition words, while she had a medium level change on number of sentences,

number of paragraphs, and number of persuasive essay parts, and a small level change in the number of words. However, the visual analysis for holistic quality score of a four at fluency maintenance testing showed improvement from her baseline mean of 2.75 ( $SD = 0.96$ ), but her fluency maintenance performance overlapped with her best baseline performance. She had 100% PNDs from fluency maintenance to baseline on all of the essays measures. Her performance at fluency maintenance was similar to her post-fluency testing in the number of words, number of sentences, and number of paragraphs, but there were slight decreases in performance at fluency maintenance on the number of transition words, number of persuasive essay parts, and holistic quality score compared to post-fluency testing.

Cassandra wrote 94 words at fluency maintenance testing compared to her post-fluency mean of 78.67 ( $SD = 7.51$ ). Her essay structure at fluency maintenance was composed of 11 sentences in three paragraphs while her post-fluency essays were composed of a mean of 11.00 ( $SD = 2.00$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. Her essay at fluency maintenance had five transition words while her post-fluency essay included a mean of 6.00 ( $SD = 1.00$ ) transition word. Cassandra's persuasive essays structure declined at fluency maintenance included five persuasive parts with a holistic quality score of four from post-fluency mean of 11.00 ( $SD = 2.00$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

Cassandra also completed one timed fluency generalization essay. A visual analysis of Cassandra's performance at fluency generalization testing compared to

baseline is similar to fluency maintenance testing. The visual analysis shows Cassandra has a large level change in the number of transition words with medium level changes on number of sentences, number of paragraphs, and number of persuasive essay parts, and a small level change in the number of words. In addition, the visual analysis for holistic quality score at fluency generalization was a four which showed an increase without a level change from her baseline mean of 2.75 ( $SD = 0.96$ ), but there was an overlap from post generalization to her best baseline performance. Cassandra had 100% PNDs from fluency generalization to baseline on all six of the essays measures. Her performance at fluency generalization was similar to her post-fluency testing and fluency maintenance testing in the number of words, number of sentences, and number of paragraphs. However, there was a slight decrease in her performance at fluency generalization testing from post-fluency on the number of transition words, number of persuasive essay parts, and holistic quality score.

Cassandra's fluency generalization essay had 57 words compared to her post-fluency essay mean of 78.67 ( $SD = 7.51$ ). Her essay structure at fluency generalization was composed of seven sentences in two paragraphs while her post-fluency essays was composed of a mean of 11.00 ( $SD = 2.00$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. Furthermore, Cassandra's essay at fluency generalization contained four transition words while her post-fluency essay contained a mean of 6.00 ( $SD = 1.00$ ) transition words. In addition, her persuasive essays declined at fluency generalization to include five persuasive parts with a holistic

quality score of four from post-fluency mean of 11.00 ( $SD = 2.00$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

*Kevin.* Kevin completed one timed fluency maintenance essay. A visual analysis of Kevin's performance at fluency maintenance testing compared to baseline demonstrated a medium level change in the number of transition words with small level changes in the number of persuasive essay parts and holistic quality score. For the other three measures visual analysis at fluency maintenance compared to baseline shows that the number of words, number of sentences, and number of paragraphs overlaps baseline around Kevin's highest baseline performance. His performance at fluency maintenance was lower than his post-fluency testing in the number of words, number of sentences, number of persuasive essay parts, and holistic quality score with similarities in number of paragraphs and number of transition words. Kevin's results were supported by 100% PNDs from fluency maintenance measures to baseline measures on number of transition words, number of persuasive essay parts, and number of holistic quality score. He had 0% PNDs for number of words, number of sentences, and number of paragraphs.

At fluency maintenance testing Kevin wrote 72 words compared to a mean of 90.00 ( $SD = 11.53$ ) words at post-fluency testing. His essay structure at fluency maintenance was composed of five sentences in one paragraph while his post-fluency essays were composed of a mean of 10.00 ( $SD = 0.00$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. In addition, his essay at fluency maintenance contained five transition words compared to a mean of 5.33 ( $SD = 0.58$ ) transition words at post-

fluency testing. Lastly, Kevin's persuasive essays decreased at fluency maintenance to include six persuasive parts with a holistic quality score of four from a post-fluency mean of 10.33 ( $SD = 1.16$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

Kevin also completed one timed fluency generalization essay. A visual analysis of Kevin's performance at fluency generalization testing compared to baseline demonstrated a small level change in the number of transition words. For the other measures visual analysis from fluency generalization to baseline showed that the number of words, number of sentences, number of paragraphs, number of persuasive essay parts, and holistic quality scores overlapped with Kevin's highest baseline performance. His performance at fluency generalization was lower than his post-fluency testing in the number of words, number of sentences, number of transition words, number of persuasive essay parts, and holistic quality score with similarities in number of paragraphs. Yet, his performance at fluency generalization testing was similar to his fluency maintenance testing. Kevin's results were supported by 100% PND from fluency generalization to baseline on the number of transition words. He had 0% PNDs for number of words, number of sentences, number of paragraphs, number of persuasive essay parts, and holistic quality score.

Kevin's fluency generalization essay had 58 words compared to a mean of 90.00 ( $SD = 11.53$ ) words at post-fluency testing. While his essay structure at fluency generalization testing was the same as his fluency maintenance testing composed of five sentences in one paragraph while his post-fluency essays were



composed of a mean of 10.00 ( $SD = 0.00$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. His fluency generalization essays contained three transition words compared to a mean of 5.33 ( $SD = 0.58$ ) transition words at post-fluency testing. Finally, Kevin's persuasive essays at fluency generalization included five persuasive parts with a holistic quality score of three while at post-fluency his mean number of persuasive essay parts was 10.33 ( $SD = 1.16$ ) with a holistic quality score mean of 7.33 ( $SD = 2.89$ ).

*Caleb.* Caleb completed one timed fluency maintenance essay. A visual analysis of Caleb's performance at fluency maintenance testing compared to baseline demonstrated medium level changes in the number of transition words and number of persuasive essay parts while he had a small level change in the number of sentences. For the other three measures visual analysis from fluency maintenance to baseline testing shows that the number of words, number of paragraphs, and holistic quality score overlaps with Caleb's highest baseline performance. His performance overall at fluency maintenance testing was similar to his post-fluency for all measures, except for persuasive essay parts and holistic quality score which were lower at fluency maintenance. These results were supported by his 100% PNDs from fluency maintenance to baseline on the number of sentences, number of transition words, and number of persuasive essay parts. While he also had 0% PNDs on number of words, number of paragraphs, and holistic quality score.

Caleb's essay performance at fluency maintenance testing contained 71 words to a post-fluency mean of 77.00 ( $SD = 9.54$ ) words. His essay structure at fluency

maintenance testing had eight sentences in one paragraph while his post-fluency essays had a mean of 10.00 ( $SD = 2.65$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. At fluency maintenance his essay contained five transition words while his post-fluency essays contained a mean of 6.00 ( $SD = 0.00$ ) transition words. Finally, Caleb's persuasive essays decreased at fluency maintenance testing to include eight persuasive parts with a holistic quality score of four from a post-fluency mean of 10.33 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 8.67 ( $SD = 0.58$ ).

Caleb also completed one timed fluency generalization essay. A visual analysis of Caleb's performance at fluency generalization testing compared to baseline demonstrated medium level changes in the number of transition words and number of persuasive essay parts while he had small level changes in number of sentences and holistic quality score. For the other two measures visual analysis from fluency generalization to baseline testing showed that, the number of words and number of paragraphs overlaps with Caleb's best baseline performance. His performance overall at fluency generalization testing was similar to his post-fluency and fluency maintenance testing. These results were supported by 100% PNDs from fluency generalization to baseline on the number of sentences, number of transition words, and number of persuasive essay parts. He had 0% PNDs on number of words, number of paragraphs, and holistic quality score.

At fluency generalization testing Caleb wrote 82 words to a post-fluency mean of 77.00 ( $SD = 9.54$ ) words. His essay structure at fluency generalization

testing had seven sentences in one paragraph while his post-fluency essays had a mean of 10.00 ( $SD = 2.65$ ) sentences in a mean of 1.00 ( $SD = 0.00$ ) paragraph. Caleb's fluency generalization essay contained five transition words while his post-fluency essays contained a mean of 6.00 ( $SD = 0.00$ ) transition words. Lastly, his persuasive essay decreased slightly at fluency generalization testing to include seven persuasive parts with a holistic quality score of six from a post-fluency mean of 10.33 ( $SD = 0.58$ ) persuasive essay parts with an overall holistic quality score mean of 8.67 ( $SD = 0.58$ ).

*Jacob.* Jacob declined to participate in maintenance and generalization testing. Therefore, there are no results for him.

***Multiple paragraph overall results for maintenance and generalization.***

Overall students' results are reported for maintenance first followed by the generalization results. The overall multiple paragraph maintenance essay results indicate large growth with greater variability across all essays measures. The essay number of words at multiple paragraph maintenance had a mean of 137.20 ( $SD = 40.81$ ) words from a baseline mean of 35.86 ( $SD = 20.23$ ) words. The students' essay structure at multiple paragraph maintenance testing had a mean of 16.20 ( $SD = 4.26$ ) sentences in a mean of 4.00 ( $SD = 1.41$ ) paragraphs while at baseline the essays contained a mean of 1.71 ( $SD = 1.11$ ) sentences in a mean of 0.29 ( $SD = 0.49$ ) paragraphs. Students' essays overall number of transition words at multiple paragraph maintenance had a mean of 8.40 ( $SD = 1.52$ ) transition words while at baseline they had a mean of 0.71 ( $SD = 0.49$ ) transition words. The students'

persuasive essays at multiple paragraph maintenance testing contained a mean of 9.60 ( $SD = 1.52$ ) persuasive essay parts with an overall holistic quality mean score of 8.40 ( $SD = 0.89$ ) while at baseline they had a mean of 3.14 ( $SD = 1.46$ ) persuasive essay parts with a holistic quality mean score of 2.57 ( $SD = 0.98$ ).

The mean scores for the overall students' performance on all the essay measures at multiple paragraph maintenance testing were statistically significant from baseline scores, as indicated in Table 3 (all  $p$ 's < .05, according to Wilcoxon matched-pairs, signed rank tests, from baseline to multiple paragraph maintenance testing). These positive results are supported by the 100% PNDs for number of words, number of sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score at multiple paragraph maintenance performance compared to baseline performance on all measures (see Table 4).

Students' overall multiple paragraph generalization essay scores also indicate large growth with greater variability across all essays measures. These results are illustrated by the means and standard deviations for the essay measures at multiple paragraph generalization testing compared to baseline testing. The number of words at multiple paragraph generalization testing resulted in a mean of 136.00 ( $SD = 33.49$ ) words from a baseline mean of 35.86 ( $SD = 20.23$ ) words. The overall students' structure for their essays at multiple paragraph generalization testing contained a mean of 17.20 ( $SD = 4.32$ ) sentences in a mean of 4.40 ( $SD = 0.89$ ) paragraphs while at baseline the essays had a mean of 1.71 ( $SD = 1.11$ ) sentences in a mean of 0.29 ( $SD$

= 0.49) paragraphs. In addition, the overall mean number of transition words at multiple paragraph generalization was 8.60 ( $SD = 1.52$ ) compared to a baseline mean of 0.71 ( $SD = 0.49$ ) transition words. The students' overall persuasive essays at multiple paragraph generalization were much stronger because they included a mean of 9.40 ( $SD = 1.34$ ) persuasive essay parts with an overall holistic quality score mean of 8.20 ( $SD = 1.30$ ) while at baseline they had a mean of 3.14 ( $SD = 1.46$ ) persuasive essay part with a holistic quality mean score of 2.57 ( $SD = 0.98$ ).

The mean scores for the overall students' performance at multiple paragraph generalization testing on all measures compared to baseline testing scores were statistically significant, as indicated in Table 3 (all  $p$ 's < .05, according to Wilcoxon matched-pairs, signed rank tests, from baseline to fluency maintenance testing). These results are supported by the 100% PNDs for number of words, number of sentences, number of paragraphs, number of transition words, number of persuasive essay parts, and holistic quality score at multiple paragraph generalization performance compared to baseline measures (see Table 4).

*Individual student multiple paragraph maintenance and generalization performance.* Overall, multiple paragraph maintenance and generalization performances on untimed persuasive essays were higher than baseline performance, as can be seen in Figures 1, 2, and 3. Students completed one untimed essay at baseline and one untimed essay each for multiple paragraph maintenance and generalization testing. Therefore, at fluency maintenance and generalization testing there was no variability in student performance.

*Garrett.* Garrett completed one untimed multiple paragraph maintenance essay. A visual analysis of Garrett's performance at multiple paragraph maintenance testing compared to baseline demonstrated a large step change in level across all essay measures with no variability. He had 100% PNDs from multiple paragraph maintenance testing to baseline on all essay measures. His performance either was similar to or higher than post-multiple paragraph testing on all measures at multiple paragraph maintenance testing. Garrett's number of words at multiple paragraph maintenance testing was 174 compared to his post-multiple paragraph mean of 147.33 ( $SD = 9.02$ ) words. His multiple paragraph maintenance essay structure was more elaborate with 22 sentences in five paragraphs compared to post-multiple paragraph testing when he had a mean of 17.67 ( $SD = 2.08$ ) sentences in a mean of 4.33 ( $SD = 1.53$ ) paragraphs. He used a similar number of transition words in his essay at multiple paragraph maintenance that included eight transition words while at post-multiple paragraph he had a mean of 10.00 ( $SD = 1.73$ ) transition words. Garrett's persuasive essay at multiple paragraph maintenance testing compared to post-multiple paragraph testing were identical with 11 persuasive essay parts with a holistic quality score of nine without any variability.

Garrett also completed one untimed multiple paragraph generalization essay. A visual analysis of Garrett's performance at multiple paragraph generalization testing compared to baseline demonstrated a large step change in level across all essay measures with no variability, which is similar to his multiple paragraph maintenance testing. His results had 100% PNDs from multiple paragraph

generalization to baseline on all of the essay measures. His performance was consistent to his multiple paragraph maintenance testing and his post-multiple paragraph testing. For Garrett's essay at multiple paragraph generalization testing contained 173 words compared to his post-multiple paragraph mean of 147.33 ( $SD = 9.02$ ) words. While his essay at multiple paragraph generalization structure was more involved with 24 sentences in five paragraphs while at post-multiple paragraph testing his essay had a mean of 17.67 ( $SD = 2.08$ ) sentences in a mean of 4.33 ( $SD = 1.53$ ) paragraphs. He had eight transition words in his multiple paragraph generalization essay while at post-multiple paragraph he had a mean of 10.00 ( $SD = 1.73$ ) transition words. Furthermore, Garrett's persuasive essay was identical at multiple paragraph generalization testing to his post-multiple paragraph testing with 11 persuasive essay parts with a holistic quality score of nine without any variability.

*Jamal.* Jamal completed an untimed multiple paragraph maintenance essay. A visual analysis of Jamal's performance at multiple paragraph maintenance testing compared to baseline demonstrated large step change in level on all essay measures with no variability. Jamal's results were further supported with 100% PNDs from multiple paragraph maintenance to baseline on all essay measures. Jamal's performance paralleled his post-multiple paragraph scores on all essay measures. His essay contained 148 words at multiple paragraph maintenance testing compared to his post-multiple paragraph mean of 159.67 ( $SD = 13.05$ ) words. While Jamal's essay structure at multiple paragraph maintenance testing included 18 sentences in five paragraphs while at post-multiple paragraph testing he included a mean of 18.67 ( $SD$

= 0.58) sentences in a mean of 6.00 ( $SD = 0.00$ ) paragraphs. His essays at multiple paragraph maintenance testing had nine transition words while he had a mean of 9.00 ( $SD = 0.00$ ) transition words at post-multiple paragraph. Jamal's persuasive essay at multiple paragraph maintenance paralleled his performance at post-multiple paragraph to include 11 persuasive essay parts with a holistic quality score of nine without any variability.

Jamal also completed one untimed fluency generalization essay. A visual analysis of Jamal's performance at multiple paragraph generalization testing compared to baseline demonstrated large step changes in level across all of the essay measures with no variability. His results were supported by 100% PNDs from multiple paragraph generalization testing to baseline testing across all essay measures. Jamal's performance also paralleled his post-multiple paragraph and multiple paragraph maintenance testing scores on all essay measures. His essay at multiple paragraph generalization contained 148 words compared to his post-multiple paragraph mean of 159.67 ( $SD = 13.05$ ) words. Jamal's essay structure at multiple paragraph generalization had 18 sentences in five paragraphs while at post-multiple paragraph testing he had mean of 18.67 ( $SD = 0.58$ ) sentences in a mean of 6.00 ( $SD = 0.00$ ) paragraphs. His essay at multiple paragraph generalization and post-multiple paragraph testing both contained or had a mean of 9.00 ( $SD = 0.00$ ) transition words. While Jamal's persuasive essay at multiple paragraph generalization testing was identical to his post-multiple paragraph testing while each had 10 persuasive essay parts with a holistic quality score of nine without any variability.



*Cassandra.* Cassandra completed one untimed multiple paragraph maintenance essay. A visual analysis of Cassandra's performance at multiple paragraph maintenance testing compared to baseline demonstrated large step changes in level across all essay measures with no variability. Her results were supported by 100% PNDs from multiple paragraph maintenance testing scores compared to baseline scores across all of the essays measures. Her performance at multiple paragraph maintenance testing was similar to her performance at post-multiple paragraph testing. Cassandra essay at multiple paragraph contained 145 words compared to her essay at post-multiple paragraph that had a mean of 177.67 ( $SD = 53.67$ ) words. While her essay structure at multiple paragraph maintenance had 17 sentences in five paragraphs which was similar to her post-multiple paragraph structure that had a mean of 17.33 ( $SD = 6.43$ ) sentences in a mean of 5.00 ( $SD = 1.73$ ) paragraphs. Her essay contained 10 transition words at multiple paragraph maintenance testing while she also had a mean of 10.00 ( $SD = 3.46$ ) transition words at post-multiple paragraph testing. Cassandra's persuasive essay at multiple paragraph maintenance had eight persuasive parts with a holistic quality score of eight which was similar to her essay performance at post-multiple paragraph contained a mean of 9.33 ( $SD = 2.89$ ) persuasive essay parts with an overall holistic quality mean score of 7.33 ( $SD = 2.89$ ).

Cassandra also completed one untimed multiple paragraph generalization essay. The visual analysis of Cassandra's performance at multiple paragraph generalization testing compared to baseline parallels her results at multiple paragraph

maintenance testing. A visual analysis of Cassandra's performance at multiple paragraph generalization testing compared to baseline demonstrated large step change in level across all essay measures. Cassandra had 100% PNDs from multiple paragraph generalization testing compared to baseline testing across all essay measures. Her performance at multiple paragraph generalization was similar to her performances at post-multiple paragraph testing and multiple paragraph maintenance testing. Cassandra's essay at multiple paragraph generalization contained 149 words compared to her post-multiple paragraph mean of 177.67 ( $SD = 53.67$ ). Her multiple paragraph generalization essay structure contained 17 sentences in five paragraphs which is identical to her multiple paragraph maintenance essay and similar to her post-multiple paragraph essays that contained a mean of 17.33 ( $SD = 6.43$ ) sentences in a mean of 5.00 ( $SD = 1.73$ ) paragraphs. Cassandra's essay at multiple paragraph generalization contained 10 transition words while her post-multiple paragraph also had a mean of 10.00 ( $SD = 3.46$ ) transition words. Cassandra's performance on her persuasive essays at multiple paragraph generalization contained eight persuasive parts with a holistic quality score of eight which is identical to her multiple paragraph maintenance testing and similar to her post-multiple paragraph testing that contained a mean of 9.33 ( $SD = 2.89$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

*Kevin.* Kevin completed one untimed multiple paragraph maintenance essay. A visual analysis of Kevin's performance at multiple paragraph maintenance testing compared to baseline demonstrated positive growth with large step changes in level

across all essay measures. Kevin's results are supported by 100% PNDs from multiple paragraph maintenance testing to baseline testing across all essay measures. His performance at multiple paragraph maintenance testing was similar to his performance at post-multiple paragraph testing. Kevin's multiple paragraph maintenance essay contained 107 words compared to his post-multiple paragraph essay mean of 163.33 ( $SD = 6.43$ ) words. His essay structure at multiple paragraph maintenance essay contained 12 sentences in two paragraphs while his post-multiple paragraph essays had a mean of 19.33 ( $SD = 0.58$ ) sentences in a mean of 6.00 ( $SD = 0.00$ ) paragraphs. In addition, his multiple paragraph maintenance and post-multiple paragraph essays had a mean of 9.00 ( $SD = 0.00$ ) transition words. Finally, Kevin's persuasive essay at multiple paragraph maintenance contained 10 persuasive parts with a holistic quality score of nine which is similar to his post-multiple paragraph performance that has a mean of 11.00 ( $SD = 0.00$ ) persuasive essay parts with an overall holistic quality score mean of 9.00 ( $SD = 0.00$ ).

Kevin also completed one untimed multiple paragraph generalization essay. A visual analysis of Kevin's performance at multiple paragraph generalization testing compared to baseline demonstrated large step change in level across all essay measures with no variability. Further, this visual analysis of Kevin's performance at multiple paragraph generalization testing is consistent to his performance at both post-multiple paragraph and multiple paragraph maintenance testing. His results were supported by 100% PNDs from multiple paragraph generalization testing from baseline testing across all essay measures. Kevin's essay at multiple paragraph

generalization testing had 126 words compared to a mean of 163.33 ( $SD = 6.43$ ) words at post-multiple paragraph testing. While his essay structure at multiple paragraph generalization testing was composed of 14 sentences in two paragraphs while his post-multiple paragraph essays were composed of a mean of 19.33 ( $SD = 0.58$ ) sentences in a mean of 6.00 ( $SD = 0.00$ ) paragraphs. His multiple paragraph generalization and post-multiple paragraph essays both contained nine transition words with no variability. Lastly, his persuasive essay at multiple paragraph generalization had 10 persuasive parts with a holistic quality score of nine, which is identical to his multiple paragraph maintenance testing and similar to his post-multiple paragraph testing had no variability with a mean of 11 persuasive essay parts with a holistic quality score mean of nine.

*Caleb.* Caleb completed one untimed multiple paragraph maintenance essay. A visual analysis of Caleb's performance at multiple paragraph maintenance testing compared to baseline demonstrated large step changes in level across all essay measures with no variability, except he had a small level change in number of words. His performance overall at multiple paragraph maintenance testing parallels his post-multiple paragraph performance across all measures. His results are supported by 100% PNDs from multiple paragraph maintenance testing to baseline testing across all essay measures. Caleb's essay at multiple paragraph maintenance testing contained 84 words compared to a post-multiple paragraph mean of 90.00 ( $SD = 37.24$ ) words. His multiple paragraph maintenance essay structure had 12 sentences in three paragraphs while his post-multiple paragraph essays had a mean of 11.33 ( $SD$

= 6.66) sentences with a mean of 3.00 ( $SD = 1.73$ ) paragraphs. In addition, his multiple paragraph maintenance essay had six transition words which is comparable to the post-multiple paragraph mean of 6.33 ( $SD = 3.79$ ) transition words. Caleb's persuasive essay at multiple paragraph maintenance testing contained eight persuasive parts with a holistic quality score of seven, which was similar to his post-multiple paragraph performance that had a mean of 8.67 ( $SD = 4.04$ ) persuasive essay parts with an overall holistic quality score mean of 8.67 ( $SD = 2.89$ ).

Caleb also completed one untimed multiple paragraph generalization essay. The visual analysis of Caleb's performance at post-multiple paragraph testing compared to baseline parallels his results at multiple paragraph maintenance testing. A visual analysis of Caleb's performance at multiple paragraph generalization testing compared to baseline testing shows large step changes across all essay measures with no variability, except there was a small level change in his number of words. His results were supported by 100% PNDs from multiple paragraph generalization testing to baseline testing across all essay measures. His performance at multiple paragraph generalization testing was similar to his post-multiple paragraph and multiple paragraph maintenance testing. At multiple paragraph generalization testing Caleb wrote 84 words compared to a mean of 90.00 ( $SD = 37.24$ ) words at post-multiple paragraph testing. While his essay structure at multiple paragraph generalization testing was composed of 13 sentences in four paragraphs his post-multiple paragraph essays contained a mean of 11.33 ( $SD = 6.66$ ) sentences in a mean of 3.00 ( $SD = 1.73$ ) paragraphs. Caleb's multiple paragraph generalization essay had six transition

words while his post-multiple paragraph essays contained a similar mean of 6.33 ( $SD = 3.79$ ) transition words. Lastly, his persuasive essay at multiple paragraph generalization testing had eight persuasive parts with a holistic quality score of six from a similar post-multiple paragraph mean of 8.67 ( $SD = 4.04$ ) persuasive essay parts with an overall holistic quality score mean of 7.33 ( $SD = 2.89$ ).

*Jacob.* Jacob declined to participate in maintenance and generalization testing. Therefore, there are no results for him.

**Self-Efficacy.** The self-efficacy measure was given with an untimed writing prompt at baseline testing, post-multiple paragraph testing, and maintenance testing. The self-efficacy measure was a task specific measure that students were asked thirteen questions about how confident they were in doing a specific component in the persuasive essay that they would write following the measure. Students rated their confidence level in a Likert scale from one that corresponded to a 0% confidence to a five that corresponded to a 100% confident.

There were too few students ( $n = 7$  at baseline and  $n = 6$  at post-multiple paragraph, and  $n = 5$  at multiple paragraph maintenance testing) to validate the self-efficacy measures. To analyze these data composite scores were computed for a total at each phase of testing.

The students overall made descriptive gains on their self-efficacy, with a baseline pretest mean of 39.93 ( $SD = 9.99$ ), post-multiple paragraph testing mean of 55.25 ( $SD = 6.35$ ), and maintenance testing mean of 59.60 ( $SD = 4.98$ ). The difference between baseline and posttest were statistically significant according to the

Wilcoxon matched-pairs, signed rank test,  $p = .03$ . In addition, the difference between baseline and maintenance testing was also statistically significant according to the Wilcoxon matched-pairs, signed rank test,  $p = .04$ . However, the difference between post-test and maintenance testing was not statistically significant according to the Wilcoxon matched-pairs, signed rank test,  $p = .18$ .

### **Social Validity, Student Interviews, and Strategy Reports**

The post-intervention student interview contained the social validity, student interview, and strategy report. The interview were conducted by the instructor with one student at a time and consisted of twelve questions that the student were asked while being audiotaped (see Appendix J for a list of the specific questions). The interview was used to determine the students' perceptions and knowledge of the POW+TREE strategy, its usefulness, the POW+TREE lessons, and which instructional phase they preferred.

All the students recalled all the parts to the mnemonic strategy, except for Garrett who forgot to mention the specific number of reasons and counter reason, and to examine. However, a few students did forget to mention components that were not part of the mnemonic strategy, but are included in the strategy. For example five students (Jamal, Kevin, Cassandra, Caleb, and Jacob) did not mention the use of transition words and three students did not mention the refute for the counter reason. Then when students were asked to draw the graphic organize which they learned independently during instruction. Most of the students did remember how to draw the graphic organizer with all the components, except Cassandra who only included three

of the eleven components on the graphic organizer. However, she was able to recall the mnemonic strategy she had learned during instruction, which contains most of the components in the learned strategy.

When students were asked what they liked about the strategy, they mentioned that the strategy was “easy” (Garrett and Caleb) and it helped them be “organized” (Jamal and Kevin). Students described what they liked personally about the strategy. Kevin reported the strategy helped him “expand on his ideas” while Jacob reported it helped him “stay on topic.” However, Cassandra was unable to think of what she liked about the strategy. Students were also asked whether the POW+TREE strategy helped them with their writing. Most students reported the strategy helped them. A few overall comments from students on how it helped them included, “It helped me write good persuasive essays (Jamal)” or “It helped make it easier to think of ideas and to write (Garrett and Jacob).” Students also reported specific components of the strategy that helped them. Jamal reported, “It kept everything organized.” Garrett reported, “It helped him avoid writers block and made his essay make more sense.” While Kevin reported, it “increased the size of his essay through elaboration.” Cassandra reported it helped her with paragraphs. Then Jacob reported the strategy helped him, “actually write the correct parts of the essay.”

Students were asked what they had learned and what helped them the most. Overall, students reported the transition word list, counter reasons, and explaining themselves was very beneficial. Kevin explained he learned to use transitions among sentences and paragraphs. Jamal reported the counter reason was the best. He



learned when persuading someone else, you need to bring up the other side, but you have to come back and say why your side is better. Furthermore, students also reported that reasons, the strategy, the graphic organizer, and self-statements also helped them. Jacob elaborated that self-statements helped him think positively, so he could write his essay.

Students were also asked what they would add or change anything about the POW+TREE lessons. All students reported they would not change anything about the lesson, but when asked what they would change a few had suggestions. Jamal suggested including a larger transition word list. Caleb suggested writing longer multi-paragraph essays.

Students were asked about specific components about the strategy. First, students were asked about their opinion on the two different instructional methods. Students were asked if they preferred the one-paragraph 10-minute essays or the multiple paragraph essays that were untimed. Half the students (Jamal, Garrett, and Kevin) reported preference for the 10-minute essay because it was shorter, faster, easier, and pushed you. While the other half of the students (Cassandra, Caleb, and Jacob) preferred the untimed essay because it gave you time to think, time to do the essay, and you do not feel pressured. Secondly, students were asked about the importance of counter reasons. Most students reported the counter reason was important. They stated counter reasons make arguments stronger, reduces bias, and presents another side to the argument.

Students were asked if they had used the strategy they learned in any other classes and whether the strategy could help other students. Half the students (Jamal, Garrett, and Cassandra) reported they had not used the strategy in any class yet, but talked about how they could in English class when they had to write persuasive essays. The other half of the students (Kevin, Caleb, and Jacob) reported that they had used the strategy outside of the study for the standardized writing assessment, and English class. Yet, all of the students reported this strategy could help other students write better persuasive essays. A number of students went on to say it could help other students with good ideas, to elaborate, organize their thoughts, and write better papers.

## 5. DISCUSSION

This study examined the potential benefits of SRSD for persuasive writing using the POW+TREE strategy with seventh and eighth grades students with EBD. Overall findings revealed, (a) all students improved from baseline to post-fluency on all essay measures, (b) all students improved from baseline to post-multiple paragraph on all essay measures, (c) maintenance and generalization performance across fluency and multiple paragraph measures was higher than baseline but somewhat lower on some measures than at post treatment, (c) students' WJF descriptively increased, but not significantly from baseline to post-test, (d) students' overall significantly improved on their persuasive writing self-efficacy from baseline to post-test and maintenance, (e) students' strategy reports indicated positive attitudes about learning the writing strategy, and (f) on-task behavior was lower than anticipated. Each of these findings is discussed separately next. Finally, the educational implications, limitation, and recommendations for future research from this study are discussed.

### **Standardized Testing**

Although students improved descriptively on the WJF, these differences were not statistically significant. The six participants obtained a 17% increase in the standard score. Since the sample was so small, there was limited power in the analysis that may have resulted in a type II error.

There are four previous studies (Mason, Kubina, & Hoover, 2011, Mastropieri et al., in press; 2010; 2009) that have used the WJF as a pre- post-test measure. Two studies (Mason, Kubina, & Hoover, 2011; Mastropieri et al., 2010) reported the pre- and post-test raw mean score of the WJF. Mason, Kubina, and Hoover (2011) study contained three high school students with EBD who had a 22% increase in the raw score of the WJF. While Mastropieri et al. (2010) study contained eight eighth grade students with EBD who had 21% increase in the raw score of the WJF.

The other two other studies (Mastropieri et al., in press; 2009) reported the pre- and post-test standard mean score of the WJF. Mastropieri et al. (2009) reported 12 eighth grade students with EBD who had a 12% increase on the standard score of the WJF. While Mastropieri et al. (in press) study had 12 middle school students with ED who increased 19% on the standard score of the WJF.

Comparing findings of the current study with the four previous studies there are differences in the fluency subtest of the WJF pre- post-test outcomes. These differences could due to differences in reporting raw scores or standard scores. The current study's finding are in alignment with the previous two studies (Mastropieri et al., in press; 2009) that reported standard scores. Those studies reported WJF standard scores increased from 12 and 19%.

### **Post-Fluency Instruction Performance**

The overall post-fluency instruction performance for students had significantly improved from baseline testing to post-fluency instruction testing on all of the essay measures. In the present study, initial participants included seven middle

school students, six male and one female, with five students in seventh grade and two students in eighth grade. In addition, these students were ethnically and racially diverse and were diagnosed with EBD with a majority also having co-morbidity (e.g., having a specific learning disability in addition to emotional disability). During instruction students were taught in small groups (2-3 students) for 27 days (ranged from 22-32) and each session was approximately 40-minute over the course of approximately eight weeks of intensive SRSD instruction and fluency instruction. Over the course of instruction students improved on their essay performance by increasing the length of their essays in number of words written, writing more sentences, and paragraphs. They also improved on their persuasive essay components by having more transition words, including more persuasive essay parts, and having higher overall holistic quality scores. Students' growth from baseline to post-fluency represented large percent increases in the number of persuasive essay parts with an increase of 231%, in the holistic quality with an increase of 242%, and in the number of words written with an increase of 96% (see Table 5).

The present study extends and replicates five previous studies listed in Table 5 in several important ways. The present study taught single paragraph fluency lessons to mastery over multiple lessons and then taught students multiple paragraph essays. Previous studies have used a multiple baseline design to teach students with learning disabilities or EBD to write one paragraph persuasive essays fluently or using a quick write (QW) model through the SRSD model.

*Table 5.* Fluency percent of increase or decrease based on overall means for number of parts, number of words, and holistic quality from baseline to post-test, maintenance, and generalization.

Studies	<u>Post-Fluency or QW</u>			<u>Maintenance</u>			<u>Generalization</u>		
	Parts	Words	Quality	Parts	Words	Quality	Parts	Words	Quality
Mason, Kubina, & Taft 2011 study 1	116%	19%	66%	127%	22%	83%	-	-	-
Mason, Kubina, & Taft 2011 study 2	96%	46%	57%	93%	48%	62%	-	-	-
Mason et al. 2010	11%	-14%	88%	7%	-24%	83%	-	-	-
Mason, Kubina, & Hoover, 2011	61%	56%	66%	65%	76%	88%	-	-	-
Mastropieri et al. 2009*	209%	326%	162%	167%	263%	150%	109%	243%	96%
Mastropieri et al. in press*	208%	133%	156%	195%	59%	123%	192%	124%	132%
Cerar Dissertation	231%	96%	242%	151%	120%	160%	138%	97%	168%

*Note:* Data were used from Mason and Kubina (2011) to calculate post-multiple paragraph percent increases. Mason and Kubina (2011) citation were updated to match the studies publication citation. Maintenance and generalization data were retrieved from the studies publication.

\*compared to untimed baseline, since there is no timed baseline point

Previous studies have taught students with disabilities to write persuasive essays QW with significantly fewer sessions. Finally, a few studies taught middle school students with EBD to write one paragraph persuasive essay fluently after they first learned to write multiple paragraph persuasive essays using SRSD model.

Mason, Kubina, and Taft (2011) taught middle school students to write one paragraph persuasive essays fluently using QW procedures (note same as fluency in the present study) taught students with learning disabilities and other high incident disabilities. This study was composed of two multiple baseline experiments. In the first experiment, participants were six seventh grade students with four males and two females. Students were taught in pairs for five to six 45-minute sessions to write one paragraph persuasive essays QW. Overall, these students improved from baseline to post-QW instruction. These students' essays increased by 116% for the number of persuasive parts, increased by 66% for the holistic quality, and increased by 19% for the number of words. In the second study, participants were ten seventh and eighth grade students with four males and six females over five to six sessions. Overall, these students also improved from baseline to post-QW instruction. These students' essays increased by 96% for the number of persuasive essay parts, increased by 57% for the holistic quality, and increased by 46% for the number of words.

There are two previous studies that have investigated the use of SRSD model to teach how to QW a one paragraph persuasive essays with students with EBD. In the first study participants were five Caucasian middle school students in an inclusive middle school (four in seventh grade and one in the eighth grade) that were composed

of four males and one female (Mason et al., 2010). These students were composed of three students with EBD; two of those students had co morbid conditions with EBD. During instruction participants were taught one-on-one for five 30-minute lessons and five 10-minute lessons over the course of two to three weeks. Overall, these students improved from baseline to post-QW instruction on most essay measures. These students' essay improved on the number of persuasive essay with an increase of 11%, and on the holistic quality with an increase of 88%, however they decreased in the number of words in their essays with a decrease of 14%.

In the second study Mason, Kubina, and Hoover (2011) replicated and extended the Mason et al. (2010) study. As in the previous study, this study was a multiple baseline design that taught students one-on-one with five to seven 30-minute sessions for SRSD instruction with five 10-minute additional sessions for QW instruction. Participants in this study were three male high school students with one ninth-grade student and two eleventh grade students who had EBD with ADHD. These students' essays improved on the overall holistic quality with an increase of 66%, on the number of persuasive essay parts with an increase of 61%, and on the number of words with an increase of 56%.

Lastly, there were two studies that investigated the used the SRSD model to teach students with EBD to first write multiple paragraph persuasive essays followed by a second phase to teach them how to write one paragraph persuasive essays fluently. In the first, study Mastropieri et al. (2009) taught 12 eighth grade students (11 males) diagnosed with EBD with co-morbidity. During instruction students were



taught in four small groups for 55 sessions of 30-minutes. Overall, these students had drastically improved from baseline (untimed measure was used as the comparison due to no timed baseline measure) to post-fluency instruction. The students' essays improved on the number of words with an increase of 326%, on the number of persuasive essay parts with an increase of 209%, and on the holistic quality score with an increase of 162%.

In the second study, Mastropieri et al. (in press) replicated and extended the above study. The participants in this study were 12 seventh and eighth grade males' students with only ED. Instruction took place in four small groups for 15 sessions for the multiple paragraph instruction with an additional 3 sessions for fluency instruction with each session was 40-minutes. Overall, these students considerably improved from baseline (untimed measure was used as the comparison due to no timed baseline measure) to post-fluency instruction. The students' essays improved on the number of persuasive essay parts with an increase of 208%, on the holistic quality with an increase of 156%, and on the number of words with an increase of 133%.

Although all studies reported gains, there are differences in the amount of gains between the present study and the previous studies. These outcome differences could be accounted for by the sample differences or duration of the interventions, or both. The current study sample contained ethnically and racially diverse population of students with EBD while Mason, Kubina, and Taft (2011) had students with other disabilities than EBD. Furthermore, Mason, Kubina, and Hoover (2011) had students

with EBD but they were in high school, Mason et al. (2010) had Caucasian students with EBD with fewer co-morbid condition, and Mastropieri and colleagues (in press; 2009) had a similar sample to the current study with ethnically and racially diverse middle school students with EBD. The differences in the studies samples could explain the differences in the finding of the various studies.

The duration of the intervention varied across the studies and could be an additional factor in the different outcomes of the interventions. The studies by Mastropieri and colleagues (Mastropieri et al., in press; 2009) and the current study had longer intervention duration with higher percentage increases than the studies that had shorter durations by Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010). Students in the current study were unable to master the SRSD and fluency instructional content in ten sessions; therefore, longer training that was more intensive was needed for students to master the material before post-fluency instruction testing. This could possibly explain the outcome differences between the current study and the Mason and colleagues studies (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010).

There were other intervention variations between the current study and the previous study that could account for some of the outcome differences between the studies. One factor could be that the current sample appeared to be lower functioning than Mason et al. (2010), but due to the lack of background information on the participants in that study it cannot be definitely said. Furthermore, the current study's intervention was delivered in small groups over the course of 27 sessions compared to

the one-on-one instruction over the course of ten session by Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010). The increased instructional time could account for the outcome difference between these studies. This appears to indicate that although SRSD for persuasive writing for fluency can be taught to students with serious issues with EBD, instruction might needs to be very intensive, recursive, and repetitive, like the current study, and other previous studies by Mastropieri and colleagues (in press; 2010; 2009).

In the present study, it is also important to note that the post-fluency results included four incomplete post-fluency essays. During post-fluency testing four out of seven students were unable to finish one of their three essays within the 10-minute time frame. Three of those four students ran out of time and were unable to finish the ending on their essay, which gave them more persuasive essay parts at post-fluency than baseline. However, the students' holistic quality scores suffered for the incomplete essay, since the ending of the essay is considered a basic component (with topic, and three reasons) were unable to receive a score above a four. This illustrates even though these students mastered the fluency writing (two completed essays with all the parts within ten minutes) during instruction, they were still learning to self-regulate to complete tasks within a certain time frame. The fourth student was unable to complete his essay, since he was unable to come up with a third reason and explanation for his topic. This student did continue writing his essay without those parts to finish his essay, but he only had two reasons. This too affected his holistic quality score, since the basic components are three reasons. Yet, like the other three

student his essay also included more persuasive essay parts than at baseline. This illustrates the complexity of fluency writing given a restricted 10-minute period to plan and write their essays.

### **Post-Multiple Paragraph Instruction Performance**

The overall post-multiple paragraph instruction performance for students had significantly improved from baseline testing to post-fluency instruction testing on all of the essay measures. All of the essay measures significantly improved between the two testing phases. After fluency instruction and post-fluency testing the multiple paragraph instruction occurred for an average of 7 days (ranged from 4-8) were students' improved over the course of instruction on their essay structure and their persuasive essay components. Overall, for the essay structure students' improved on the length of their essay in number of words written, number of sentences, and number of paragraphs. They also improved on their persuasive essay components by using more transition words, using more persuasive essay parts, and having higher overall holistic quality scores. To illustrate the students' growth from baseline to post-multiple paragraph instruction students had large percent increases in the number of words written increased by 288% , in the number of persuasive essay parts increased by 227%, and in the overall holistic quality of increased by 218% (see Table 6).

The present study extends and replicates previously conducted research in several important ways with respect to multiple paragraph instruction. First, previous

*Table 6.* Percent of increase or decrease based for multiple paragraph studies on overall means for number of parts, number of words, and holistic quality from baseline to post-test in previous and current studies.

Studies	<u>Post-Multiple Paragraph</u>			<u>Maintenance</u>			<u>Generalization</u>		
	Parts	Words	Quality	Parts	Words	Quality	Parts	Words	Quality
Mastropieri et al. 2010	95%	179%	70%	67%	110%	46%	-	-	-
Mastropieri et al. 2009	198%	394%	153%	-	-	-	-	-	-
Mastropieri et al. in press	219%	246%	159%	227%	203%	175%	219%	203%	173%
Cerar Dissertation	227%	288%	218%	206%	283%	227%	199%	279%	119%

*Note:* Data were used from Mason and Kubina (2011) to calculate post-multiple paragraph percent increases. Mason and Kubina (2011) citation were updated to match the studies publication citation. Maintenance and generalization data were retrieved from the studies publication.

studies by Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) for one paragraph QW instruction has not had a second phase that extended fluency instruction to multiple paragraph instruction. However, previous studies by Mastropieri and colleagues (Mastropieri et al., in press; 2009) taught multiple paragraph instruction followed by a second instruction phase for fluency writing. In addition there was also one other study (Mastropieri et al., 2010) that only taught multiple paragraph instruction.

Mastropieri et al. (2009) used a multiple baseline design and taught multiple paragraph instruction for persuasive writing followed by a second fluency instruction phase using SRSD. The participants were 12 eighth grade students with EBD and all but one had a co-morbid condition that attended a public day school. Instruction for these students occurred in four small groups that were composed of 55 session that lasted 30-minutes each. The student growth overall was substantial from baseline to post-multiple paragraph instruction. These students essay improved on the number of word by an increase of 394%, on the number of persuasive essay parts by an increase of 198%, and on the overall holistic quality score by an increase of 153%.

Mastropieri et al. (in press) replicated and extended the Mastropieri et al. (2009) study in an inclusive middle school. The participants in this study were 12 seventh and eighth grade students with ED and/or co morbid conditions that attended an inclusive middle school. During instruction students were taught in four small groups for 15 sessions that lasted approximately 40-minutes each for multiple paragraph instruction. Overall student growth was extensive from baseline to post-

multiple paragraph instruction like the previous study. Students essay improved on the number of words with an increase of 246%, on the number of persuasive essay parts with an increase of 219%, and on the overall holistic quality with an increase of 159%.

In conclusion, multiple paragraph instruction for persuasive writing results in high percent increase in performance across the four studies with slight variations. The variations in the results can be due to study design, school setting, or intervention duration. The study design varied between the design experiment and multiple baseline design, which the multiple baseline design studies resulted in higher percent increases overall. The design also varied in the order of the instructional phases with the current study taught multiple paragraph instruction second while the Mastropieri and colleagues studies (in press; 2010; 2009) taught multiple paragraph instruction first. The current study resulted in highest percent increases in persuasive essay parts and holistic quality, while Mastropieri et al. (2009) resulted in the largest increase for number of words. However, these outcome differences could also be due to participant difference or the differences in intervention duration. It is important to note that all studies demonstrated increases in performance across written essay measures regardless of order of instruction or number of instructional sessions. This may be an important finding for teachers when deciding whether to introduce single or multiple paragraphs first.

## **Fluency Maintenance and Generalization Performances**

The overall fluency maintenance performance for students in the present study all improved substantially and significantly from baseline to maintenance testing on most fluency essay measures. Fluency maintenance testing occurred approximately a five weeks after the final post-multiple paragraph testing phase. Overall, students improved on the number of words written, number of sentences written, number of transitions words used, number of persuasive components written, and overall essay quality. However, there was not substantial improvement in the number of paragraphs composed. Since fluency instruction only taught students how to write a one paragraph persuasive essay, these findings were expected. The students' improvements were large in the percent increases that ranged from 120% to 160% for words, persuasive essay parts and holistic quality. These results are a little depressed from post-fluency testing, but still significantly better than baseline (see Table 5). These results replicate and extend previous research conducted by Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) and Mastropieri et al. (in press; 2009) in several ways.

In both this study and the Mason and colleague studies (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) taught students to write one paragraph essays using the POW+TREE strategy. In the present study, students received maintenance testing after approximately five weeks after post-testing while the Mason and colleagues studies students received maintenance testing after two to three weeks after post-testing. The Mason and colleagues studies found



percent increase for words, essay parts, and holistic quality ranged from -24% to 127%, with most under 90% (see table 5). Findings in the current study demonstrate higher performance than baseline after a longer delay in maintenance testing.

The present study and the Mastropieri and colleagues studies (in press; 2009) taught student to write both one paragraph essays and multiple paragraph essays using the POW+TREE strategy. In the Mastropieri et al. (2009) study students received maintenance testing 11.5 weeks after post-fluency testing and found percent increase for words, essay parts, and holistic quality ranged from 150% to 263%. Then in the Mastropieri et al. (in press) study students again received maintenance testing after post-fluency testing, but only after a two week delay. This study found that students' percent increases for words, essay parts, and holistic quality ranged from 59% to 195%. These findings are similar to the current study, but the outcomes vary along with the delay time between post-testing and maintenance testing.

In summary, the maintenance performance of student varies in two ways. First, maintenance testing has been assessed at different intervals from two to 11.5 weeks after the final post-testing was completed. Second, the results of the maintenance testing varied across studies. This could be due to the delay variation between the study and it could be due to the variation in the duration of the intervention itself. The present study indicated students retained a great deal of the learning over a five-week delay interval.

The overall fluency generalization performance for students in the present study all improved substantially and significantly from baseline to generalization

testing on most fluency essay measures. Generalization testing occurred immediately following maintenance testing five weeks after post-testing. Overall, students improved on the number of words written, number of transitions words used, number of persuasive components written, and overall essay quality. It was not surprising that there was not substantial improvement in the number of words or the number of paragraphs composed because students were taught to write a single paragraph only. Students' improvement was illustrated with large percent increase that ranged from 97% to 168% on the number of words, number of persuasive essay parts, and holistic quality. Again, these results are a little depressed from post-fluency testing and maintenance testing, but still significantly better than pre-testing.

These results replicate and extend previous research conducted Mastropieri et al. (in press; 2009) in several ways. The Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) studies did not include a generalization assessment of the students. In both the present study and the Mastropieri and colleagues studies (in press; 2009) taught student to write both one paragraph essays and multiple paragraph essays, but in reverse order of each other. In the Mastropieri et al. (2009) study students received generalization testing after maintenance testing at 11.5 weeks after post-testing. This study found students' essays improved with large percent increase that ranged from 96% to 243% for the number of words, the number of persuasive essay parts, and holistic quality (see Table 5). Then in the Mastropieri et al. (in press) study students again received generalization testing after maintenance testing after a delay of two week after post-

testing. This study found that students' essays were also large with percent increases that ranged from 124% to 192% for number of words, number of persuasive essay parts, and holistic quality. These findings are similar to the current study, but the outcomes vary along with the delay time between post-testing and generalization testing.

In summary, the generalization performance of student varied between the three studies. First, generalization testing has been assessed at varied intervals between two and 11.5 weeks after the final post-testing phase was completed. Second, the results of the generalization testing vary across the studies. This could be due to the delay variation between the study and it could be due to the variation in the participants or duration of the intervention in the studies. However, most importantly all studies reported marked improvements over baseline performances.

### **Multiple Paragraph Maintenance and Generalization Performance**

The overall multiple paragraph maintenance performance for students in the present study was substantially and significantly from baseline to both maintenance and generalization testing for multiple paragraph instruction on all essay measures. Multiple paragraph maintenance and generalization testing occurred after the fluency maintenance and generalization testing which was five weeks after the final post-multiple paragraph testing was completed. Overall, for the essay structure students' improved on the number of words written, number of sentences written, number of paragraphs written, number of transition words used, number of persuasive essay parts used, and overall holistic quality. Student growth was substantial and included

percent increase that ranged from 119% to 283% for the number of words, number of persuasive essay parts, and holistic quality. These results replicate and extend previous research conducted by Mastropieri et al. (in press; 2010) in a few ways.

Mastropieri and colleague (in press, 2010) studies taught students how to write multiple paragraph persuasive essays followed by a second phase that taught one paragraph persuasive essays fluently. After the final post-testing phase for fluency instruction was completed there was a delay of 3 months (Mastropieri et al., 2010) or 2 weeks (Mastropieri et al., in press) before maintenance testing occurred. Mastropieri et al. (in press) study also assessed for generalization after maintenance testing. The findings from the maintenance assessment from Mastropieri et al. (2010) study found student essay had increased moderately with percent increase that ranged from 46% to 110% for the number of words, number of persuasive essay parts, and holistic quality. While the findings from the maintenance and generalization assessment in Mastropieri et al. (in press) study had more substantial growth with students' essays improving with large percent increases that ranged from 173% to 227% for the number of words, number of persuasive essay parts, and holistic quality scores (see Table 6).

In conclusion, there are a limited number of studies that have taught students how to write multiple paragraph persuasive essays and have assessed maintenance and/or generalization of students' learning. The maintenance and generalization results vary from a minimum maintained growth of 46% to a maximum maintained growth of 283% across the studies. This variation in performances could be due to

the delay difference in the studies that range from two weeks to three months. It is also important to note that the intervention duration in these studies also varied from 15 days to 50 days, which could have impacted the students' ability to recall and generalize learning.

### **Self-Efficacy**

Overall the students made statistically significant gains on the self-efficacy measures from baseline to post-multiple paragraph test and from baseline to maintenance testing, but not from post-multiple paragraph to maintenance testing. These findings are interesting because this measure was not used previously. There is heuristic value in finding that students with EBD performed significantly better on this measure following treatment and retained that significance difference at maintenance testing. A change in self-efficacy for students with EBD is important and has the potential to impact other aspects of their life. Many previous researchers have not reported self-efficacy measure results (e.g. Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010; Mastropieri et al. in press; 2010; 2009). One exception is Mastropieri et al. (2012) in which a different self-efficacy measure was employed. Those findings were mixed in that some students with ED obtained significant self-efficacy gains while others did not. The present study offers positive preliminary results, but awaits future replication using this measure. Finally, since the sample size in the present study was small, replication of this measure is needed to validate the measure.

## **Social Validity**

Students' social validity revealed that the students learned the strategy by recalling the mnemonic; they knew how to plan for a persuasive essay, and felt the strategy helped their writing. Furthermore, the students found the strategy "easy", helped them be "organized", and "helped them write good persuasive essays." All the students mentioned the strategy could help other students like them and would help them in a similar manner that the strategy helped them. Half of the students reported they had used the strategy outside the study, while the other half said they could use the strategy in English class when they get to writing persuasive essays in class. These findings are similar to findings reported by Mastropieri and colleagues (in press; 2010; 2009).

## **On-Task Behavior**

Overall, the current study found a lower percentage (68%) of on task engagements than anticipated. This is less than desirable for students with EBD and can directly influence the amount of intervention time is required for the students to master instructional material. In this study 34 days of instruction was used in this intervention for students to master the materials in both instructional phases.

This study had one of the lower percentages of on-task behavior compared to three previous studies (Mastropieri et al., in press; 2010; 2009). These previous studies found that middle school students with EBD percentage of on-task behavior during their SRSD intervention had a mean of 62% - 94% (Mastropieri et al., in press; 2010; 2009) on-task behavior. One difference between the current study and the

previous studies is that the current study taught student to use the SRSD strategy fluently first. While the previous studies taught students to use the SRSD strategy to write multiple paragraph response and a few of the studies (Mastropieri, et al., in press; 2009) followed the multiple paragraph instruction with the fluency instruction. They maybe a contributing factor to the difference in on-task behavior. In addition, the time on-task from the current study was similar to the Mastropieri et al. (2009) 72% and Mastropieri et al. (2010) 62%. Both of the Mastropieri et al. (2010; 2009) studies were conducted at a specialized school for students with EBD, while the current study was conducted at an inclusive middle school. However, the student characteristics of the current study are closer to the student characteristics in the Mastropieri et al. (2009) study then to the student characteristics in Mastropieri et al. (in press). It seems severity of the EBD characteristic and lower academic functioning may be among the greater predictors of lower academic engagement.

### **General Discussion**

Overall findings from this study add significantly to the growing body of research on instructional strategies with students with serious emotional and behavioral disabilities (EBD) in several important ways. This study found that SRSD instruction using the POW+TREE mnemonic successfully improved persuasive writing skills of middle school students with serious EBD issues. Students wrote significantly better single paragraph and multiple paragraph essays following an average of 34 days of total instruction time across the two phases in the intervention study, followed by a five-week delayed testing phase that included both maintenance

and generalization measures for both instructional phases. Students also reported learning and seeing the benefits of the strategy.

These results replicated and extended the work of Mason and colleagues (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) by introducing the instruction of single paragraph essays within a ten-minute planning and writing framework. Mason and colleagues and this study reported positive immediate benefits of similar instruction. However, the present study extended the Mason and colleagues work (Mason, Kubina, & Hoover, 2011; Mason, Kubina, & Taft, 2011; Mason et al., 2010) in several ways. Mason implemented relatively brief instructional sessions (2 weeks), administered maintenance two weeks post instruction, but did not administer generalization measures. Maintenance findings were equivocal in the Mason studies, perhaps due to the reduced instruction time. The length of training was significantly longer and more intense in this study. Moreover, maintenance and generalization measures were administered approximately five weeks post a second instructional phase in this study intervention. Both maintenance and generalization findings were somewhat lower than immediate post intervention measures in this study, but substantially higher than baseline measures and students had higher percentage increases on essay measures than in the Mason studies. Finally, this study taught students to apply the strategy to writing multi paragraph essays following the 10-minute fluency writing. With the addition of that instructional phase, it was observed that students applied the single paragraph



essay to multi-paragraph essays. Again student performance differences were substantial from baseline to post intervention.

The present study also extended the work of Mastropieri and colleagues work with teaching persuasive writing to middle school students with significant EBD issues (Mastropieri et al., in press; 2010; 2009). Mastropieri and colleagues studies consistently introduced and taught multi-paragraph essays first and once students had mastered that, introduced the 10-minute single paragraph essay writing phase. Interestingly, both approaches have yielded similar positive benefits for middle school students with EBD.

### **Educational Implications**

Educational implications from this study add to the continuing literature on successful teaching of persuasive writing to middle school students with EBD. Taken together these findings indicate teachers can select to teach either single paragraph or multiple paragraph essays to students using the SRSD approach and obtain similarly positive findings. Moreover, instruction can be sequence such as that single or multiple paragraph instruction can be introduced either first or second with students performing optimally with either instructional sequence. Depending on the students initial writing abilities teacher may need extend intensive instruction as in the present study and Mastropieri et al., (2010; 2009).

### **Limitations**

Limitations of this study include: (a) small sample size, (b) attrition rate, and (c) instruction was limited to the SRSD strategy, POW+TREE. The final sample size

in this study was six students. Although this was a single-subject design and the present sample size is considered adequate any study with such a small sample size awaits replication. The present study lost three students to attrition. This was unfortunate but unavoidable. Students with EBD are a vulnerable population who are at risk for being suspended or expulsion from school and having serious mental health issues. In this study participants were lost due to both reasons, which limited the small size to six. Finally, a single strategy was taught in the present study. In reality the present sample requires a much more comprehensive written expression instruction program. For example, grammar, syntax, and multiple writing genres instruction would have benefited all participants.

### **Recommendations for Future Research**

Future research could address the limitations of this study. First, replication and extension of this study using a larger sample size would provide more validity to the present findings. An experimental design with a larger sample size using random assignment of participants to experimental and comparison conditions would allow greater generalization of findings. In addition a study containing a more comprehensive writing instruction program could address the serious writing need of students with serious EBD.

APPENDIX A

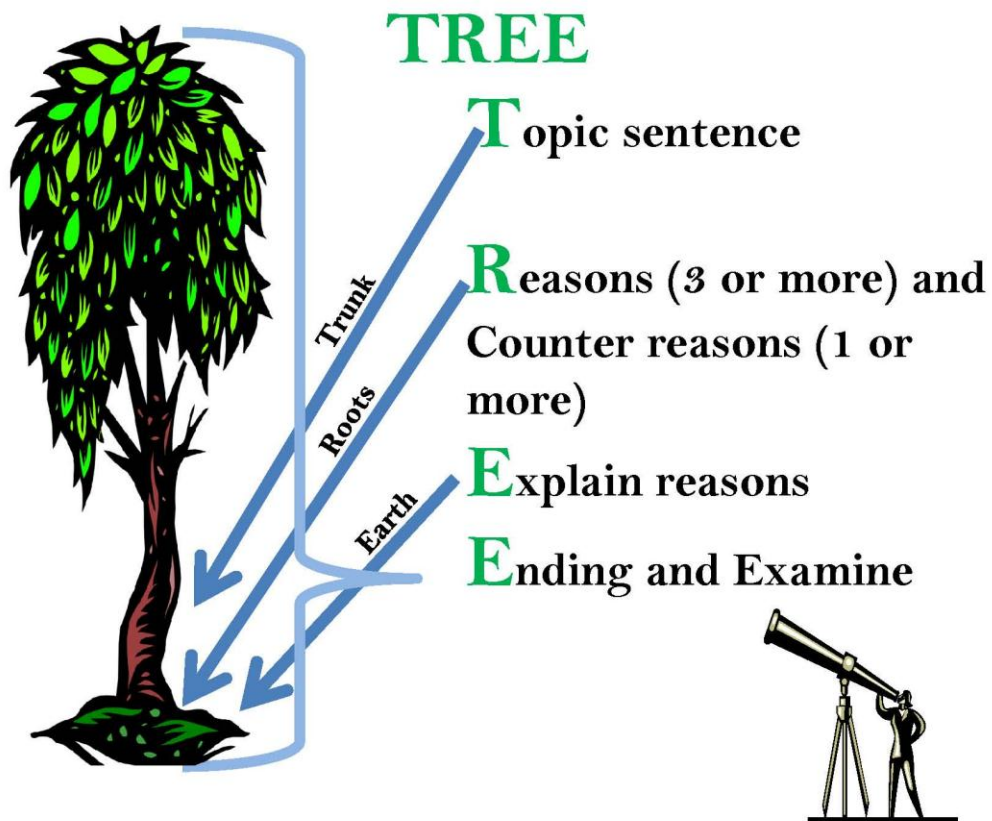
# POW

**P**ick my idea

**O**rganize my notes

**W**rite and say more

---



## APPENDIX B

# POW + TREE

**T** **TOPIC** Sentence  
What do I believe?

--

**TW** **R** **REASONS** -3 or more  
Why do I believe this?  
Will my readers believe this?

**E** **EXPLANATIONS**  
Say more about each reason.  
What details will persuade my reader?


**CR** **Counter Reason**- 1 or more.  
Who might disagree and why?

**E** **EXPLANATIONS**  
Say more about the counter reason(s).


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


**E** **ENDING**  
What do I want my reader to remember?

--	--

**EXAMINE** Check my paper again. Do I have all my parts? Yes \_\_\_\_\_ No \_\_\_\_\_  
Does each of my paragraphs have at least 3 sentences? Yes \_\_\_\_\_ No \_\_\_\_\_

## APPENDIX C

# Transition Words

**Words you can use to show a reason**

<b>First</b>	<b>Second</b>	<b>Third</b>	<b>In addition</b>
<b>Another</b>	<b>To begin</b>	<b>Also</b>	<b>Furthermore</b>
<b>Next</b>	<b>Finally</b>	<b>My final</b>	<b>Lastly</b>
_____	_____	_____	_____

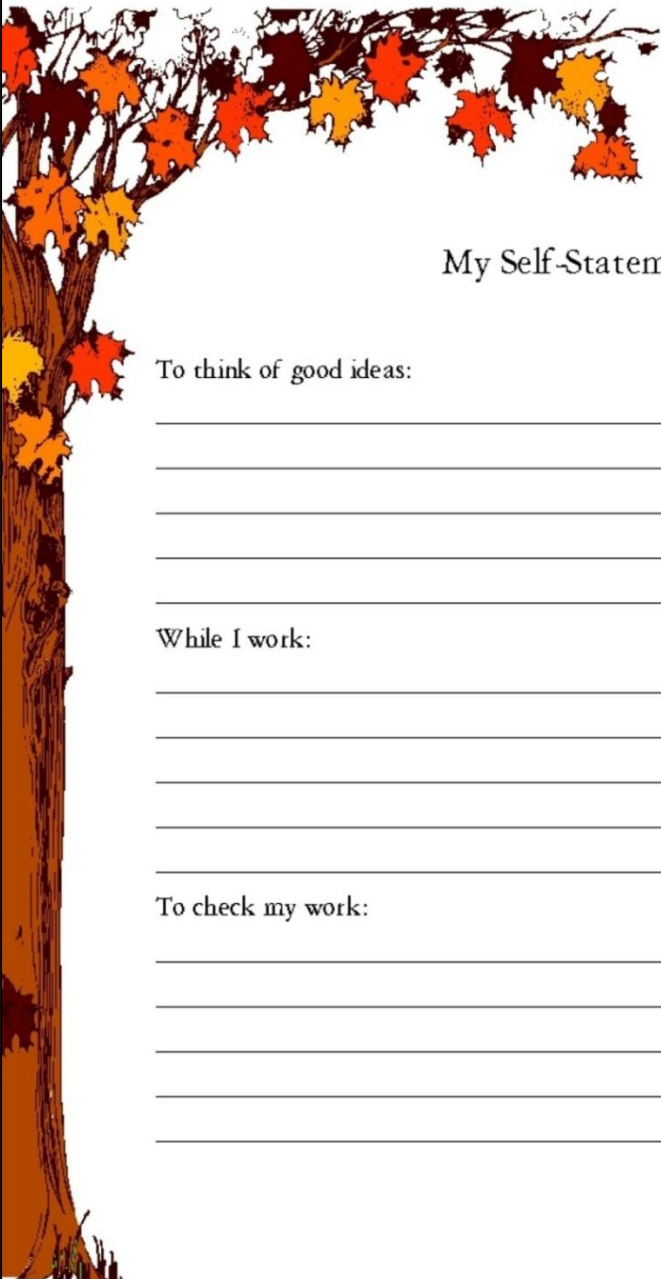
**Words you can use to show a counter reason**

<b>However</b>	<b>Nevertheless</b>	<b>Conversely</b>	<b>On the contrary</b>
<b>Yet</b>	<b>Instead</b>	<b>On the other hand</b>	
_____	_____	_____	

**Words you can use to conclude your essay**

<b>In conclusion</b>	<b>In short</b>	<b>On the whole</b>
<b>To summarize</b>	<b>In general</b>	<b>In sum</b>
_____	_____	_____

## APPENDIX D



My Self-Statements

To think of good ideas:

---

---

---

---

---

---

---

---

While I work:

---

---

---

---

---

---

---

---

To check my work:

---

---

---

---

---

---

---




---

## APPENDIX E

Name: \_\_\_\_\_

## POW+TREE Record Sheet + Effort

Use this sheet to track how you did. Look at the example on the first line.

1) Rate how well you did today -     
3 2 1

Goal: \_\_\_\_\_

[illegible]

# of parts: \_\_\_\_\_

## APPENDIX F

1

### Fluency Lesson 6

#### Objective

Students will:

- recognize instances where a one-paragraph written response is required;
- complete a 10-minute written response and
- Graph essay on the fluency sheet and set goal for next essay
- Students try to plan and writing essay in 10 minutes

#### Materials

POW+TREE mnemonic chart, POW+TREE graphic organizer, sample essay questions with sample responses, chart paper, markers, sample Student Record Sheet

#### Procedures

**I.** Introduce the idea that some writing tasks require a shorter response (i.e., not an entire essay), but that the response still needs to present good ideas and be well organized. The teacher could say something like the following:

**You have learned how to write good persuasive essays using POW+TREE. You are experts at it. Sometimes in school, you have to write, but you don't have to write an entire essay. What are some times in school that you might have to write a good paragraph, but not an entire essay?**

*(Allow time for students to respond. Possible examples include: essay questions on a test, answering questions from the social studies/science textbook for a homework assignment, in-class assignments, writing for a flier or brochure)*

**When you get these assignments, like writing an essay question on a test, you may not have the whole class period to do an essay. You may only have 10 minutes. So we are going to practice writing good paragraphs quickly.**

Example of an essay question:  
*Describe the three branches of government.*



<p>_____ <b>II.</b> Model writing a 10-minute prompt with think aloud. Include self-statements as you model. Solicit ideas from students so that you take up the entire 10 minutes. Model prompt: Should students your age have cell phones?</p> <ol style="list-style-type: none"> <li>1. Set the timer for 10 minutes.</li> <li>2. Post the prompt and read it out loud.</li> <li>3. Jot down ideas quickly – write “TREE” on paper and one word for each part, rather than phrases or sentences. Don’t forget about the counter reason and refute.</li> <li>4. Begin drafting on chart paper. Be sure to think aloud about transitions.</li> <li>5. If you finish before the timer goes off, say: <b>I still have more time so I can add another reason.</b> or <b>I still have time, so I can read it over and check my work.</b></li> <li>6. When timer goes off, put marker down.</li> <li>7. Discuss which parts were included, the order they are in the essay, and what was missing. Point out the use of transition words.</li> </ol> <p>_____ <b>IV.</b> Model filling in the individual Student Record Sheet. Make a goal of having more parts for the next quick write. Say:</p> <p><b>This time, I had _____ parts in my paragraph. Next time, I want to have _____ parts.</b></p> <p>_____ <b>V.</b> Now it’s your turn to plan and write your persuasive essay in 10 minutes. Quick-write</p> <p>Pass out papers face down, and pencils. Prompt: Should students your age be allowed to buy CD’s with warning labels that the</p>	<p>Prompt: Model prompt: Should students your age have cell phones?</p>
---	---

lyrics to songs may not be suitable for children? OR Should students recycle?

1. Read prompt aloud then have students flip over their papers as you start the time for 10 minutes.
2. Give two minute warning before timer goes off.
3. When timer goes off, ask students to put their pencils down. If students finish early, give them the full 10 minutes. Do not collect papers early.
4. Collect pencils and hand out pens.

**VI. Charting and goal setting.**

1. Students examine their papers and count number of parts.
2. Fill in the individual Student Record Sheet as a group.
3. Help students set realistic goal for next essay (e.g., two more parts). Write the goal on the chart above the column for the next day.

Lesson wrap-up.

**“Tomorrow we will review quickly. Then you will practice writing another paragraph in 10 minutes, and charting the number of parts.”**

## APPENDIX G

Lesson 1- POW+TREE Multiple Paragraph	Materials
<p><b>Purpose:</b> Modeling the entire process for writing a persuasive essay.</p> <p><b>Objectives:</b> The students will orally review the mnemonic for POW + TREE and state what each letter stands for. The students will attend to the teacher’s modeling lesson. The students will write self-statements for the POW + TREE writing strategy.</p> <p>_____ <b>A. Discuss the agenda for the day</b></p> <p>_____ <b>Today’s agenda is:</b></p> <ol style="list-style-type: none"> <li>1. Review POW+TREE, transition words, self-statements</li> <li>2. Observe the teacher model how to write an essay</li> <li>3. Write your own self-statements</li> <li>4. Graph the essay</li> </ol> <p>_____ <b>B. Briefly review POW+TREE, transition words, and self-statements (are things you say to yourself to help you to start or keep working; it's not always necessary to think out loud, you can think these in your head)</b></p> <p><b>C. Model the Strategy-</b> “Today I am going to model how to write an essay using all the materials we have been using. Pay attention to the self-statements that I use to talk to myself in positive ways.”</p> <ul style="list-style-type: none"> <li>• Read aloud the practice prompt: “Is it better to use EMAIL or SNAILMAIL?”</li> </ul> <p>_____ <b>I. Pick my Idea</b></p> <p>Lay out a copy of the TREE GRAPHIC ORGANIZER. Then explain:  <b>“What is it I have to do? I have to write a good opinion essay and good opinion essays tells the reader what you believe, give at least three reasons why, give an explanation for each reason , provide at least one counter reason with an explanation and refute the counter reason and have an ending sentence. My essay needs to makes sense and have all the parts. Remember P in POW - pick my idea - let my mind be free. (Pause) Take my time, what I</b></p>	<p><u>Materials:</u></p> <ol style="list-style-type: none"> <li>1. Mnemonic charts, GO chart,</li> <li>2. Transition words chart</li> <li>3. Self-statement chart</li> <li>4. Student record sheet poster and handouts</li> <li>5. Paper, pencils, and folders</li> </ol>

believe and good reasons why will come to me.” (Pause)

Dialogue: “I am going to model out loud writing an essay”

\_\_\_\_ “First, what do I believe - what do I want to tell the reader I believe?” (Now - talk out and figure out which side you are going to argue for). “I believe email is better than snail mail”  
“Good idea!”

## II. Organize my Notes

- The second letter in POW is O- ORGANIZE my NOTES. I will use POW + TREE to help me. I can write down ideas/notes for each part. But I don’t have to write ideas down in order, but I can jump around the GO as I think of ideas. This will help me organize my notes. (be sure to model moving out of order during your planning).
- You will do this too the next time you write an essay.
- Model the entire process for organizing your Notes **by completing the entire GO.** Use problem definition, planning, transition words, self-evaluation, and self-statements as you go. Follow the steps and statements below, filling in, ad lib statements where indicated. Ask the students to help you with ideas and the writing, but be sure you are in charge of the process:

Dialogue:

\_\_\_\_ “First, what do I believe - what do I want to tell the reader I believe?” (Now - talk out and fill in notes for Topic Sentence). “I believe email is better than snail mail” “Good I got the first thing done!”

\_\_\_\_ “Now I better figure out at least 3 reasons and give an explanation for each reason. Let my mind be free, think of good ideas.” (Now talk out and briefly write notes for at least 3 reasons- not in full sentences - use coping statements at least twice.)

“Hmmm, my essay would be even stronger if I think about other points of view. For example in this case: Who would disagree with me? Who might think snail mail is better than email? I know my grandmother! She might think snail mail is better b/c it is more practical and convenient for her to use regular mail b/c she doesn’t have access to a computer. Well, that is a good reason; however these days there are many public places that have access to computers.”

TW	Counter reason- snail mail is better Grandmother	Explanation	Refute it- Turn back
Some people might think	More practical and convenient	No access to a computer	But these days many public places have access to computers

After generating notes for all essay parts say – “Now I can look back at my notes and see if I can add more notes for my essay parts” (actually do this - model it - use coping statements). “I can also look for ideas for good word choice or million dollar words” (do this).

III. Write and Say More

A. Say, “Now I can do W in POW - write and say more. I will use the notes in the GO to help me write my essay. I can write my opinion essay and think of more good ideas or million dollar words as I write.” (Now - talk yourself through writing the essay; the students can help). Use a clean piece of paper and print. **ALSO, MODEL HOW TO USE THE CHECKLIST FOR WRITING**

**PARAGRAPHS** Start by saying:

### **DIALOGUE**

\_\_\_ **"How shall I start? I need to tell the reader what I believe, I need a topic sentence."** Then pause and think, then write out the sentence. So in P1 – You first write the topic sentence-This is where you ask, **"What am I writing about?"** Use your topic sentence to answer this question. Tell the reader what you believe, which is the main idea of your essay. Then on this paragraph I am only going to mention my reasons and I will use TW to introduce each reason. **MARK THE PARAGRAPH CHEKLIST AS YOU GO**

I believe email is better than snail mail. There are three reasons why I believe this. First Email is faster, second you can send attachments instantly, and third you save money.

\_\_\_ **"Good. Now I have to restate each reason I came up and explain it more."** So in P2 I will start my essay with a TW, then I will write my reason. To support my reason I will give an explanation or example to support this reason. **MARK THE PARAGRAPH CHEKLIST AS YOU GO**

First, email is faster than snail mail. When you send an email the other persons receives the email within minutes. You don't have to wait the couple of days it takes to receive a letter.

\_\_\_ **"Now, on P3 I have to write my next reason. Again I will start with a TW, then my reason and my explanation. MARK THE PARAGRAPH CHEKLIST AS YOU GO**

Next, in an email you can send attachments instantly. For example you can send pictures, documents, and even music and you don't have to wait for the other people to receive it. Also you don't have to pay extra to send these attachments.

\_\_\_ Good I am doing an excellent job. Now on P4 I will have to write my third reason. Again I will start with a TW, my reason, and explanation. **MARK THE PARAGRAPH CHEKLIST AS YOU GO**

In addition, you save money by using email. You don't have to spend money on gas driving to the post office. Also, you don't have to buy envelopes and stamps.

\_\_\_\_ Excellent I am doing an excellent job. I wrote my three reasons and explained each of them. Now, I better think about the counter argument. Who might have a different opinion than mine? I know-- my grandmother! She might think snail mail is better. On P5 I will begin with a TW for showing the opposite view and then I will state what other's position might be and then provide an explanation for that position.

On the contrary, some people might think snail mail is better than email because some people don't have access to a computer at home. (Good I wrote the counter reason and explained it, now I have to turn back and provide reasons to refute other people position) However, nowadays there are many public places like libraries where people can have access to computers.

#### **MARK THE PARAGRAPH CHEKLIST AS YOU GO**

\_\_\_\_ Now, on P6 I will write my ending or summary of the three reasons you have given. Restate your topic statement or belief and follow with the summary of your three reasons. Wrap it up!

In conclusion, I believe email is better than snail mail. It is faster, you save money and it is more practical. People should use email!

\_\_\_\_ Review by asking, "**Have I shown the reader all my reasons with explanations or examples?"**  
**Do I have all the parts? Do I have at least 3 sentences in each paragraph?"**

#### **D. Self-Statements**

\_\_\_\_ **I. Ask the students if they can remember:**

- 1) things you said to yourself to get started?
- 2) things you said while you worked,

3) things you said to yourself when you finished.  
Jot their ideas on the self-statement chart.

II. Ask the students to write some things they could say to themselves on their individual SELF-STATEMENT SHEET, using the chart as a reference.

- What to say to get started. This must be along same lines as "What is it I have to do? I have to write an opinion essay using TREE." – be sure students use their own words.
- Things to say while you work: self-evaluation, coping, self-reinforcement, and any others he/she likes (in students' own words).
- Things to say when you're finished (in students' own words).
- Note that we don't always have to think these things out loud; once we learn them we can think in our heads or whisper to ourselves.

#### **A. GRAPH THE ESSAY**

I. Model how to graph this essay on the Student Record Sheet. Ask students, **“Does this essay have at least 10 parts?”**

#### **B. Lesson wrap-up check off agenda**



## APPENDIX H

Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Persuasive Writing Self-Efficacy

Directions: Read the essay prompts below. Circle the one you select to write about. Then, answer the items below based on how confident you are on your ability to do what the question asks.

**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?**

	0% Confident	25% Confident	50% Confident	75% Confident	100% Confident
1. How sure are you that you can write an essay that contains at least 10 persuasive essay parts?	1	2	3	4	5
2. How sure are you that you can pick one side of the topic to argue in your essay?	1	2	3	4	5
3. How sure are you that you can make a plan before writing your essay?	1	2	3	4	5
4. How sure are you that you can write enough on the major points?	1	2	3	4	5
5. How sure are you that you can write an essay that is organized into good paragraphs?	1	2	3	4	5
6. How sure are you that you can plan an essay containing good reasons to persuade the reader?	1	2	3	4	5
7. How sure are you that you can write your position in your essay?	1	2	3	4	5
8. How sure are you that you can plan arguments for both sides of the topic for your essay?	1	2	3	4	5
9. How sure are you that you can write good transitional sentences in your essay?	1	2	3	4	5
10. How sure are you that you can use good self-statements while doing your essay?	1	2	3	4	5
11. How sure are you that you can examine and revise your essay?	1	2	3	4	5
12. How sure are you that you can plan an ending for your essay?	1	2	3	4	5
13. How sure are you that if you get stuck you can continue to work on your essay?	1	2	3	4	5

## APPENDIX I

### Social Validity Questionnaire

\* developed and used by Mastropieri et al. (in press; 2009)

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. (looking for POW+TREE and what each step means Be sure to prompt here with “can you tell me more” to ensure you obtain all student knows about the strategy -- remember we are also looking for counter arguments here now, too)
2. Draw a picture of the graphic organizer we used (ask student to label the parts or you write in labels if the GO is unclear)
3. What did you like most about this strategy?
4. Has using the POW+TREE strategy helped you become a better writer? How?
5. What did you learned when working with your writing teacher?
6. How do you think POW+TREE could help other students?
7. If you were the teacher, would you add anything to help students learn to write?
8. If you were the teacher, what would you change in the POW+TREE lessons? Why?
9. From the POW+TREE lessons, what things have most helped you become a better writer?
10. We used different writing time periods. One type allowed you as much time as you wanted to write an essay. The other type only allowed you ten minutes. Tell me which method you preferred and why.

11. Have you used POW+TREE in any other classes? If yes, ask, what other classes or assignments and how has it helped? (e.g., what class or classes? How did you do on those assignments? Better or worse than before?)
12. Tell me how you have used counter arguments in your writing. Why are counter arguments important?

## REFERENCES

## REFERENCES

- Adkins, S. H. (2005). *Self-regulated strategy development and generalization instruction: Effects of story writing among second and third grade students with emotional and behavioral disorders* (Doctoral dissertation). Retrieved from ProQuest. (AAT 3202407)
- Albertson, L. R., & Billingsley, F. F. (1997). *Improving young writers' planning and reviewing skills while story-writing*. Retrieved from ERIC database (ED406687)
- Baker, S., Chard, D., Ketterlin-Geller, L., Apichatabutra, C. & Doabler, C. (2009). Teaching writing to at-risk students: The quality of evidence for self-regulated strategy development. *Exceptional Children*, 75, 303-318.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp.307-337). United States of America: Information Age Publishing, Inc.
- Chard, D., Vaughn, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, 35, 386-406. doi:10.1177/00222194020350050101
- Code of Federal Regulations, Title 34, Section 300.7(b) (9).
- De La Paz, S. (1999). Self-regulated strategy instruction in regular education settings: Improving outcomes for students with and without learning disabilities. *Learning Disabilities Research & Practice*, 14, 92-106. doi:10.1207/sldrp1402\_3
- De la Paz, S. (2001). Teaching writing to students with attention deficit disorders and specific language impairment. *The Journal of Educational Research*, 95, 37-47. doi:10.1080/00220670109598781
- Donovan, M., & Cross, C. (2002). *Minority students in gifted and special education*. Washington, D.C.: National Academy Press.

- Gersten, R., & Baker, S. (2001). Teaching expressive writing to students with learning disabilities: A meta-analysis. *The Elementary School Journal*, 101, 251-272. doi:10.1086/499668
- Graham, S. (1990). The role of production factors in learning disabled students' composition. *Journal of Educational Psychology*, 82, 781-791. doi:10.1037/0022-0663.82.4.781
- Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. MacArthur, S., Graham, & J., Fitzgerald (Eds.), *Handbook of writing research* (pp. 187-207). New York: Guilford.
- Graham, S., & Harris, K. R. (1989). Improving learning disabled students' skills at composing essays: Self-instructional strategy training. *Exceptional Children*, 56, 201-214.
- Graham, S., & Harris, K. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 383-402). New York, NY: Guilford.
- Graham, S., Harris, K. R., MacArthur, C. A., & Schwartz S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disabilities Quarterly*, 14, 89-114. doi:10.2307/1510517
- Graham, S., Harris, K., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated writing development. *Contemporary Educational Psychology*, 30, 207-241. doi:10.1016/j.cedpsych.2004.08.001
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99, 445-476. doi:10.1037/0022-0663.99.3.445
- Guzel-Ozmen, R. (2006). The effectiveness of modified cognitive strategy instruction in writing with mildly mentally retarded Turkish students. *Exceptional Children*, 72, 281-297.
- Hammill, D., & Larsen, S. (1996). *Test of Written Language-3*. Austin, TX: Pro-ED.
- Harris, K. R., & Graham, S. (1999). Programmatic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly*, 22, 251-262. doi:10.2307/1511259

- Harris, K., Graham, S., & Mason, L. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal*, 43, 295-340. doi:10.3102/00028312043002295
- Harris, K., Graham, S., Mason, L., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore, MD: Brookes.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71, 165-179.
- Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, 58, 702-714. doi:10.1598/RT.58.8.1
- Kennedy, C. H. (2005). *Single-case designs for education research*. Boston, MA: Allyn and Bacon.
- Lane, K., Harris, K., Graham, S., Weisenback, J., Brindle, M., & Morphy, P. (2008). The effects of self-regulated strategy development on the writing performance of second grade students with behavioral and writing difficulties. *Journal of Special Education*, 41, 234-253. doi:10.1177/0022466907310370
- Little, M. A., Lane, K. L., Harris, K. R., Graham, S., Story, M., & Sandemel, K. (2010). Self-regulated strategies development for persuasive writing in tandem with schoolwide positive behavioral supports: Effects for second-grade students with behavioral and writing difficulties. *Behavioral Disorders*, 35, 157-179.
- Mason, L. H., Kubina, R. M., & Hoover, T. (2011). Effects of quick writing instruction for high school students with emotional disturbances. *Journal of Emotional and Behavioral Disorders*. Advance online publication. doi:10.1177/1063426611410429
- Mason, L. H., Kubina, R. M., & Taft, R. J. (2011). Developing quick writing skills of middle school students with disabilities. *Journal of Special Education*, 44, 205-220. doi:10.1177/0022466909350780
- Mason, L. H., Kubina, R. M., Valasa, L. L., & Cramer, A. M. (2010). Evaluating effective writing instruction for adolescent students in an emotional and behavior support setting. *Behavioral Disorders*, 35, 140-156.

- Mason, L., & Shriner, J. (2008). Self-regulated strategy development instruction for six elementary students with emotional behavioral disorders. *Reading and Writing: An Interdisciplinary Journal*, 21, 71-93. doi:10.1007/s11145-007-9065-y
- Mastropieri, M. A., Scruggs, T. E., Cerar, N. I., Bronaugh, D. A., Thompson, C., Guckert, M., Leins, P., Hauth, C., & Cuenca-Sanchez, Y. (in press). Fluent persuasive writing with students with emotional disturbance: Developing arguments and counterarguments. *Journal of Special Education*.
- Mastropieri, M. A., Scruggs, T. E., Cerar, N. I., Guckert, M., Thompson, C., Bronaugh, D. A., Jakulski, J., Abdulalim, L., Mills, S., Evmenova, A., & Cuenca-Sanchez, Y. (2012). *Classroom application of persuasive writing instruction with students with emotional and behavioral disabilities*. Manuscript submitted for publication.
- Mastropieri, M. A., Scruggs, T. E., Cuenca-Sanchez, Y., Irby, N., Mills, S., Mason, L. H., & Kubina, R. (2010). Persuading students with emotional disabilities to write: Findings from a design experiment. In T.E. Scruggs & M. A. Mastropieri (Eds.), *Literacy and learning: Advances in learning and behavioral disabilities: Vol. 23* (pp. 237-268). Oxford, UK: Emerald.
- Mastropieri, M. A., Scruggs, T. E., Mills, S., Irby, N., Cuenca-Sanchez, Y., Bronaugh, D. A., Thompson, C., Guckert, M., & Regan, K. (2009). Teaching students with emotional disabilities to write fluently. *Behavioral Disorders*, 35, 19-40.
- Mooney, P., Epstein, M. H., Reid, R. & Nelson, R. (2003). Status and trends in academic intervention research for students with emotional disturbance. *Remedial and Special Education*, 24, 273-287. doi:10.1177/07419325030240050301
- Mooney, P., Ryan, J.B., Uhing, B. M., Reid, R., & Epstein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education*, 14, 203-221. doi:10.1007/s10864-005-6298-1
- Morgan, P. & Sideridis, G. (2006). Contrasting the effectiveness of fluency interventions for students with or at risk for learning disabilities: A multilevel random coefficient modeling meta-analysis. *Learning Disabilities Research & Practice*, 21, 191-210. doi:10.1111/j.1540-5826.2006.00218.x.



- Nelson, R., Benner, G., Lane, K., & Smith, B. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children, 71*, 7159-7173.
- Oakes, W. P., Mathur, S. R., & Lane, K. L. (2010). Reading interventions for students with challenging behavior: A focus on fluency. *Behavioral Disorders, 35*, 120-139.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly, 19*, 139-158.  
doi:10.1080/1057356039014308
- Pajares, F., & Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 159-170). New York, NY: The Guildford Press.
- Reid, R., Gonzalez, J., Nordness, P., Trout, A., & Epstein, M. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38*, 130-143.  
doi:10.1177/00224669040380030101
- Rogers, L.A. & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*, 879-906.  
doi:10.1037/0022-0663.100.4.879
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition*. Itasca, IL: Riverside Publishing.
- Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Assessment Service Bulletin Number 2: WJ III technical abstract*. Itasca, IL: Riverside Publishing.
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial and Special Education, 24*, 252-261.  
doi:10.1177/07419325040250040801
- Trout, A., Nordness, P., Pierce, C., & Epstein, M. (2003). Research on the academic status of children with emotional and behavioral disorders: A review of the literature from 1961 to 2000. *Journal of Emotional and Behavioral Disorders, 11*, 198-210. doi:10.1177/10634266030110040201

- U.S. Department of Education. (2009). *Twenty-eighth annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2006* (vol.1) (Contract No. ED01CO0082/0008). Washington, DC: Author.  
Retrieved from <http://www2.ed.gov/about/reports/annual/osep/2006/parts-b-c/28th-vol-1.pdf>
- Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M.H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13*, 79-96.  
doi:10.1177/10634266050130020201
- Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., & Kaufman, M. J. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders, 4*, 194-209. doi:10.1177/106342669600400401
- Walker, H. M., & Severson, H. (2001). Developmental prevention of at-risk outcomes for vulnerable antisocial children and youth. In K. L. Lane, F. M. Gresham, & T. E. O'Shaughnessy (Eds.), *Interventions for children with or at risk for emotional and behavioral disorders* (pp. 177-194). Boston: Allyn & Bacon.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children (4<sup>th</sup> ed.)*. San Antonio, TX: The Psychological Corporation.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001a). Examiner's manual. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001b). *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

## CURRICULUM VITAE

Nancy Irby Cerar is a 2001 graduate of Park View High School. She received her Bachelor of Science degree in Psychology from George Mason University in 2005. She worked closely with Dr. Patricia Moyer-Packenham and Dr. Anastasia Kitsantas on a NSF grant while completing her Master of Education in Educational Psychology from George Mason University in 2007. Then she worked closely with Dr. Margo A. Mastropieri on IES grants while completing her Doctorate of Philosophy in Education from George Mason University in 2012.