THE EFFECTIVENESS OF A GOAL SETTING INTERVENTION THAT INCORPORATES PERFORMANCE FEEDBACK, SELF-GRAPHING, AND REINFORCEMENT ON IMPROVING THE WRITING SKILLS OF HIGH SCHOOL STUDENTS

A dissertation submitted to the Kent State University College of Education, Health, and Human Services in partial fulfillment of the requirements for the degree of Doctor of Philosophy

By

Katherine Bobak Lavik

May 2014
A dissertation written by

Katherine Bobak Lavik

B. A., Case Western Reserve University, 2008

M. Ed., Kent State University, 2009

Ph. D., Kent State University, 2014

Approved by

__________________________________, Director, Doctoral Dissertation Committee
Frank J. Sansosti

__________________________________, Member, Doctoral Dissertation Committee
Richard J. Cowan

__________________________________, Member, Doctoral Dissertation Committee
Andrew L. Wiley

Accepted by

__________________________________, Director, School of Lifespan Development and Educational Sciences
Mary Dellmann-Jenkins

__________________________________, Dean, College and Graduate School of Education, Health, and Human Sciences
Daniel F. Mahony
The present study addressed gaps in the current literature on writing interventions by conducting a goal setting writing intervention with high school students and by utilizing correct writing sequences (CWS) as the outcome measure. The intervention included performance feedback, self-graphing, and reinforcement as supporting components. The intervention was conducted in a one-on-one setting with six high school students and utilized changing criterion single-subject research methods. CWS were used to evaluate the writing samples and to set goals. Participant writing samples were also graded for total words written, correct minus incorrect writing sequences, and correct punctuation marks in order to further discussion of the utility of these curriculum-based measures at the high school level. Descriptive and visual analysis and effect size data revealed positive effects for five of the six participants and mixed results for the sixth participant. Effect sizes were reported as percent nonoverlapping data (PND) and nonoverlap of all pairs (NAP). Treatment integrity data revealed high adherence to intervention protocol. The study suggests several important findings, in spite of some limitations. First, the study supports goal setting as an effective intervention for improving high school students’ writing. Second, participants found the treatment acceptable, supporting its social validity. Limited acceptability data was available from
the participants’ teachers, however. Third, generalization and maintenance data were generally positive, with some exceptions. Fourth, the utility of various curriculum-based measures for high school students and of PND and NAP in single-subject research are discussed. Additional implications for research and practice are also included.
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all of the family, friends, colleagues, and professors who have supported and encouraged me during my studies. First, I would like to thank my dissertation committee. A sincere thank you goes to Dr. Frank Sansosti for all his edits, support, encouragement, and time spent reviewing and revising my dissertation. Thank you to Dr. Richard Cowan for support and for feedback on many revisions of this dissertation, as well. I also thank Dr. Andrew Wiley for bringing his perceptions of the study as someone outside of the field of school psychology. Additionally, I thank Drs. Karla Anhalt and Caven McLoughlin, who, while not on my dissertation committee, also deserve recognition for the research experiences and training they provided me.

Second, I would like to thank those who assisted me in setting up, conducting, and editing my research and writings. Thank you to my participants—without you six, I would not have had a study! Thank you to the teachers and district personnel that worked with me on finding participants and space in which to conduct the study. I thank Alicia, Lisa, and Erin for their help in data collection. I also thank my husband, mom, and brothers for providing editing eyes to multiple versions of my dissertation.

Third, I would like to thank my internship supervisors, Dr. Bradd Falkenberg, Mrs. Aimee Kirsch, Mr. Jeff York, Dr. Rob Kubick, and Dr. Perry Clark, who helped prepare me well for employment as a school psychologist. I would also like to thank my practicum, advanced practicum, and PSI supervisors for providing me with diverse
learning and practical experiences. Thank you to my co-workers in Akron for their support, as well, during my first year of employment.

Last, but most importantly, I would like to thank my family. Thank you to my mom, dad, and stepmom for instilling in me an insatiable need to achieve and perform to the best of my abilities. Thank you to my brothers and sister for supporting me with love and friendship. Thank you to my in-laws for also being great supports. Thank you to my husband, Andrew, for his constant encouragement, support, and love through graduate school and beyond.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>The Importance of Writing</td>
<td>1</td>
</tr>
<tr>
<td>Interventions to Assist Struggling Writers in High School Settings</td>
<td>5</td>
</tr>
<tr>
<td>Strategy instruction</td>
<td>6</td>
</tr>
<tr>
<td>Summarization</td>
<td>6</td>
</tr>
<tr>
<td>Peer-assisted writing</td>
<td>7</td>
</tr>
<tr>
<td>Goal setting</td>
<td>7</td>
</tr>
<tr>
<td>Handwriting, spelling, or typing</td>
<td>7</td>
</tr>
<tr>
<td>Word processing</td>
<td>8</td>
</tr>
<tr>
<td>Sentence combining</td>
<td>8</td>
</tr>
<tr>
<td>Inquiry</td>
<td>9</td>
</tr>
<tr>
<td>Prewriting</td>
<td>9</td>
</tr>
<tr>
<td>Process writing</td>
<td>9</td>
</tr>
<tr>
<td>Writing-to-learn</td>
<td>10</td>
</tr>
<tr>
<td>Model study</td>
<td>10</td>
</tr>
<tr>
<td>More on Goal Setting</td>
<td>11</td>
</tr>
<tr>
<td>Performance feedback</td>
<td>14</td>
</tr>
<tr>
<td>Self-graphing</td>
<td>15</td>
</tr>
<tr>
<td>Reinforcement</td>
<td>16</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>16</td>
</tr>
<tr>
<td>Research Questions</td>
<td>19</td>
</tr>
<tr>
<td><strong>II. LITERATURE REVIEW</strong></td>
<td>20</td>
</tr>
<tr>
<td>Background</td>
<td>20</td>
</tr>
<tr>
<td>Assessing Writing Skills</td>
<td>20</td>
</tr>
<tr>
<td>What is Measured?</td>
<td>20</td>
</tr>
<tr>
<td>Curriculum-Based Measurement and Assessment</td>
<td>23</td>
</tr>
<tr>
<td>CBM, fluency, and improved academic performance</td>
<td>26</td>
</tr>
<tr>
<td>Curriculum-based measurement in writing</td>
<td>26</td>
</tr>
<tr>
<td>Scoring curriculum-based assessment measures with high school students</td>
<td>28</td>
</tr>
</tbody>
</table>
Future practice and research using writing CBA with high school students. ................................................................. 31
Evidence-Based Practices in Writing Intervention ........................................................................................................... 34
When Is an Intervention Considered Evidence-Based? ...................................................................................................... 34
Summary. ............................................................................................................................................................................ 39
Evidence-Based Practices for Students in Need of Individualized Interventions ............................................................... 39
A call for more single-subject research. ............................................................................................................................. 41
Writing Intervention Research with High School Students ................................................................................................. 42
Meta-Analyses of Writing Intervention Research ........................................................................................................... 43
Strategy instruction. ............................................................................................................................................................ 62
Self-regulated strategy development. ........................................... 68
Editing .............................................................................................................................................................................. 69
Paragraph construction............................................................................................................................................................ 70
Planning/drafting ................................................................................................................................................................. 71
Direct instruction: Expressive Writing. ........................................... 72
Goal setting ............................................................................................................................................................................ 73
Sentence construction .............................................................................................................................................................. 74
Reinforcement ......................................................................................................................................................................... 74
Prewriting activities ............................................................................................................................................................... 75
Critique and Analysis of the Research ........................................................................................................................................ 75
Effective studies ........................................................................................................................................................................ 76
Ineffective studies ....................................................................................................................................................................... 78
Effectiveness by intervention type ...................................................................................................................................... 78
Overall Recommendations ......................................................................................................................................................... 79
Future Directions for Writing Interventions for High School Students .................................................................................. 83
Summary ................................................................................................................................................................................ 90
Conclusion .................................................................................................................................................................................. 91

III. METHODOLOGY .......................................................................................................................................................... 94
Method ................................................................................................................................................................................... 94
Setting ................................................................................................................................................................................... 94
Participants ............................................................................................................................................................................ 97
Researcher and data collectors ................................................................. 100
Materials ................................................................................................................................................................................ 100
Dependent Variables .............................................................................................................................................................. 101
Curriculum-based assessments .......................................................................................................................................... 101
Primary outcome measure and goal setting measure ...................................................................................................... 101
Additional outcome measures ........................................................................................................................................ 101
Generalization measures .................................................................................................................................................. 103
Classroom performance ............................................................................................................................................. 103
Participant samples .............................................................................................................................................................. 104
Cover letter sample ............................................................................................................................................................. 104
Independent Variable .......................................................................................................................................................... 104
Data Collection ........................................................................................................ 105
Inter-rater reliability .............................................................................................. 105
Additional writing probe administration .............................................................. 106
Experimental Design ............................................................................................. 107
Procedure ................................................................................................................. 108
Participant recruitment .......................................................................................... 108
   Inclusion criteria .................................................................................................. 109
   Exclusion criterion ............................................................................................. 110
   Response rate ...................................................................................................... 110
   Teacher consent ................................................................................................. 111
   Institutional research approval ......................................................................... 111
Baseline .................................................................................................................. 111
Introductory session ............................................................................................... 116
   Performance feedback ....................................................................................... 117
   Self-graphing ...................................................................................................... 118
   Goal setting ......................................................................................................... 119
      End-of-intervention goal ............................................................................... 119
      Initial goal ........................................................................................................ 119
      Remaining goals ........................................................................................... 120
   Reinforcement ...................................................................................................... 123
   Generalization prompt ....................................................................................... 124
Intervention .............................................................................................................. 125
   Maintenance and generalization ....................................................................... 126
      Maintenance .................................................................................................... 126
      Generalization ................................................................................................ 127
Treatment Integrity and Interobserver Agreement ................................................. 128
Social Validity .......................................................................................................... 129
Data Analysis ........................................................................................................... 131
   Research Questions ........................................................................................... 131
   Generalization ..................................................................................................... 137
Other data ................................................................................................................ 138

IV. RESULTS ........................................................................................................... 139
Participants’ Improvement in Written Expression .................................................. 139
   Variability ............................................................................................................ 143
   Magnitude and Immediacy of Change ................................................................ 146
      Criterion changes ........................................................................................... 147
      End-of-intervention goal ............................................................................... 151
   Rate of Change .................................................................................................... 152
      Trend lines ......................................................................................................... 152
      Rate of improvement ....................................................................................... 158
   Overall Pattern of Change ................................................................................... 159
      PND ...................................................................................................................... 160
      NAP ..................................................................................................................... 160
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Kyrie’s Results</td>
<td>140</td>
</tr>
<tr>
<td>2. Alonzo’s Results</td>
<td>140</td>
</tr>
<tr>
<td>3. Omri’s Results</td>
<td>141</td>
</tr>
<tr>
<td>4. Anderson’s Results</td>
<td>141</td>
</tr>
<tr>
<td>5. Tyler’s Results</td>
<td>142</td>
</tr>
<tr>
<td>6. CJ’s Results</td>
<td>142</td>
</tr>
<tr>
<td>7. Kyrie’s Trends</td>
<td>153</td>
</tr>
<tr>
<td>8. Alonzo’s Trends</td>
<td>154</td>
</tr>
<tr>
<td>9. Omri’s Trends</td>
<td>155</td>
</tr>
<tr>
<td>10. Anderson’s Trends</td>
<td>156</td>
</tr>
<tr>
<td>11. Tyler’s Trends</td>
<td>157</td>
</tr>
<tr>
<td>12. CJ’s Trends</td>
<td>158</td>
</tr>
<tr>
<td>13. Kyrie’s Other Data</td>
<td>172</td>
</tr>
<tr>
<td>14. Alonzo’s Other Data</td>
<td>173</td>
</tr>
<tr>
<td>15. Omri’s Other Data</td>
<td>173</td>
</tr>
<tr>
<td>16. Anderson’s Other Data</td>
<td>174</td>
</tr>
<tr>
<td>17. Tyler’s Other Data</td>
<td>174</td>
</tr>
<tr>
<td>18. CJ’s Other Data</td>
<td>175</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1. Example of Scoring Correct Writing Sequences (CWS) in Curriculum-Based Assessment of Writing</td>
<td>29</td>
</tr>
<tr>
<td>2. Effective Writing Interventions as Identified by Rogers and Graham (2008)</td>
<td>53</td>
</tr>
<tr>
<td>3. Number of Writing Intervention Studies Conducted with High School Students</td>
<td>55</td>
</tr>
<tr>
<td>4. Quality Indicators for Single-Subject Research</td>
<td>61</td>
</tr>
<tr>
<td>5. Demographic Information of Single-Subject Writing Interventions with Secondary Students</td>
<td>63</td>
</tr>
<tr>
<td>6. Demographic and Basic Information for Single-Subject Writing Interventions with Secondary Students</td>
<td>65</td>
</tr>
<tr>
<td>7. Six Stages of Self-Regulated Strategy Development</td>
<td>68</td>
</tr>
<tr>
<td>8. Intervention Rankings</td>
<td>77</td>
</tr>
<tr>
<td>9. State and District Demographics</td>
<td>95</td>
</tr>
<tr>
<td>10. Participant Demographics</td>
<td>97</td>
</tr>
<tr>
<td>11. Teacher Demographics</td>
<td>112</td>
</tr>
<tr>
<td>12. Student Probe Sequences</td>
<td>114</td>
</tr>
<tr>
<td>13. Stability in Baseline Phase</td>
<td>116</td>
</tr>
<tr>
<td>14. Variability in the Baseline Phase</td>
<td>143</td>
</tr>
<tr>
<td>15. Variability in Intervention Criterion Subphases</td>
<td>144</td>
</tr>
<tr>
<td>16. Mean and Median of Baseline and Intervention Phases</td>
<td>146</td>
</tr>
<tr>
<td>17. Comparison of CWS Percentiles from Winter Benchmark to Intervention</td>
<td>152</td>
</tr>
</tbody>
</table>
18. Rate of Improvement of Correct Writing Sequences from Baseline to Intervention .................................................................159

19. Ratings for Each Item on the Social Validity Scale for Students ..................163

20. Variability in the Maintenance Phase ...........................................................164

21. Mean and Median of Baseline and Maintenance Phases ............................165

22. Rate of Improvement of Correct Writing Sequences from Baseline to Maintenance .................................................................................................................................167

23. Data Summary ................................................................................................177
CHAPTER I

INTRODUCTION

Background

The Importance of Writing

Writing has been an essential component of everyday life ever since it was invented. If someone possesses some knowledge to pass on to others, one of the best ways to make sure that everyone receives the same message for generations to come is to write it down so it can later be read and understood. For example, Egyptian kings and queens wanted future generations to be aware of their accomplishments, their greatness, and their divine heritage (Brown, 2009). To accomplish this goal, they lined their temples, palaces, and burial chambers with hieroglyphs that described what they did, how well they did it, and how they had descended from the gods. These writings were considered so important and so influential that rival pharaohs would scratch out important parts of the writings of their predecessors after they had passed away so that their own stories and accomplishments would be told and remembered instead (Brown, 2009).

Once we developed the ability to read and write, we instantly expanded our capacity for knowledge and our ability to innovate by building upon previous knowledge. As such, the ability to write is as important now as it has been historically. Education depends on the written word (ACT, 2011; Patterson & Duer, 2006). Employment depends on the written word (National Commission on Writing, 2004). The acquisition and expansion of knowledge depends on the written word. Unfortunately, research in the field of education has focused much more on the ability to read the written word than the
ability to write (National Commission on Writing, 2004). Within schools, students
generally learn to read and write at the same time, but the primary focus of education and
research has been on reading. Like reading, proficient writing performance is an
important life skill for high school students as they prepare to enter college and/or the
working world. Unfortunately, many adolescents do not attain proficiency in writing
(Graham & Perin, 2007a; National Center for Education Statistics, 2012).

In 2011, the National Center for Education Statistics’ (NCES) National Assessment of Educational Progress (NAEP) administered writing assessments to students across the United States (NCES, 2012). Students received a narrative, informative, or persuasive writing task. In 2011, the NAEP writing assessment was given on the computer for the first time (it had previously been a paper and pencil test). Prompts were presented as text with accompanying audio, photographs, video, and/or animation. Students composed their responses on laptop computers and had 30 minutes to write. They were allowed standard editing (e.g., spell check, thesaurus, cut, paste, delete), formatting (e.g., bold, italics), and text viewing (e.g., collapse prompt, collapse response) functions. Subsequently, the students’ writings were graded using a rubric and were rated as basic (“partial mastery of the prerequisite knowledge and skills that are fundamental for proficient work at each grade;” NCES, 2012, p. 2), proficient (solid academic performance), or advanced (superior performance). Overall, 54% of eighth grade students achieved the basic level of proficiency while 20% achieved below this level. In grade 12, 52% of students achieved the basic level of proficiency while 21% achieved below the basic level. These statistics imply that 74% of eighth grade students
and 73% of 12th grade students performed below the proficient level. Moreover, only 3% of eighth grade students and 3% of 12th grade students achieved the advanced level of writing performance. The NAEP reports have noted that some states may define their assessment standards differently. For example, some states may define proficient as the “standard for promotion to the next grade,” rather than using the NAEP’s definition of proficient as competency and solid performance (Salahu-Din, Persky, & Miller, 2008, p. 6). Despite such a difference in the definition of “proficient,” high school students should strive to meet both given definitions of proficiency. By passing on to the next grade through solid academic performance and becoming competent in writing, students likely will possess the skills necessary to be successful in college and/or the workplace.

In addition to the NAEP study, there is other similar evidence that high school students do not possess a basic aptitude in writing. ACT, Inc., in their report on The Condition of College & Career Readiness (ACT, 2011), stated that only 66% of ACT-tested high school graduates met the ACT’s benchmarks for college readiness in freshman English composition courses. These empirically-based benchmarks represent the minimum ACT subject test scores needed to predict about a 75% chance of earning a C or higher in the first-year college course for that subject. These national assessments reveal writing as an academic area in which high school students are not meeting the proficiency standards necessary in a competitive global environment.

Poor writing ability not only affects school performance—it also impacts performance in the workplace, where the ability to write is a valued skill (National Commission on Writing, 2004). An employee’s (or prospective employee’s) ability to
write is often considered when making hiring and promotion decisions (National Commission on Writing, 2004). Employers expect their employees to have the ability to express themselves appropriately through writing, but some businesses continue to need to invest in remedial writing programs for their employees. In fact, the National Commission on Writing (NCOW) conducted a survey of 120 major American corporations and asked about the importance of their employees’ writing skills and the cost of remediating their employees’ writing deficiencies (NCOW, 2004). The companies surveyed employed nearly 8 million people and reported that as much as 3.1 billion dollars are spent annually on writing remediation. The time, money, and productivity lost when employees have to be retrained to write can amount to substantial losses for companies. Altogether, the ability to write proficiently is an integral part of society, and being unable to write proficiently can hinder a person’s achievement. Although nearly 70% of major companies in the United States described writing as an essential responsibility of professional, salaried positions in their companies, they reported that only 33% of their new hires possessed the writing skills their companies valued (NCOW, 2004).

Writing is a skill that is important throughout life. Proficiency in writing is necessary in primary, secondary, and postsecondary education for multiple reasons: (a) through writing, students can demonstrate to the teacher that they have learned and understand the concept that was taught; (b) businesses, companies, and other work places use written products to evaluate an applicant’s prospective ability to perform a job; (c) work places use written products when making promotional decisions; and (d) socially,
letters, email, texts, social media, and other methods of correspondence are used to communicate, express and share ideas, and forge and maintain relationships. Unfortunately, contemporary evidence suggests that many high school students lack proficiency in writing, leading to decreased academic performance in school (as well as in the workplace). Such limited proficiency in writing not only negatively impacts performance in school and/or at work, but also indicates the need for better instruction and intervention in writing for high school students before they seek employment and/or apply to colleges. High schools specifically should be focused on improving writing because students there are in transitional years, getting ready either to go to college or enter the workforce. High schools should investigate and implement evidence-based strategies for helping their students to improve their writing skills. Despite the disheartening statistics pertaining to the writing proficiency of students, a host of interventions have been demonstrated to be successful at improving writing skills and curtailing the long-term negative outcomes associated with poor writing proficiency.

**Interventions to Assist Struggling Writers in High School Settings**

While academic research has not focused as much on writing as on reading and math, there have been some studies to suggest effective methods for improving the writing skills of high school students. In their meta-analysis on writing interventions, Rogers and Graham (2008) list in their introduction the 12 writing interventions that are the best supported by current research. They are, in order of magnitude of effectiveness: (a) strategy instruction, (b) summarization, (c) peer assistance, (d) goal setting, (e)
handwriting/spelling/typing instruction, (f) word processing, (g) sentence combining, (h) inquiry, (i) prewriting, (j) process writing, (k) writing-to-learn, and (l) model study.

**Strategy instruction.** Strategy instruction interventions focus on teaching students specific strategies for different aspects of writing, such as planning, writing, revising, and editing. These strategies typically involve multiple steps that are followed and memorized using mnemonic devices, such as STOP: suspend judgment, take a side, organize ideas, and plan more as you write (Graham, 2006). Within the category of strategy instruction interventions, there is a large subset of studies that specifically follow methods for self-regulated strategy development (SRSD). Throughout SRSD, students are taught skill-specific strategies, self-regulation strategies, and declarative, procedural, and conditional knowledge about the strategies. The students actively engage with the teachers and their peers. Writing goals are set and monitored throughout SRSD. Throughout each step of SRSD, the teacher scaffolds student learning, gradually helping to increase student independence and mastery of the strategy. In Graham and Perin’s (2007a) meta-analysis of writing interventions for adolescents, strategy instruction interventions were demonstrated to be highly effective (average weighted effect size = 0.82).

**Summarization.** Summarization interventions teach students how to summarize text and focus on the explicit and systematic teaching of summarization skills (Graham & Perin, 2007a). These interventions employ techniques such as the fading of use of model expert summaries, teaching of summarization rules, teaching of methods for synthesizing information from multiple sources, and focusing on main ideas. Graham and Perin’s
Peer-assisted writing. Writing interventions that use peer-assisted techniques involve students working with each other to plan, draft, and/or revise their written products (Graham & Perin, 2007a). Students provide each other with feedback, help each other choose topics, compose together on computers, and help each other revise and edit text. Graham and Perin’s (2007a) meta-analysis revealed peer-assisted writing as a moderately effective intervention (average weighted effect size = 0.75).

Goal setting. In goal setting writing interventions, students establish goals for their written products (Graham & Perin, 2007a). For example, students might set goals for the number of paragraphs they write, the number of supporting details they include, the types of sentences they write, or steps they will follow to self-edit. These goals are often set by the instructor, by the student, or collaboratively. Before writing, the instructor and student discuss what the goal for that session will be. The student writes and then the written work is reviewed. A discussion of whether or not the student’s goal was met then takes place. This discussion often includes graphing progress on a chart, earning points or other reinforcements for meeting goals, and/or discussing the goal for the next session. Graham and Perin’s (2007a) meta-analysis revealed goal setting as another moderately effective intervention (average weighted effect size = 0.70).

Handwriting, spelling, or typing. Some writing interventions focus on specific mechanical aspects of writing, such as handwriting, spelling, or typing. These interventions typically involve direct, explicit, and systematic instruction in handwriting,
spelling, or typing. These interventions typically do not assess how instruction affects student writing, but instead focus on the sole intent of the intervention, such as how much the student’s handwriting, spelling, or typing improved (Rogers & Graham, 2008). Graham, McKeown, Kiuhara, and Harris’s (2012) meta-analysis of writing instruction for elementary school students revealed that handwriting/spelling/typing interventions were moderately effective interventions (average weighted effect size = 0.55).

**Word processing.** Word processing interventions focus on using computers and word processing programs as instructional tools (Bangert-Drowns, 1993). While a computer is only as intelligent as the person using it is, word processing programs lessen some of the mental burden that writing places on the brain. Users can edit, plan, reorder, and rearrange more easily than they can while writing on paper. Spelling and grammar checking features further reduce mental burdens. The theory about the effect of word processors on writing is that word processors “allow the user to attend to higher order decisions…by removing the mechanical difficulties involved in changing text” (Bangert-Drowns, 1993, p. 72). By decreasing the mental burden of making simple text changes, word processors allow users to “compose longer documents and engage in more revision” (Bangert-Drowns, 1993, p. 72). Additionally, some word processing programs can even have built-in prompts reminding users to check certain aspects of their paragraph and essay structure. Graham and Perin’s (2007a) meta-analysis revealed word processing as a moderately effective intervention (average weighted effect size = 0.55).

**Sentence combining.** Sentence combining interventions are sometimes thought of as an alternative way of teaching grammar (Graham & Perin, 2007a). Sentence
combining interventions teach students how to combine simple sentences into complex sentences. These interventions can include peer assistance and extended practice. Graham and Perin’s (2007a) meta-analysis revealed sentence combining as a moderately effective intervention (average weighted effect size = 0.50).

**Inquiry.** In inquiry interventions, students are taught how to analyze data by engaging in activities designed to help them generate ideas (Graham & Perin, 2007a). The activities involve collecting data, generating hypotheses, making inferences, and responding to critical thinking questions. The students then use this information to inform their writing. Graham and Perin’s (2007a) meta-analysis revealed inquiry as an intervention with a small positive effect (average weighted effect size = 0.32).

**Prewriting.** Prewriting interventions generally involve students engaging in brainstorming activities designed to assist them in idea generation (Graham & Perin, 2007a). Activities include creating semantic webs, using graphic organizers, modeling, practice, independent use, and small group discussions about brainstormed ideas. These tools assist students with writing. Graham and Perin’s (2007a) meta-analysis revealed prewriting as an intervention with a small positive effect (average weighted effect size = 0.32).

**Process writing.** Process writing interventions generally involve increasing the number of writing opportunities, providing students with real audiences, going through the writing process (i.e., planning, revision, production) many times, involving peer support and interaction, and personalizing instruction (Graham & Perin, 2007a). These interventions also typically include a supportive writing environment and self-
reflection/evaluation. One specific type of intervention that uses the process writing approach is the writer’s workshop (e.g., Berne, 2009). Graham and Perin’s (2007a) meta-analysis revealed process writing as an intervention with a small positive effect (average weighted effect size = 0.32).

**Writing-to-learn.** Writing-to-learn interventions came about from research that promoted writing as a way to improve retention and understanding of content in all subject areas (Bangert-Drowns, Hurley, & Wilkinson, 2004). The writing process would encourage active learning, personalization of content, and self-regulation of learning. The writing product would then allow students to engage in dialogues with their instructors about what they had learned. An example of a writing-to-learn intervention is one designed by McCrindle and Christensen (1995) that required students in a university-level biology class to write weekly journals that reviewed content they had learned. These reviews needed to include a description about the ways in which they were learning the content and an assessment of the processes they used to learn the content. Bangert-Drowns et al.’s (2004) meta-analysis revealed writing-to-learn as an intervention with a very small positive effect (average unweighted effect size = 0.26, average weighted effect size = 0.17).

**Model study.** Model study interventions involve providing students with good examples of writing and having them study which components make those examples good (Graham & Perin, 2007a). Students are taught how to evaluate the models and imitate certain patterns in the models. Model study typically involves scaffolding techniques where the models are slowly faded from use. Graham and Perin’s (2007a)
meta-analysis revealed model study as an intervention with a small positive effect (average weighted effect size = 0.25).

More on Goal Setting

Of these interventions, goal setting appears to stand out as a flexible, effective intervention that already is or may be included as a main or complementary component within all of these interventions. As an example, goal setting is already an integral component of strategy instruction-based interventions, specifically within SRSD interventions. In SRSD, students set goals and discuss these goals with teachers and/or peers throughout many of SRSD’s six stages (Graham, 2006). Some peer-assisted writing interventions also have included goal setting as an integral component. For example, Yarrow and Topping (2001) conducted a peer-assisted writing intervention that followed the Paired Writing System (Topping, 1995), which includes peer assistance with clarifying goals that students have set. These goals could focus on whatever issues with which the students are having problems, such as spelling, sentence length, type of sentence, editing strategies, or any variety of other goals. Goal setting is also easily incorporated into inquiry interventions. For example, Hillocks’ (1982) inquiry-based intervention included a component where writing instruction was focused on specific goals and these goals were reviewed with the students through teacher feedback. As an example of including goals within model study, Knudson (1989) taught students how to study literary models in order to improve their writing abilities. Knudson described model study as a product-oriented instructional strategy, in which students should develop clear goals they wish to achieve with their writing. In model study interventions,
then, students could be involved by setting goals for how well they imitate the model. A review of the existing literature indicates that goal setting is useful as a standalone intervention technique and as an enhancement to other effective interventions. Even though goal setting is such an effective and flexible intervention technique, there has been less focus in the literature on goal setting interventions than on many of these other effective writing interventions (Graham & Perin, 2007a, 2007b; Rogers & Graham, 2008).

To date, goal setting interventions have been shown to be effective at improving many areas of student performance, including behavior (e.g., Frayne & Latham, 1987; Schnoll & Zimmerman, 2001), reading (e.g., Conte & Hintze, 2000; Eckert, Ardoin, Daly, & Martens, 2002), math (e.g., Barry & Messer, 2003; Codding, Chan-Iannetta, Palmer, & Lukito, 2009), and writing (e.g., Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977). Within writing interventions specifically, goal setting has been demonstrated to have large effects on improving student writing in both the experimental literature base and in the single-subject research base. Graham & Perin (2007a, 2007b) demonstrated that goal setting interventions improve writing for adolescents significantly (average weighted effect size = 0.70). The studies included in their meta-analysis included setting goals to add more ideas during revision, to write a specific kind of paper, and to include specific structural elements while writing. Although Graham and Perin noted that there were only five studies that focused on setting goals, they concluded that these studies were methodologically sound, suggesting that goal setting interventions tend to be solidly designed and produce desirable effects. Likewise, goal setting has been
validated as an effective intervention when researchers use single-subject research methods. Rogers and Graham (2008) found seven single-subject method studies that investigated goal setting as a writing intervention across grade levels. The median effect size of these studies was 91% nonoverlapping data between the baseline and intervention conditions, which would be considered a large effect. These interventions incorporated goal setting as related to how many words to write, how many words and adjectives to write, how many writing lessons to complete, and percentage of writing assignments completed. Many of these interventions reviewed by Rogers and Graham also received high scores on measures that assessed the quality of the studies’ designs, again suggesting that goal setting interventions have strong design backgrounds and produce positive effects on student writing.

An important element to consider when designing a goal setting intervention is who will set the goals. The majority of prior research on goal setting interventions has focused on those approaches whereby either (a) a student sets his or her own goals, or (b) a teacher (or other interventionist, including peers) sets the desired goals for or in consultation with a student. Across the extant literature, both participant-set and interventionist-set goals appear to be effective at improving student performance (Alberto & Troutman, 2009); and, one method has not been supported over the other as being more effective (e.g., Johnson & Graham, 1990; Locke, Shaw, Saari, & Latham, 1981). Despite these findings, Johnson and Graham (1990) suggest using participative goal setting, which involves a collaborative interaction between the participant and the interventionist while setting goals (i.e., a self selected goal with interventionist feedback).
Such a combination is hypothesized to bolster the outcomes of low-achieving students because it increases the level of commitment required by the participant and allows for structuring goals that are not set too high or too low. In participative goal setting, then, the researcher assists the student in determining an ambitious but achievable goal. By setting that goals are challenging but attainable, student performance improves more than in conditions in which students goals are set without regard to ambitiousness and achievability (e.g., L. S. Fuchs, Fuchs, & Deno, 1985; L. S. Fuchs, Fuchs, & Hamlett, 1989). Regardless of who assigns goals, it is vital that the goals are specific and challenging (Locke et al., 1981).

Another important element to consider when designing a goal setting intervention is whether or not to include additional components. Goal setting interventions appear to be most effective when they incorporate other approaches (Locke et al., 1981). A review of the literature indicates that the majority of goal setting writing interventions include components from successful behavior change interventions, such as performance feedback, self-graphing, and reinforcement (e.g., Barry & Messer, 2003; Codd et al., 2009; Conte & Hintze, 2000; Eckert et al., 2002; Frayne & Latham, 1987; Schnoll & Zimmerman, 2001; Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977). A brief review of these elements follows.

Performance feedback. Performance feedback has been a predominant component of effective goal setting interventions (e.g., Chandler, 2003; Krohn, Skinner, & Fuller, 2012; Van Houten & McKillop, 1977). Specifically, prior research has demonstrated that positive feedback (i.e., informing a participant what he or she did
correctly) improves student outcomes and negative feedback (i.e., informing a participant what he or she did incorrectly) can lead to increased student effort (Kluger & DeNisi, 1996). Moreover, the extant literature suggests that feedback is most effective when it is provided immediately (Codding, Feinberg, Dunn, & Pace, 2005) and is direct and explicit (Li, 2010). Unfortunately, prior studies reporting the effectiveness of goal setting interventions simply have indicated that performance feedback was provided. Without a level of specificity of the exact approaches/methods used (i.e., positive versus negative, immediacy of feedback), it is difficult to ascertain if the results obtained in the previously reported studies were optimal.

**Self-graphing.** Self-graphing is a method of self-monitoring that requires a participant to monitor and record his or her own performance (Mace & Kratochwill, 1988), and charts his or her performance on a graph (Figarola et al., 2008). “Performance” in self-graphing could be academic or behavioral performance that is being measured. Not only have graphic displays of performance been shown to improve student performance (L. S. Fuchs & Fuchs, 1987), but self-graphing can also be used to improve student performance (DiGangi, Maag, & Rutherford, 1991). Aside from the research supporting its use among a wide variety of academic and behavioral strategies, self-graphing is also one method of providing immediate performance feedback. That is, the participant graphs his or her performance immediately after completing a task and visually analyzes that performance (DiGangi et al., 1991). When combined, graphing with performance feedback reveals a greater impact on intervention outcomes for a student than simply informing them about their performance (L. S. Fuchs & Fuchs,
1986). Academic fluency also can be improved when self-graphing is combined with goal setting (Figarola et al., 2008).

**Reinforcement.** With regard to the use of reinforcement within goal setting interventions, the extant literature emphasizes reinforcing goals, not simply the reinforcement of participation in the intervention. Such emphasis has extended from the body of experimental results investigating the application of reinforcement in applied contexts. Specifically, fundamentals of applied behavioral research indicate that contingent reinforcement targeted at changing a behavioral response should be provided only when a participant exhibits the desired change in behavior (Alberto & Troutman, 2009; Skinner 1970, 1974). In goal setting interventions with reinforcement, goal attainment becomes the behavior required for contingent reinforcement (Locke et al., 1981), whether the desired goal is participation or an academic skill. Extending from these fundamental principles, studies examining the use of reinforcement merely for participation have indicated negative effects on participant motivation and performance (e.g., Kruglanski et al., 1975; Schunk, 1983). Taken together, there has been a precedent in the goal setting literature that supports the use of reinforcement of goals, and not the reinforcement of simple participation (e.g., Seabaugh & Schumaker, 1994).

**Purpose of the Study**

The purpose of the present study is to investigate the effects of a goal setting writing intervention for high school students at the individualized level of support. As an approach, goal setting interventions have been demonstrated to be effective for improving the writing performance of high school students (e.g., Seabaugh & Schumaker, 1994;
Van Houten & McKillop, 1977). Moreover, the existing literature reveals that the majority of goal setting interventions incorporate multiple treatment components, such as reinforcement, self-monitoring, and performance feedback (Locke et al., 1981). Thus, a study that incorporates multiple components into one intervention approach appears valid and also typical of goal setting interventions. Specifically, the proposed goal setting study will utilize several research-validated approaches, including performance feedback, self-graphing, and reinforcement, as supporting components to the goal setting. Performance feedback, self-graphing, and reinforcement all have been identified as effective methods for improving performance. Because goal setting interventions tend to incorporate these components, it is not the intention of the present study to investigate each component separately. Rather, it is the intention of the present study to demonstrate how a goal setting intervention that incorporates these components affects student writing performance.

In addition to demonstrating the effectiveness of a goal setting intervention that incorporates reinforcement, self-graphing, and performance feedback, the present study will address several major gaps in the extant literature. First, this study expands the current body of research by demonstrating the effectiveness of a writing intervention for high school aged participants, which can be of direct benefit to educators who desire to employ evidence-based strategies. In recent years, there has been a call for increased research on the application of and outcomes associated with writing interventions for school-aged populations (Graham & Perin, 2007a, 2007b; NCOW, 2003, 2004; Rogers & Graham, 2008) that can be promoted for use within school-based contexts. The Institute
of Education Sciences’ (IES) What Works Clearinghouse (IES, 2011), which is heralded as a leader in promoting evidence-based practices for education, lists only two writing interventions (i.e., Read Naturally, Lindamood Phoneme Sequencing) on its website (IES, 2010a, 2010b). Unfortunately, both of these interventions are indicated as having only limited effectiveness (in fact, the Lindamood Phoneme Sequencing intervention is reported to have a negative Improvement Index). Although more recent studies have begun investigating the effectiveness of writing interventions (Rogers & Graham, 2008), only a few of these studies have been conducted with high school students. On average, only about 19% of the studies included in meta-analyses of writing interventions published within the past 20 years have examined outcomes with high school populations. Second, this study utilizes more robust progress monitoring measures than previous goal setting interventions have used. Much of the prior research has used outcome measures that have been demonstrated to be less valid and less reliable for progress monitoring high school students’ responses to writing interventions (e.g., total words written, holistic ratings, grammar, number of sessions completed), raising concerns regarding the applicability of these studies’ findings to practice (Espin, Scierka, Skare, & Halverson, 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). Instead of using these less age-appropriate measures, this study utilized correct writing sequences (CWS) as the primary outcome measure, as CWS currently is the most valid and reliable outcome measure for monitoring the progress of high school students’ writing (Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). Other curriculum-based measures also are reported in
order to further the discussion of the appropriateness of these measures as progress-monitoring tools for high school aged populations. Third, this study expands the current body of research by using more statistically sound strategies for measuring the impact of the intervention. For decades, prior research has used percent nonoverlapping data (PND) as a measure of effect size. Recently, PND has come under increased scrutiny as an inappropriate measure, with some researchers suggesting the use of nonoverlap of all pairs (NAP) because it has more statistical soundness (e.g., Parker & Vannest, 2009; Petersen-Brown, Karich, & Symons, 2012). As such, the present study will report both measures to demonstrate the rate of improvement for students receiving the intervention. Fourth, there has been a specific call for more single-subject studies of increased variety in the literature on academic interventions (Kratochwill & Stoiber, 2000a, 2000b, 2000c, 2002). The present study will use a changing criterion design, which is both methodologically sound and clinically useful to educators (Kazdin, 2010) because it demonstrates behavior changes gradually over the course of the intervention.

**Research Questions**

1. Does a goal setting intervention that includes performance feedback, self-graphing, and reinforcement improve the writing ability of high school students?
2. Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by the participants?
3. Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by the participants’ teachers?
4. Will intervention effects maintain following the intervention?
CHAPTER II
LITERATURE REVIEW

Background

The core literature behind the present study focuses mainly on individualized writing interventions conducted with high school aged students. Before delving into a review of this core literature, some background is provided on how to assess students’ writing abilities, especially at the high school level. As such, the assessment of writing will be discussed first as background for the subsequent literature review on writing interventions. Through this review of literature, the appropriateness of different assessment and intervention methods for students in high school will be highlighted.

Assessing Writing Skills

What is Measured?

When assessing student performance in writing, both the process of writing and the final written product are considered (Berninger, Garcia, & Abbott, 2009). The process of writing involves text planning, generation, and revision. The writing product involves the production of written words through spelling, grammar usage, punctuation, sentences, paragraphs and other aspects of writing evident in the final written product (Berninger et al., 2009). Writing assessment of the process and product can take two forms: qualitative and quantitative.

Qualitative measures of writing can focus on either the process or the product. Several examples of qualitative writing measures include portfolio appraisals and holistic and analytic rubrics (Berninger et al., 2009). Writing portfolio appraisals assess the
entire writing process, from planning to production to editing to final product. In a portfolio appraisal, such as that used in Berne’s (2009) writing workshop, students collect all stages of a writing assignment and put them into a portfolio. This portfolio begins with samples of a student’s planning or brainstorming pages, includes all drafts of the writing, and ends with the student’s final draft. The teacher then appraises the portfolio. Often, these appraisals include one or more rubrics focused on the entire portfolio and may be broken down into assessing different parts (e.g., Berne, 2009; Knoch, 2009). Despite the relative face value of such an approach, the portfolio appraisal likely is subjective. While useful in the writing workshop, portfolios may not be useful for activities such as progress-monitoring that focus on demonstrating student improvement over a specified period of time (Deno, 1985).

Holistic and analytic rubrics are other qualitative examples of writing assessment that usually are focused on the final writing product. Holistic rubrics are used when a teacher wishes to grade the overall quality of writing while analytic rubrics are used to delve into a more detailed analysis of specific writing traits (Knoch, 2009). Holistic rubrics are quicker to administer, and, as such, would be appropriate as a screening measure of general writing proficiency (Knoch, 2009). State and national writing proficiency assessments, too, are usually based on holistic rubrics. Analytic rubrics would be most appropriate for an in-depth diagnostic analysis of the writer’s strengths and weaknesses. Understandably, analytic rubrics are more time-consuming to complete, but they offer much more useful information for evaluating the effectiveness of writing interventions. One disadvantage to the use of rubrics is that they tend to be subjective.
(i.e., two teachers may use the same rubric and give the same essay different scores, resulting in poor inter-rater reliability). As an example, Stellmack, Konheim-Kalkstein, Manor, Massey, and Schmitz (2009) developed a rubric to assess the introductory paragraphs of research papers. The raters in the study agreed perfectly with each other only 37% of the time. Research on rubrics often discusses poor exact inter-rater reliability and contrasts that finding with better inter-rater reliability when a difference of one or two points is taken into consideration, such as in the Stellmack et al. (2009) study. One or two points of difference may not be substantial on a 0-100 point grading system, but rubrics tend to have a range of only a few points. If a rubric only has six or eight points, a one-point difference in scoring is a large jump on such a small scale. Another concern related to the use of rubrics is that they are not appropriate to use as a way to monitor student progress through interventions. Because rubrics have a fixed set of scores and are rated subjectively, rubrics likely will not show growth over time because they cannot measure small changes in writing performance, which is a requirement for monitoring progress over time (Deno, 1985).

Quantitative writing assessments also can focus on both the writing process and the writing product. The writing process can be evaluated by counting the steps taken in planning, text generation, counting the types of revisions made, or other countable indices that focus on planning, text generation, and/or revision. Similarly, writing products can be evaluated through word counts, frequency counts of specific types of words or phrases, grammatical analysis, or paragraph and sentence lengths. While quantitative measures may not give an overall sense of the writing and certainly have
their own statistical flaws (which will be discussed), they appear to provide more objective comparisons within the context of progress monitoring applications.

One example of a product-focused writing assessment is curriculum-based measurement (CBM). CBM is a set of brief (i.e., one-minute, three-minute) standardized measures of basic academic skills designed to be used as indicators of overall skill within a basic-skills domain. With regard to writing, CBM can be used to assess aspects of writing such as word counts, grammar, and spelling. Overall, CBM offers critical information useful for screening, goal setting, and progress-monitoring. It is the area of progress monitoring that separates CBM from the other product-focused rubrics. Unlike portfolios and rubrics, CBM can be used to monitor student progress, as it is sensitive to small amounts of change in performance (Deno, 1985).

**Curriculum-Based Measurement and Assessment**

In order to evaluate the progress of a student in the general educational curriculum and to assess the effectiveness of an intervention, a measure is needed that is simple and efficient to administer, easily understood, based on the curriculum in the classroom, reliable, and valid. CBM is a method of evaluating student performance that meets these requirements (Deno, 1985; Shapiro, 2004; Shinn, 1989). Curriculum-based assessment (CBA) uses the same measurement procedures as CBM, but the purpose of CBM is to identify students at-risk for failure, whereas the purpose of CBA is to monitor student progress (i.e., to monitor a student’s response to an intervention). For the sake of simplicity, when describing the procedures and probes used in CBM and CBA, both CBM and CBA will be referred to as CBM in the following paragraphs.
Deno (1985) described how CBM was developed as a simple, efficient, and inexpensive way to assess student performance. It is less time-consuming to train educators to score student essays using writing CBM measures than it is to train them to score essays using other methods, such as rubrics. CBM also can be used to screen for students who are at risk for failing end-of-the-year and high stakes assessments (Hintze & Silberglitt, 2005). Students identified as at-risk through CBM are then provided with interventions and their responses to these interventions can be monitored using CBA. A student’s response to the intervention is monitored to predict whether they will fail or succeed with continued intervention services. By telling teachers that a student can read only a certain number of words per minute, they more readily will understand what that student’s abilities are than by reporting standard scores for reading fluency (Deno, 1985). In addition, teachers can observe direct results when monitoring and graphing words read per minute. Such data gathering through CBM facilitates collaboration between general education and special education teachers (Deno, 1985). As such, CBM and CBA are useful tools for problem-solving teams in schools.

Through conducting assessments derived from the student’s current curriculum, teachers more easily can tailor interventions to meet students’ needs. Norm-referenced assessments, such as the *Woodcock-Johnson Tests of Achievement—Third Edition* (WJ-ACH-III; Woodcock, McGrew, & Mather, 2001), are not curriculum-based and, as such, do not provide teachers with clear and direct links to the material that is being taught in the classroom. Moreover, each form from norm-referenced assessments can only be given once per year; otherwise, the students may perform better simply because they are
already familiar with the questions that were given. When screening student performance and monitoring student progress, multiple assessment forms are needed because student progress will be measured multiple times throughout the year. Norm-referenced assessments often require time between subtest administration, and would be appropriate if progress monitoring probes were only going to occur once every three months, for example (Fletcher, Lyon, Fuchs, & Barnes, 2007). For more frequent assessment, however, student performance on norm-referenced assessments simply cannot provide the data necessary for monitoring a student’s response to an intervention.

Using CBM to screen for academic problems (i.e., assessing students before they are referred for special education evaluation) is the first step in developing an effective method of problem-solving in the schools (Deno, 1989). Recent educational initiatives have helped to promote the use of CBM. For example, the Individuals with Disabilities Education Improvement Act [IDEIA] of 2004 (IDEIA, 2004) requires schools to implement interventions in the classroom and measure a student’s response to those interventions prior to evaluating the student for special education eligibility determination. CBM is reliable, valid, and useful as a formative assessment of academic progress (Deno, 1989; Shinn, 1989). CBM also permits direct and repeated assessment of student performance in the curriculum for the purpose of gaining information useful for educational decision-making (Shinn, 2008). More specifically, CBM identifies students who are at-risk for academic failure by assessing the performance of all students at least three times per year (e.g., in the fall, winter, and spring; Deno, 1985; Shapiro, 2004; Shinn, 1989). CBM has been designed to assess performance in reading skills,
mathematical skills, and writing skills (including spelling). The procedures may sound complicated, but are easy to learn and administer once training has been received. The next section contains a literature-derived discussion of instructional assessment. Because the focus of the present study is on writing, the descriptions for reading and mathematics are not included.

**CBM, fluency, and improved academic performance.** Methods that utilize CBM as an outcome measure typically monitor fluency skills. Using CBM as a way to monitor progress allows teachers to quickly gauge student fluency and provide immediate feedback about performance. Through CBM and immediate feedback, fluency improves. When fluency improves, academic performance also improves on high-stakes tests, end of the year assessments, and other assessments (Deno, 1985; Deno et al., 2002). CBM not only allows teachers to provide immediate feedback, it allows teachers to use data to change an intervention based on a student’s response to that intervention (Deno, 1985; Shinn, 1989; Shapiro, 2004). This effect has been demonstrated to improve student writing fluency on end of the year writing assessments when teachers use CBM to inform feedback and interventions about writing (Deno, 1985; Deno, Marston, & Mirkin, 1982; McMaster & Campbell, 2008; McMaster & Espin, 2007; Tindal & Parker, 1991).

**Curriculum-based measurement in writing.** CBM in writing generally focuses on analysis of student writing samples. Writing samples are conducted by providing students with age-appropriate prompts, generally referred to as story starters. The most commonly used story starters are meant to easily stimulate writing. For example, students may be given a narrative story prompt, such as, “It was a dark and stormy
night…. Older students may be provided with prompts similar to in-class essays and/or state assessments that are designed to elicit persuasive or expository writing skills (e.g., López & Thompson, 2011; McMaster & Campbell, 2008). Regardless of the type of prompt, CBM in writing is conducted by providing students one minute to think about their story and three minutes to write (Shapiro, 2004). The resulting writing samples can be scored for many different types of quantifiable elements. Specifically, the three most common ways to score CBM writing probes include total words written (TWW), words spelled correctly (WSC), and correct writing sequences (CWS). Because TWW and CWS have been shown to be the most reliable and valid measures of writing performance (Deno et al., 1982; Espin, Scierka, Skare, & Halverson, 1999; Jewell & Malecki, 2005; McMaster & Espin, 2007; Tindal & Parker, 1991), they are described further below.

To measure TWW, a student’s sample simply is counted for how many words it contains (Shapiro, 2004; Shinn, 1989). A “word” is any combination of letters and/or non-letters separated by a space, and does not have to be spelled correctly to count in the TWW score (Shapiro, 2004; Shinn, 1989). TWW has been demonstrated to be a reliable and valid measure of the writing performance of elementary students that may be used to monitor a student’s progress and predict student performance on end of the year assessments (Deno, 1985; Deno et al., 1982; McMaster & Campbell, 2008; McMaster & Espin, 2007; Tindal & Parker, 1991). The extant literature has revealed, however that TWW is not appropriate for older students (e.g., Amato & Watkins, 2011; Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007; Tindal & Parker, 1991). The lack of effectiveness of TWW for older students relates to
the fact that such students who write longer responses may not take sufficient time to plan, edit, and revise their work. As a result, their responses may be somewhat disorganized and contain more errors in semantics and syntax, in addition to spelling. For students in grades six and above, then, CWS is a more reliable and valid measure of performance because it accounts for semantics, syntax, and spelling (Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). Given that struggling writers in high school often experience difficulty with the planning and organizational requirements of writing tasks, the use of CWS will provide a more valuable score. The use of CWS as it relates directly to high school students is described in much greater detail below.

**Scoring curriculum-based assessment measures with high school students.**

CWS is defined as, “two adjacent, correctly spelled words that are acceptable within the context of the phrase to a native speaker of the English language” (Videen, Marston, & Deno, 1982; p. 11). As such, CWS indirectly measures spelling, punctuation, and grammar/usage (Hessler & Konrad, 2008). CWS is measured by putting carets between correct writing sequences in the writing sample (see Table 1 for scoring examples). One caret is placed at the beginning of a sentence if the first word is: (a) preceded by an indent for the first word in a paragraph or preceded by an ending punctuation mark (e.g., period, exclamation point, or question mark) and a space for all other sentences, (b) spelled correctly, and (c) capitalized. Likewise, a caret is put after the last word in a sentence (and before the ending punctuation mark) if the word is spelled correctly and is followed by grammatically correct ending punctuation.
Table 1

Example of Scoring Correct Writing Sequences (CWS) in Curriculum-Based Assessment of Writing

<table>
<thead>
<tr>
<th>Sentence</th>
<th>CWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>^The^cat^was^hungry^,^He^went^into^the^kitchen^to^get^some^milk^.</td>
<td>15</td>
</tr>
<tr>
<td>^The^cat wuz hungry^,^He^went^into^the ketchen to^get sum milk^.</td>
<td>9</td>
</tr>
<tr>
<td>the cat wuz hungry he^went^into^the ketchen to^get sum milk</td>
<td>4</td>
</tr>
<tr>
<td>the cat hungry went^into ketchen get^milk</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. In the first sentence, everything is spelled correctly, is punctuated correctly, and makes grammatical sense. In the second sentence, everything is punctuated correctly and makes grammatical sense, but there are some spelling mistakes. In the third sentence, everything makes grammatical sense, but there are spelling mistakes and punctuation is missing. In the fourth sentence, there are spelling mistakes, punctuation is missing, and there are writing sequences that do not make grammatical sense.

For the remaining words in the sentence, one caret goes in between each word as long as the words are spelled and capitalized correctly, separated by spaces, and make sense in the grammatical context of the sentence. No caret is placed and the sequence is counted as incorrect if these criteria are not met. Because the grammatical rules surrounding the use of mid-sentence punctuation marks (e.g., commas, dashes) in the English language are very detailed and specific, some scoring standards (e.g., AIMSweb; Powell-Smith & Shinn, 2004) do not consider commas in their scoring guides. If the use of commas, dashes, or other mid-sentence punctuation marks is a specific skill that is being targeted by the intervention, however, the correct and incorrect use of these marks should be counted, so long as the instructor grading these samples has a clear understanding of the rules surrounding proper, improper, and grammatically acceptable ways of using mid-sentence punctuation marks (e.g., knowing the rules surrounding the
use or disuse of the Oxford/serial comma). Regardless of mid-sentence punctuation, the CWS score is the total number of carets in a sample. Because CWS can differentiate among students at different skill levels and is sensitive to performance (Marston, Deno, & Tindal, 1983), not only is it useful as a progress-monitoring tool, it is also a reliable and valid predictor of success on high-stakes assessments, holistic measures, norm-referenced tests, and classroom grades for older students (Espin et al., 2000; Espin et al., 2008; López & Thompson, 2011; Tindal & Parker, 1989; Videen et al., 1982).

Some researchers have used other variations of CWS to assess performance. For example, Tindal and Parker (1989) used the percentage of CWS (%CWS) in the writing sample. This calculation involves counting the number of CWS and dividing that number by the total number of word sequences. A positive aspect of %CWS is that it is a production-independent measure, which means that it does not depend on the total number of words a student writes. Production-independent measures may be more appropriate ways of measuring the writing performance of older students (Jewell & Malecki, 2005). As students get older, writing more words does not necessarily result in better writing; students may not plan what they will write. As a result, they may write many words, but their writing will lack purpose, substance, and coherence. Older students who take the time to plan what they write likely will be more succinct, purposeful, and articulate. Although %CWS yields useful information and is a good predictor of student performance on holistic measures (Amato & Watkins, 2011), the fact that it is a production-independent measure means that it is not always sensitive to change and would not be useful as a progress monitoring tool (Tindal & Parker, 1989).
Correct minus incorrect writing sequences (CIWS) is another variation of CWS (Espin et al., 2000). CIWS involves counting the number of correct and incorrect writing sequences, and then subtracting the number of incorrect writing sequences from the number of correct writing sequences. For example, a written product might contain 40 total writing sequences. The student wrote 25 of those sequences correctly and 15 of them incorrectly. Those 15 incorrect sequences are subtracted from the 25 correct sequences, yielding a CIWS of 10. CIWS is reliable and valid for both narrative and expository prompts, which indicates its usefulness for older students, and is a reliable and valid predictor of performance on high stakes tests (Espin et al., 2008). Unlike %CWS, CIWS is appropriate for progress monitoring because it is sensitive to growth (McMaster & Campbell, 2008).

**Future practice and research using writing CBA with high school students.**

When utilizing writing CBA as part of an overall systemic intervention process, educators need an outcome measure that is efficient and linked directly to the curriculum, and can predict performance and monitor the progress of a student over time. Having too many measures is time-consuming, inefficient, and confusing, and likely will lead to abandonment by educators. Educators should use those measures that currently have the most research support and meet the criteria outlined above (Espin et al., 2000; Espin et al., 2005; Espin, De La Paz, Scierka, & Roelofs, 2008; López & Thompson, 2011; Marston et al., 1983; McMaster & Campbell, 2008; Tindal & Parker, 1989; Videen et al., 1982; Weissenburger & Espin, 2005). Investigators who are interested in employing writing CBA in their research with high school students may have more time and
flexibility with their studies than teachers have in their classrooms. As such, investigators may be interested in examining and comparing student performance on other promising outcome measures as well.

The measures with the most support for use with high school students include CWS, %CWS, and CIWS (Amato & Watkins, 2011; Espin et al., 2000; Espin et al., 2005; Espin et al., 2008; Jewell & Malecki, 2005; López & Thompson, 2011; Marston et al., 1983; McMaster & Campbell, 2008; Tindal & Parker, 1989; Videen et al., 1982; Weissenburger & Espin, 2005). Although CWS, %CWS, and CIWS are all useful for measuring overall writing performance and predicting future performance, CWS and CIWS have been demonstrated to be valid tools for progress monitoring. Additionally, researchers may want to investigate other outcome measures that have less, but still promising, research support. Mean length of CWS (MLCWS; Espin et al., 1999), number of characters per word (characters/word; Espin et al., 1999), total sentences written (Espin et al., 1999), incorrect writing sequences (IWS; Diercks-Gransee, Weissenburger, Johnson, & Christensen, 2009; Videen et al., 1982), and number of correct punctuation marks (CPM; Diercks-Gransee et al., 2009; Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002) all have promising reliability and validity as writing CBM outcome measures for older students; however, few studies have employed such measures.

Researchers interested in writing CBM have called for many expansions to the literature. For example, McMaster & Campbell (2008) called for more research examining the sensitivity of writing CBM measures to student growth when strong
interventions are in place. They questioned if one sole measure is able to demonstrate
growth over brief periods, and, thus, would be useful to teachers for instructional
decision-making. Overall, McMaster and Campbell suggest that multiple measures of
writing performance would be more useful than just one. Espin et al. (1999) agree,
stating that more complex measures or combinations of measures should be developed
and investigated, but that the feasibility and sustainability of CBM should be retained.
They offered MLCWS, CWS, characters/word, and number of sentences as reliable and
valid measures that are sensitive to small differences in performance of high school
students. McMaster and Espin (2007), subsequently, added CIWS to this list.

It is important to note, however, that support for all the various measures of
writing is mixed. For example, CIWS and CWS were found to be reliable and valid in
various studies by Espin et al. (2000, 2005, 2008), among others (e.g., López &
Thompson, 2011; McMaster & Campbell, 2008), but Amato and Watkins (2011) did not
find that these measures could reliably predict performance on the Test of Written
Language, Third Edition (TOWL-3; Hammill & Larsen, 1996), a common norm-
referenced writing assessment. Additionally, Deno et al. (1982) found that some product-
dependent indices (e.g., TWW) possessed criterion-related validity for sixth-grade
students while Jewell and Malecki (2005) did not. Clearly, the utility of writing CBM for
progress monitoring and predicting performance is an area in need of more research.
Future research and practice should examine the various CBM measures highlighted
within this review in order to determine which are the most useful for progress-
monitoring for high school students. Additionally, future research and practice should
investigate the use of persuasive and expository prompts over the use of narrative prompts for CBM with high school students (McMaster & Campbell, 2008). Future research and practice also should investigate the use of a 5-minute time limit for writing samples for high school students versus the standard 3-minute time limit, since it has been suggested that such students may produce better writing when provided with a 5-minute time limit (McMaster & Campbell, 2008).

Evidence-Based Practices in Writing Intervention

When Is an Intervention Considered Evidence-Based?

There are many ways to determine whether or not an intervention is evidence-based. Cook, Tankersley, & Landrum (2009) described the quantity of studies needed to call an intervention “evidence-based,” noting the differences touted by various professional organizations. These organizations have different requirements for findings from group-based experimental designs and single-subject research designs. Regarding group-based designs, the Division 12 Task Force of the American Psychological Association (APA) calls for two well-designed between-group experiments to support an intervention in order to call the intervention well-established and either two group experiments that used a wait-list control group or one group experiment that met all quality criteria—except the criterion of having multiple investigators—to call a treatment possibly efficacious (Chambless et al., 1996). The What Works Clearinghouse (WWC) (IES, 2011) requires a positive effect intervention to have support from two or more group-based (i.e., true experimental) design studies, of which at least one study must
meet the WWC’s quality standards without reservations. The WWC states that there also must be no studies that demonstrate significant negative effects of the intervention.

As for single-subject design studies, the Division 12 Task Force of the APA calls for nine single-subject studies to support an intervention in order to call the intervention well-established and three single-subject studies to call a treatment possibly efficacious (Chambless et al., 1996). The WWC, part of the Institute of Education Science (IES), only considers true-experimental designs as being able to stand alone as support for an intervention as having positive effects (IES, 2011). To the WWC, then, single-subject studies would need to be supplemented with group-based designs (i.e., randomized controlled studies) in order for an intervention to be considered as having positive effects; single-subject studies alone could only meet evidence standards with reservations (IES, 2011). Despite the WWC limits on single-subject studies, others have advocated for the usefulness of single-subject studies in establishing intervention practices as evidence-based. For example, Horner et al. (2005) outlined standards for single-subject design studies, calling for a minimum of five single-subject studies conducted by at least three different investigators in at least three different geographical locations involving a total of at least 20 participants that demonstrate positive effects.

It is important to consider these standards when determining whether or not an intervention is evidence-based. Some interventions do not meet these standards simply because their literature base is too small. An insufficient literature base does not necessarily mean that the intervention is ineffective—simply that more replications should be conducted in order to determine its effectiveness. While numerous meta-
analyses and literature reviews have been conducted on writing in general, the number of studies behind one single intervention (e.g., goal setting, peer-assisted strategies) is so small that the interventions barely have enough research behind them to meet the quantitative requirements to be considered evidence-based. When breaking down the literature by specific groups of students (e.g., high school students, struggling students, students with disabilities), the base becomes even smaller.

The WWC does not endorse any writing interventions as having unreservedly positive effects on improving writing achievement for students at any grade level. When conducting a search for writing achievement interventions on the WWC’s website, only two interventions are to be found on the Find What Works page: Read Naturally (Read Naturally, Inc., 2013) and Lindamood Phoneme Sequencing (LiPS; P. C. Lindamood & Lindamood, 2011). The WWC only lists studies with students in the fourth through sixth grades for Read Naturally, and the effects are only potentially positive for improving writing performance (IES, 2010a). For LiPS, the WWC only lists studies with students from kindergarten through sixth grade, and its effects on writing performance are potentially negative (IES, 2010b). Much more research is needed on all writing interventions, then, in order to meet the WWC’s criteria for calling an intervention evidence-based. Other standards, such as those described by Horner et al. (2005), might be useful for current determinations of what interventions are effective at improving the writing performance of high school students.

Other important issues to consider when investigating the effectiveness of or implementing a literature-derived intervention are to monitor treatment integrity and
assess social validity. Treatment integrity is defined as the degree to which an intervention is carried out the way it was designed (Lane, Bocian, MacMillan, & Gresham, 2004). Accuracy and consistency are core components in determining treatment integrity (Lane et al., 2004). In order to determine if an intervention, as it was designed, was effective, an evaluation of treatment integrity is necessary (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000; Lane et al., 2004). Treatment integrity can be monitored in many ways: (a) direct observations of the interventionist; (b) indirect measures, such as self-monitoring, self-reporting, and post-intervention interviews between interventionist and researcher; (c) permanent products from the intervention; (d) manualized treatment protocol; and (e) consultant feedback (Gresham et al., 2000; Lane et al., 2004). Self-report methods are commonly used in intervention research, but they should be confirmed with observations in order to deter against the interventionist’s overinflation of self-reports of integrity (Lane et al., 2004). By monitoring and intervening to foster treatment integrity, a researcher can be certain that the intervention’s internal validity (i.e., the intervention, as it was designed, caused the change), external validity (i.e., the intervention, as it was designed, could be useful for other students with the same problem), and criterion validity (i.e., the intervention, as it was designed, affected what it was supposed to affect) are not negatively impacted by poor implementation of the intervention (Gresham et al., 2000). If any of these types of validity are threatened, it is difficult to draw conclusions about the effectiveness of the treatment (i.e., independent variable)—even in cases resulting in improvement with regard to the dependent variable.
Social validity is defined as the degree to which someone—often a key stakeholder in consultation (e.g., teachers, parents)—perceives an intervention to be desirable, useful, and important (Wolf, 1978). Social validity can be assessed by asking teachers to rate interventions on a variety of levels, such as time commitment of the intervention, intervention outcomes, intervention complexity, and appropriateness of the intervention for the desired outcome of the intervention. Treatment acceptability, an important construct related to social validity, may be defined as, “consumers’ judgments about treatment procedures” (Eckert & Hintze, 2000, p. 124). According to Eckert and Hintze’s (2000) review of research on treatment acceptability, studies may have raters rate interventions based on interventions they themselves performed, on interventions that were provided to them, or on interventions that were provided by someone else. Reimers, Wacker, and Koeppel (1987) argue that treatment acceptability may influence the treatment agent’s (e.g., teacher, parent) level of treatment integrity, which in turn may impact treatment outcomes. Many factors may influence treatment integrity, including the time required to implement the intervention, the complexity of the intervention, the amount of materials and people needed for the intervention, the effectiveness of the intervention, and teacher motivation (Lane et al., 2004).

Research on the correlation between perceptions of social validity (e.g., treatment acceptability) and treatment integrity has been mixed. For example, Eckert and Hintze (2000) revealed that some studies supported a link between higher social validity and higher treatment integrity, while other studies did not support such a link. Nevertheless, social validity is considered by many researchers (e.g., Gresham & Lopez, 1996; Wolf,
1978) to be an important outcome variable to be measured for any intervention research. Assessing stakeholders’ perceptions regarding the social significance of the treatment goals, procedures, and effects of the intervention provides research-practitioners with additional information to consider when further developing research agendas and developing applied interventions.

**Summary.** To summarize, there are many opinions about how much and what kinds of research are needed to determine whether an intervention is evidence-based. Some researchers, such as Horner et al. (2005), lay out requirements that single-subject studies may stand alone in support of an intervention. Other researchers, such as the WWC, do not consider single-subject studies to be able to support an intervention at the highest rating they bestow upon interventions. Social validity and treatment integrity also are also important when considering the overall effectiveness and utility of an intervention. An evidence-based intervention will only work the way it was designed to work if the treatment protocol is followed exactly, and interventionists may only implement an intervention correctly if they perceive the intervention to be effective (Lane et al., 2004; Reimers et al., 1987).

**Evidence-Based Practices for Students in Need of Individualized Interventions**

The present study will involve conducting an individualized writing intervention with high school students who are struggling with writing and who are in need of individualized interventions. The main focus of the literature review, then, will be on single-subject studies because the findings of those studies will be most directly comparable to the outcomes of the present study. When working with students at an
individualized level of support, educators need to employ interventions that have been shown to work with small groups of students or individual students. Unfortunately, the small sample sizes required by such studies often do not contain enough statistical power for the use of true experimental research methods and true experimental effect size measures (Kazdin, 2010). Because the findings of studies using other quantitative research methods (e.g., group-based designs) also are valid and applicable to the findings of the present study, their results will be discussed, albeit in limited detail.

The practices included in this review highlight effective practices for students at all levels of intervention, but mostly at the universal (i.e., building-wide) level and for students of all ability levels, including average and high-achieving students (Rogers & Graham, 2008). As such, there is a need for more research at the individualized level of writing interventions for high school students. If an intervention is said to be evidence-based and the only research support behind that intervention was conducted with large groups of students, it does not necessarily follow that that intervention will be effective at the individualized level. This statement is especially true if an intervention involves group work or peer editing. However, if an intervention is shown to be effective in studies using true experimental, quasi-experimental, single-subject, and other research methods, that intervention’s claims for effectiveness among many groups and types of students are further reinforced (Rogers & Graham, 2008).

In order to determine evidence-based practices at the individualized level of support, educators may wish to look to studies that have used single-subject research methods (Horner et al., 2005). Single-subject research is an experimental research design
that can provide information on interventions for small groups and individual students. These studies do not have the statistical power to establish control through experimental and control groups. Rather, experimental control may be established in multiple ways: (a) by holding some subjects or conditions in baseline conditions while others move into the intervention phase of the study, (b) by returning subjects to the baseline condition following an intervention condition; (c) by changing the criteria or goals of an intervention, so that each criterion change acts as a type of baseline to the criterion that follows it; (d) by alternating among various treatments within the intervention phase, though never administering more than one treatment at the same exact time; or (e) any combination of the aforementioned designs (Kazdin, 2010). Experimental control can be enhanced in single-subject research through repeated measurement of performance that allows for the comparison of patterns of performance before and after intervention, resulting in “at least three demonstrations of experimental effect at three different points in time” (Horner et al., 2005, p. 169). Thus, looking to single-subject research in addition to group-based research is useful for investigating the evidence base for interventions designed to work with individual students.

**A call for more single-subject research.** Over the past decade, there has been a call for more single-subject research to be conducted in order to enhance the literature base behind interventions (Kratochwill & Stoiber, 2000a, 2000b, 2000c, 2002). For example, Kratochwill and Stoiber (2000a) called for an expansion of the types of research that should be conducted and used as evidence to support an intervention’s effectiveness, including single-subject methodology included. In a similar vein, Kratochwill and
Stoiber (2000b) highlighted the need for a stronger link between research-based interventions and their subsequent implementation in real world settings, which sometimes compromises the treatment integrity and validity of the research base. Kratochwill and Stoiber’s (2000c) further emphasized the need for multiple types of research designs to be employed in studies in order to determine an intervention’s effectiveness, and for the improvement of the link between research and practice. They suggested that single-subject designs may be a useful way to bring research into practice, especially because of the ability of single-subject designs to be more responsive to changes in the environment that are more common in practice than in well-controlled research studies. Finally, Kratochwill and Stoiber (2002) promoted the use of single-subject research methods for improving the transition of an intervention from research to practice, as well as for improving the communication among researchers to improve intervention effectiveness. Moreover, they indicated that single-subject research methods also may improve the data supporting an intervention’s effectiveness with culturally diverse students. Kratochwill and Stoiber’s (2000a, 2000b, 2000c, 2002) repeated call for increased single-subject research conducted within applied contexts may be used as an impetus for further action-based research pertaining to effective interventions at all levels of support.

Writing Intervention Research with High School Students

To date, a plethora of writing intervention research has been conducted, though less specifically at the high school level. As a result, there have been ten meta-analyses of the extant literature that span grade levels and intervention types (Bangert-Drowns,
1993; Bangert-Drowns, Hurley, & Wilkinson, 2004; Goldberg, Russell, & Cook, 2003; Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a, 2007b; Graham, McKeown, Kiuhara, & Harris, 2012; Graham & Sandmel, 2011; Hillocks, 1986; Morphy & Graham, 2012; Rogers & Graham, 2008). The findings of these meta-analyses will be discussed first, in order to provide a general review of the writing intervention literature. Specific descriptions of the interventions that these meta-analyses demonstrated as effective will follow this review. These studies will be discussed in chronological order, beginning with Hillocks (1986), with the exception of ending with Rogers and Graham (2008), as the Rogers and Graham meta-analysis is the most relevant to the present study. Finally, a review of the literature on single-subject writing interventions for high school students will round out the literature review. Common themes throughout this review will focus on the applicability of findings to high school students and the applicability of findings to interventions at the individualized level of intervention support, which are two key features of the present study.

**Meta-Analyses of Writing Intervention Research**

Hillocks (1986) conducted one of the first detailed and comprehensive meta-analysis of writing instruction over a 20-year period. Specifically, Hillocks investigated how multiple kinds of writing interventions improved writing quality for students in grade 3 through college. Hillocks was driven to conduct this meta-analysis because of a turn away from experimental research in writing at that time, combined with negative opinions about writing research as being unapproachable and useless to the classroom teacher. As such, Hillocks set out to pull together all writing research and find out which
techniques were successful and which were not. The study designs included in Hillocks’
review included studies with experimental and quasi-experimental research methods.
What Hillocks found was that sentence combining, model study, student evaluation (i.e.,
peer-assisted evaluation), and inquiry interventions improved writing quality (as
measured by rubrics) whereas grammar and free-writing interventions did not improve
writing quality. In fact, Hillocks revealed that grammar and mechanic instruction, at
times, brought down the quality of student writing because of the increased focus that
students and teachers paid to mechanics and correctness, instead of content. Of interest
to the present study is how many of the interventions in Hillocks’ review were conducted
with high school students. In his analysis, Hillocks categorized students by elementary
(grades 3 through 6), secondary (grades 7 through 12), college, and mixed elementary
and secondary. Of the 69 studies included in his review, 30 of them were categorized as
including secondary school students. Unfortunately for the interest of the present study,
Hillocks’ review did not specify which studies fell into the secondary grade level
category. While the text of the review indicates the grade levels for some of the studies,
the text does not specify the grade level of the participants in every study included in the
review. As such, the exact number of studies that only included high school students is
not able to be calculated.

Bangert-Drowns (1993) conducted a meta-analysis of 33 studies investigating the
use of word processing (i.e., use of computers) for writing instruction. The advent of
word processing permitted students to focus on writing content instead of writing
mechanics, word processing tapped into different aspects of the writing process.
Bangert-Drowns hypothesized that because computers made editing easier, they could improve students’ writing abilities. Specifically, Bangert-Drowns theorized that writing quality would improve because students would more easily be able to edit their writing, would be free from worrying about mechanics, and would edit more often. Being freed of these concerns, students would be able to engage in higher-level thinking (e.g., organization, clarity). As such, Bangert-Drowns examined experimental studies that compared students who used word processors to students who did not use word processors when receiving writing instruction.

Overall, Bangert-Drowns (1993) revealed that the use of word processing in remedial writing instruction improved the quality of student writing (as measured by holistic rubrics) and reduced variance among writing quality scores. Unfortunately, there was not enough information provided in the studies for Bangert-Drowns to draw conclusions about whether instructional or student factors influenced the relationships among writing quality, word processing, and ability. Despite this fact, Bangert-Drowns’s meta-analysis outlined three specific effects when using word processing in writing instruction. First, students who had struggled with writing and were receiving remedial writing instruction saw their writing improve when using word processing. Second, the use of word processing with students receiving remedial instruction decreased the variability in writing quality among these students, as they all improved their scores to more consistent levels, whereas the control group that did not use a word processor exhibited higher variability in the quality scores of their writing. Third, the length of time that students used word processing did not affect writing quality. Taken together,
the results of Bangert-Drowns’ meta-analysis indicated that the use of word processing had a positive effect on student outcomes, which may, in part, been a result of students’ ability to engage in more reflective writing. Although these results are encouraging, they have limited application to the current study since only two of the 33 studies in this meta-analysis included high school student participants.

Ten years later, Goldberg, et al. (2003) expanded upon Bangert-Drowns’s (1993) meta-analysis by investigating experimental research that examined the effect of computers on writing between 1992 and 2002. Goldberg et al. limited their search to 1992-2002 because of the technological advances that had been made in word processing since the 1990s. Within their search criteria, they found 26 studies that compared students who wrote using computers with students who wrote on paper. Results of the meta-analysis revealed that students who used computers wrote longer and better pieces, though they did not necessarily revise their work more. Goldberg et al. concluded that the use of computers for writing not only improved writing (as measured by word counts and rubrics for writing quality), but also fostered increased engagement and motivation in students. Despite these findings, the results are limited in their application to high school students because only seven of the 14 studies involved high school students.

In the same year, Graham and Harris (2003) conducted a meta-analysis of single-subject studies that incorporated the use of the self-regulated strategy development (SRSD) model with students with specific learning disabilities (SLD). They focused on students with learning disabilities because the writing of these students often is a “condensed” version of the writing of skilled writers (p. 323). That is, students with SLD
condense the essential components of writing, including: planning, text generation, and revision. These researchers argued that writing instruction for students with SLD should focus on expanding these specific areas. As a result, Graham and Harris identified 18 studies on SRSD that focused on writing, 13 of which focused on populations of students with SLD. Overall, SRSD instruction produced large effect sizes. For students with SLD, average effect sizes ranged from 1.14 to above 2.0, depending on which element of writing was targeted by the intervention (e.g., quality, length, grammar, story elements). Moreover, SRSD appeared to be effective in promoting maintenance and generalization, highlighting its long-term potential to improve student performance. Despite such positive findings, none of the studies included in this meta-analysis included high school students.

Bangert-Drowns et al. (2004) conducted a meta-analysis on experimental writing-to-learn interventions and their effect on academic performance. Writing-to-learn interventions use writing as a method for learning academic material, which allows students to be exposed to more content and spend more time on task. In this way, writing helps students to use cognitive strategies that enhance learning, such as rehearsal, elaboration, organization, and revision. Bangert-Drowns et al. identified 48 studies that utilized writing-to-learn strategies for students from elementary school through college grade levels. The academic subjects that were being targeted through the writing-to-learn interventions covered a wide variety of subjects, including math, algebra, earth science, chemistry, biology, social studies, world history, natural resources, and some college-level courses. Results of the meta-analysis demonstrated that when the students were
provided with meta-cognitive prompts and spent a longer time in intervention, their academic achievement (as measured by final grades, final examinations, standardized tests, tests from textbooks, locally-developed content quizzes, or rubrics) improved significantly. As was the case with prior meta-analyses investigating writing interventions, however, only ten out of the 48 studies included populations of high school students.

Graham (2006) conducted a meta-analysis of 39 studies (including group-based and single-subject research designs) that used strategy instruction in writing. Overall, results of this meta-analysis revealed that strategy instruction improved writing performance, with large effect sizes evident in the group-based studies (overall mean effect size = 1.15) and strong effect sizes (overall mean effect size PND = 89%) for single-subject studies. Measures of writing performance mainly included word count, number of target elements contained within the written text, and rubrics. Graham’s findings indicated that strategy instruction improved the writing for students of all ages and grade levels, regardless of writing genre (i.e., narrative or expository). Graham cautioned that an important component in any intervention is whether a student is able to maintain skills over time and generalize skills to different settings. Despite this assertion, only 54% of the studies assessed generalization or maintenance. For those that did assess generalization or maintenance, effect sizes remained strong. As was the case of several meta-analyses conducted on writing interventions up to this point, only four of the 39 studies included data applicable to high school students, thereby limiting the interpretation of potential impact to secondary students.
In a more comprehensive meta-analytic design, Graham and Perin (2007a, 2007b) examined the collective effects of experimental and quasi-experimental writing interventions for students in grades 4 through 12 (their study was first published in an academic journal and then expanded upon later that same year in a report to the Carnegie Corporation of New York). The authors categorized the writing interventions that they investigated into four groups: (a) the process writing approach; (b) explicit teaching of skills; (c) scaffolding; and (d) alternative composition methods (i.e., word processing). The studies had to include a quantitative measure of writing quality (e.g., rubric) to be included in the analysis. Overall, several important results emerged. First, the process writing intervention studies yielded an average weighted effect size of $d = 0.32$, which is considered to be small. Second, the interventions for explicit teaching of skills yielded medium to large effect sizes with sentence combining ($d = 0.50$), strategy instruction ($d = 0.82$), and summarization ($d = 0.82$), each demonstrating overall positive effects. Explicit skills interventions for grammar, on the other hand, yielded a small and possibly deleterious effect ($d = -0.32$). Third, scaffolding interventions, which focused on providing students with assistance in prewriting, inquiry, peer assistance, model study, and goal setting, revealed mixed results. Specifically, the peer assistance and goal setting interventions ($d = 0.75$ and $0.70$, respectively) demonstrated effectiveness. The prewriting, inquiry, and model study interventions ($d = 0.32$, 0.32, and 0.25, respectively), however, proved to be limited in effecting change. Fourth, alternative composition interventions that utilized computer word processing programs had a medium effect on writing quality ($d = 0.55$). Overall, the results of this investigation
indicated that explicit teaching of strategy instruction, sentence combining, and summarization; scaffolding through peer assistance and goal setting; and word processing programs were effective interventions. As was the case with all of the prior meta-analyses conducted to date, a small percentage (only 36 of 127) of the studies focused on high school students.

Focusing specifically on the process approach to writing (e.g., writing workshop-type approaches), Graham and Sandmel (2011) conducted a meta-analysis of 29 experimental and quasi-experimental studies that included students in grades 1 through 12 (only 11 studies sampled high school populations). The authors were interested in interventions that included a quantitative measure of writing quality (e.g., rubric, norm-referenced assessment such as TOWL-3) and/or motivation. The meta-analysis yielded an overall effect size (ES) that was small but statistically significant (ES = 0.34) for students in general education classes, suggesting that the process approach is effective at improving writing quality for these students. While the process writing approach demonstrated some effectiveness for students in general education classes, it was not effective for improving the writing quality of struggling writers (ES = 0.29, with a confidence interval that crossed 0). The motivation to write was not improved for general education students or for struggling writers.

Morphy and Graham (2012) investigated those studies that used word processing to improve a variety of aspects of student writing. Studies had to employ an experimental, quasi-experimental with a pretest, or counter-balanced within-participants design. Specifically, they analyzed 27 studies with students ranging from grades 1
through 12 who were struggling with reading and/or writing (10 of the 27 studies included high school students). Of those factors relevant to this literature review, the use of a word processor improved the development/organization ($d = 0.66$), mechanics ($d = 0.61$), quality ($d = 0.52$), and length ($d = 0.48$) of student writing. Three of the studies included in their review employed a specific word processing program that provided feedback to the students as they wrote, as opposed to simply providing the students with a computer instead of a pen/pencil and paper. Feedback included comments about quality or prompts to instigate planning, drafting, or revising. When examined separately, these three studies produced a large overall effect size ($d = 1.46$).

Most recently, Graham et al. (2012) focused on reviewing writing instruction in the elementary grades. The span of “elementary grades” was defined by the school’s designation as an elementary school, which were typically either grades 1 through 5 or 1 through 6. The authors located 115 experimental and quasi-experimental (with a pretest) writing instruction studies that met their inclusion criteria, and focused on explicit teaching of skills, scaffolding, word processing, extra writing, and comprehensive writing. Analysis of these studies revealed that explicit teaching through strategy instruction (ES = 1.02), creativity/imagery instruction (ES = 0.70), text structure instruction (ES = 0.59), SRSD (ES = 0.50), and transcription skills instruction (ES = 0.22) demonstrated statistically significant effects on improving the quality of student writing (as measured by rubrics). Explicit instruction of grammar, however, did not. Scaffolding writing through peer assistance (ES = 0.89), goal setting (ES = 0.76), prewriting (ES = 0.54), and assessment (ES = 0.42) also demonstrated significantly
improved writing quality. Student writing quality also was improved significantly through word processing (ES = 0.47), comprehensive writing (ES = 0.42), and extra writing (ES = 0.30) instruction. Because this meta-analysis focused on elementary school students, none of the studies incorporated into this meta-analysis included high school students. Results of this meta-analysis further illustrate how strategy instruction, creative/imagery instruction, scaffolding through peer assistance, and goal setting demonstrate large effects on improving students’ writings, and how text structure instruction, SRSD, and prewriting demonstrate moderate effects.

Because the present study will utilize single-subject research methods with high school students, a relatively recent—and a very comprehensive—meta-analysis by Rogers and Graham (2008) is described in much greater detail in the context of this review. Based on the prior meta-analyses of experimental, quasi-experimental, and single-subject research in writing interventions conducted previously (Bangert-Drowns, 1993; Bangert-Drowns et al., 2004; Goldberg et al., 2003; Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a, 2007b; Hillocks, 1986), Rogers and Graham revealed, in order of magnitude of their impact, the 12 interventions that have been shown to improve the writing for elementary and secondary students (see Table 2).
Table 2

*Effective Writing Interventions as Identified by Rogers and Graham (2008)*

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Strategy Instruction</td>
<td>Strategy instruction interventions focus on teaching students specific strategies for different aspects of writing, such as planning, writing, revising, and editing.</td>
</tr>
<tr>
<td>2. Summarization</td>
<td>Summarization interventions teach students how to summarize text and focus on the explicit and systematic teaching of summarization skills.</td>
</tr>
<tr>
<td>3. Peer-Assisted Writing</td>
<td>Peer-assisted writing interventions involve students working with each other to plan, draft, and/or revise their written products.</td>
</tr>
<tr>
<td>4. Goal Setting</td>
<td>In goal setting interventions, students set increasing goals, based on previous performance, for their written products.</td>
</tr>
<tr>
<td>5. Handwriting, Spelling, or Typing</td>
<td>Handwriting, spelling, or typing interventions involve direct, explicit, and systematic instruction in handwriting, spelling, or typing.</td>
</tr>
<tr>
<td>6. Word Processing</td>
<td>Word processing interventions focus on using computers and word processing programs as instructional tools.</td>
</tr>
<tr>
<td>7. Sentence Combining</td>
<td>Sentence combining interventions teach students how to combine simple sentences into complex sentences.</td>
</tr>
<tr>
<td>8. Inquiry</td>
<td>Inquiry interventions teach students how to analyze data by engaging in activities designed to help them generate ideas.</td>
</tr>
<tr>
<td>9. Prewriting</td>
<td>Prewriting interventions involve students in brainstorming activities designed to assist them in idea generation.</td>
</tr>
<tr>
<td>10. Process Writing</td>
<td>Process writing interventions involve increasing the number of writing opportunities, providing students with real audiences, improving peer support and interaction, and personalizing instruction.</td>
</tr>
<tr>
<td>11. Writing-to-Learn</td>
<td>Writing-to-learn interventions focus on using writing as a way to improve retention and understanding of content in all subject areas.</td>
</tr>
<tr>
<td>12. Model Study</td>
<td>Model study interventions involve providing students with good examples of writing and having them study which components make those examples good.</td>
</tr>
</tbody>
</table>

*Note.* Compiled from Rogers and Graham’s (2008) ranking of effective writing interventions, which was based on the findings of previous meta-analyses. The more recent meta-analyses (Graham & Sandmel, 2011; Graham et al., 2012; Morphy & Graham, 2012) support the effectiveness and magnitude of impact of these interventions.
Of these interventions, only four (strategy instruction, process writing, word processing, and writing-to-learn) were based on 10 or more studies (Rogers & Graham, 2008). This number of studies barely meets commonly applied criteria for a sufficient evidence base, such as Chambless et al. (1996), who require at least nine single-subject studies conducted by at least two different investigators or research groups, and which incorporate: (a) experimental designs of good quality, (b) comparison to another treatment or placebo, (c) treatment manuals, and (d) demographic descriptions of participants, in order for that intervention to be considered an intervention with a well-established research base. Additionally, few of these studies were conducted with high school students. For example, of the 127 effect sizes reported in Graham and Perin’s review (2007a) of writing interventions with students in grades 4 through 12, only 36 originated from studies that included high school students (see Table 3). Further, most of these independent studies utilized true and quasi-experimental research designs and were not focused solely on struggling students (i.e., students in need of individualized interventions). In order to determine which studies provide evidence-based information for use with writers at the individualized level in high schools, a broader net needs to be cast.

While a host of prior meta-analyses (e.g., Hillocks, 1986; Graham & Perin, 2007a, 2007b) and comprehensive literature reviews (e.g., Mason & Graham, 2008) have been conducted they have tended to exclude most single-subject studies due to the difficulty of comparing data from single-subject and true experimental methods within the same study.
Table 3

<table>
<thead>
<tr>
<th>Meta-Analysis</th>
<th>Number of Studies&lt;sup&gt;a&lt;/sup&gt; Conducted with High School Students</th>
<th>Total Number of Studies</th>
<th>Percentage of Studies Conducted with High School Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangert-Drowns (1993)</td>
<td>2</td>
<td>33</td>
<td>6%</td>
</tr>
<tr>
<td>Goldberg, Russell, &amp; Cook (2003)</td>
<td>4</td>
<td>14</td>
<td>29%</td>
</tr>
<tr>
<td>Graham (2006)</td>
<td>4</td>
<td>39</td>
<td>10%</td>
</tr>
<tr>
<td>Graham &amp; Harris (2003)</td>
<td>0</td>
<td>18</td>
<td>0%</td>
</tr>
<tr>
<td>Graham &amp; Sandmel (2011)</td>
<td>11</td>
<td>29</td>
<td>38%</td>
</tr>
<tr>
<td>Graham et al. (2012)</td>
<td>0</td>
<td>115</td>
<td>0%</td>
</tr>
<tr>
<td>Graham &amp; Perin (2007a, 2007b)</td>
<td>36</td>
<td>127</td>
<td>28%</td>
</tr>
<tr>
<td>Hillocks (1986)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>39</td>
<td>60</td>
<td>65%</td>
</tr>
<tr>
<td>Morphy &amp; Graham (2012)</td>
<td>10</td>
<td>27</td>
<td>37%</td>
</tr>
<tr>
<td>Rogers &amp; Graham (2008)</td>
<td>16</td>
<td>76</td>
<td>21%</td>
</tr>
</tbody>
</table>

<sup>a</sup>Number of studies calculated either by author report of number of studies included or by counting the number of studies listed in the meta-analysis table.

<sup>b</sup>Students defined as secondary students in grades 7-12, not only 9-12. A breakdown of studies by specific grade level was not provided. The percentage for Hillocks (1986) is thus difficult to compare to the other percentages of studies conducted with high school students.

Rogers and Graham (2008) conducted their meta-analysis of writing interventions while focusing on single-subject research methodology. Overall, they found that strategy instruction for planning/drafting, grammar, and goal setting interventions were effective and the studies behind these interventions were well-designed, utilizing 11 indicators (specified in the Methods chapter) of study quality outlined by Horner et al. (2005).
These indicators focus on descriptions of the study that would allow for other researchers to replicate the study and that allow for those reading the study to understand the methods and calculations behind the study and its results. In addition, Rogers and Graham indicated that strategy instruction for editing, word processing, reinforcement, prewriting, and sentence combining, and strategy instruction for paragraph writing interventions also were effective, but the studies behind these interventions were of lower quality, based on the 11 indicators. Typically, these studies lacked sufficient information to describe methods, demographic information, and results, which would prevent other researchers from being able to replicate the study and scrutinize the results. As a result, Rogers and Graham tempered the findings of these studies. Taken together, the overall impact of the Rogers and Graham analysis indicates that writing interventions with the highest likelihood of impact include strategy instruction for planning/drafting, grammar, and goal setting interventions. Despite these findings, the interpretation of the Rogers and Graham study should be considered carefully since only 16 of the 83 studies included students in high school or of high school age. Therefore, the focus of the next portion of this literature review will highlight these 16 studies, as well as provide additional information on writing interventions that have been conducted with high school students incorporating single-subject research methodology.

Since the Rogers and Graham (2008) study, there likely has been further research investigating the use of writing interventions for high school students. Rogers and Graham described January 2007 as the end date for their search for studies to include in their meta-analysis. As such, any studies about writing interventions for high school
students using single-subject research methods that were published before January 2007 are assumed to have been identified by Rogers and Graham’s broad, rigorous search methods. Rogers and Graham were very comprehensive and searched across multiple databases, numerous journals (both online and in print), many peer-refereed and non-peer-refereed sources, dissertations and theses, book chapters, and reference lists from located studies. Within their search, the terms they used were broad (e.g., summary writing, summary instruction, and summary strategies, which would all be variations on the terminology summary intervention). Aside from this specificity, they did not include within their meta-analysis any intervention that had fewer than four studies published (although such studies were described descriptively).

Following this logic, a search was conducted in order to identify additional studies that had been published after the publication of Rogers and Graham’s study and to identify those interventions that Rogers and Graham found but did not include in their review due to insufficient numbers of studies (i.e., fewer than four). A study was to be included for review if it met the following criteria: (a) participants were in high school or of high school age, (b) single-subject research methods were utilized, and (c) the dependent variable was a measure of writing performance. The measure of writing performance should be either a writing quantity-based measure (e.g., TWW, CWS, number of genre elements) or a writing quality-based measure (e.g., holistic rubric). As in Rogers and Graham’s meta-analysis, the setting was allowed to be in a public school, private school, parochial school, clinic, behavior center, or anywhere where students receive instruction. Because students who require individualized writing intervention and
support are not always be served in general education settings in public schools, this broad inclusion of settings appeared most appropriate.

The search was conducted in four steps. First, the entry for Rogers and Graham’s (2008) study was located on the Science Citation Index database in July 2012. This database allows for the search of publications that cite a given study. The option to search for all articles that cited the Rogers and Graham study was selected. The results list \(N = 18\) was searched for studies that met inclusion criteria. Only two studies (Jacobson & Reid, 2010; Kiuhara, O’Neill, Hawken, & Graham, 2012) met criteria during this search. Second, the APA citation for the Rogers and Graham study was also submitted to the Google Scholar search engine in an attempt to find any more studies that may have been overlooked by the Science Citation Index database. One additional study meeting inclusion criteria was found, out of 61 total hits (Viel-Ruma, Houchins, Jolivette, Fredrick, & Gama, 2010). Third, a search was conducted across several databases and given the confines of searching for studies published since January 2007 in order to find studies that were published after Rogers and Graham completed their search. This search simply included the words *writing* and *intervention*, in order to search as broadly as possible. Out of the 1,411 studies that were found, no studies other than those already identified were found that met the inclusion criteria. Fourth, an attempt was made to find and include in this review the interventions mentioned by Rogers and Graham as being reported in fewer than four studies. Rogers and Graham, consistent with other meta-analyses (e.g., Graham & Perin, 2007a, 2007b), only wanted to draw conclusions about interventions that had four or more studies that used “conceptually similar” outcome
measures (p. 881). The following interventions were identified by Rogers and Graham as being reported in only one or two studies: strategy instruction for revising, goal setting for grammar/sentence construction, direct instruction of a broad array of skills, word processing plus (e.g., word processing plus text read aloud by computer), dialogue journals, direct teaching of self-regulation strategies, repeated writing, and verbal encouragement. These terms were searched and one additional study was found that met inclusion criteria during this process (Walker, Shippen, Alberto, Houchins, & Cihak, 2005).

Common reasons for exclusion of a study during the search process to this point included: (a) participants being younger or older than high-school-aged (e.g., elementary students, college students, older adults), (b) the use of non-single-subject research methods (e.g., randomized controlled trials), (c) the subject of results to be about something other than writing interventions (e.g., writing-based interventions aimed at improving quality of life or counseling outcomes, rather than academic outcomes), or (d) the intervention did not target writing (e.g., delinquency interventions aimed at improving academics).

Overall, four studies were found in addition to the 16 studies reported by Rogers and Graham (2008) as using high school students. The reference lists of these four studies were searched further in an effort to identify additional studies that could meet the criteria. As a result, one additional study was found (Walker, Shippen, Houchins, & Cihak, 2007). Therefore, a total of 21 studies are included in this review (the 16 reported in Rogers and Graham’s original study and five subsequent studies). The same quality
indicators used by Rogers and Graham were applied to the studies selected for this review in order to facilitate comparison with the meta-analysis studies. Specifically, Rogers and Graham applied Horner et al.’s (2005) 21 quality indicators for single-subject research to the studies in their review. Rogers and Graham, however, rearranged these 21 indicators into 11 indicators. The quality indicators for single-subject research as described in the Rogers and Graham meta-analysis can be viewed in Table 4.

These 11 indicators were reviewed and discussed by the researcher and a graduate student, who acted as raters. Four of the studies from Rogers and Graham (2008) were randomly chosen and the two raters independently applied the 11 indicators as practice. One rater agreed 100% with two of the four ratings, was within 0.50 points of the third, and was within 1.09 points of the fourth. The other rater agreed 100% with three of the four ratings and was also within 1.09 points of the fourth study. The raters discussed these discrepancies. The researcher then rated the five new studies. The graduate student chose two of the studies at random, thus offering inter-rater reliability for 40% of the studies. The two raters agreed 100% on the ratings for the two studies. This procedure for establishing inter-rater reliability was similar to the procedures that Rogers and Graham.
Table 4

*Quality Indicators for Single-Subject Research*

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Point Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant description</td>
<td>One point gained for providing information on participants’ ages or grades, sex, socioeconomic status, ethnicity or race, pre-intervention writing achievement, and disability status. This total score was then divided by 6. The resulting score between 0.00 and 1.00 is then given for this indicator. If students with disabilities were not specifically included in the study, the sixth category would be removed and the total score would be divided by 5.</td>
</tr>
<tr>
<td>2. Participant selection</td>
<td>One point gained for describing how participants were selected in enough detail that another investigator would be able to reproduce the selection procedures.</td>
</tr>
<tr>
<td>3. Location description</td>
<td>One point gained for describing the location in which the intervention took place “in enough detail that it could easily be visualized” (Rogers &amp; Graham, 2008, p. 884).</td>
</tr>
<tr>
<td>4. Operational definition of dependent variable</td>
<td>One point gained for describing the methods for scoring the dependent variable in enough detail that another investigator could use the same methods. If more than one dependent variable was measured, all of them need to meet this indicator to receive a score of one. If only one of the dependent variables meets this indicator, a score of 0.50 is given. If none of the dependent variables meet this indicator, a score of 0.00 is given.</td>
</tr>
<tr>
<td>5. Inter-rater reliability</td>
<td>One point gained if the inter-rater reliability of the measures of all the dependent variables equaled or exceeded 0.60. If only some of the dependent variables meet this indicator, then the number of dependent variables meeting this indicator is divided by the total number of dependent variables. If none of the dependent variables meet this indicator, the study receives zero points for this indicator.</td>
</tr>
<tr>
<td>6. Multiple baseline data points</td>
<td>One point gained for all baseline measures containing at least three data points. If only some of the baseline measures meet this indicator, then the number of baseline measures meeting this indicator is divided by the total number of baseline measures. If none of the baseline measures meet this indicator, the study receives zero points for this indicator.</td>
</tr>
<tr>
<td>7. Multiple intervention data points</td>
<td>One point gained for all intervention measures containing at least three data points. If only some of the intervention measures meet this indicator, then the number of intervention measures meeting this indicator is divided by the total number of intervention measures. If none of the intervention measures meet this indicator, the study receives zero points for this indicator.</td>
</tr>
<tr>
<td>8. Treatment description</td>
<td>One point gained for describing the treatment (i.e., intervention) in sufficient detail as to allow another investigator to apply the treatment.</td>
</tr>
<tr>
<td>9. Treatment integrity</td>
<td>One point gained for a demonstration of treatment integrity.</td>
</tr>
<tr>
<td>10. Testing procedure description</td>
<td>One point gained for describing the procedures for administering the treatment in enough detail so that another investigator could apply the treatment in the same way.</td>
</tr>
<tr>
<td>11. Social validity or maintenance</td>
<td>One point gained for provision of data on social validity (from students, teachers, or other relevant parties) or for data showing that gains made during treatment were maintained for at least three weeks following the intervention.</td>
</tr>
</tbody>
</table>

*Note.* Adapted from Rogers and Graham (2008).
In their meta-analysis, Rogers and Graham (2008) described interventions with quality scores of 8.00 or greater as high quality. When a quality score fell below this mark, they cautioned against interpreting high PNDs as highly effective due to concerns about the study’s quality. The quality of each of the studies included in this review will be discussed and compared (see Table 5 for the studies’ demographic information and Table 6 for intervention descriptions, including the quality scores and outcomes of every study). Overall, the interventions in this review fell into six general categories: (a) strategy instruction, (b) direct instruction, (c) goal setting, (d) sentence construction, (e) reinforcement, and (f) prewriting, all of which had been identified by previous meta-analyses (Rogers & Graham, 2008). Of interest to the reader may be the qualitative indicators for the levels of effectiveness of the PND scores reported for each study: 50.0% or below is ineffective, 50.1% to 70.0% is small, 70.1% to 90.0% is moderate, and above 90.0% is large (Scruggs, Mastropieri, Cook, & Escobar, 1986).

**Strategy instruction.** Ten (roughly half) of the identified studies focused on strategy instruction. Strategy instruction interventions focus on teaching students specific strategies for different aspects of writing. By compiling information of descriptions from Harris and Graham (1996), Jacobson and Reid (2010), and Kiuahara et al. (2012), strategy instruction includes six stages: (a) develop background knowledge, (b) discuss it (c) model it, (d) memorize it, (e) support it, and (f) independent performance (see Table 7 for a detailed description of the six stages).
Table 5
Demographic Information of Single-Subject Writing Interventions with Secondary Students

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Grade (Age)</th>
<th>Sex (% male)</th>
<th>Race/Ethnicity</th>
<th>Setting</th>
<th>Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beals (1983, study 1)*</td>
<td>9</td>
<td>10 (15-16)</td>
<td>56</td>
<td>AA (n = 4)</td>
<td>Urban</td>
<td>SLD (n = 3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beals (1983, study 2)*</td>
<td>9</td>
<td>10 (15-16)</td>
<td>56</td>
<td>AA (n = 4)</td>
<td>Urban</td>
<td>SLD (n = 3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bording et al. (1984)*</td>
<td>9</td>
<td>NR (12-16)</td>
<td>100</td>
<td>CA</td>
<td>NR</td>
<td>EBD (n = 5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SLD (n = 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CD (n = 2)</td>
</tr>
<tr>
<td>Dowell et al. (1994)*</td>
<td>3</td>
<td>9 to 11 (14-16)</td>
<td>67</td>
<td>NR</td>
<td>Urban</td>
<td>SLD</td>
</tr>
<tr>
<td>Hopman &amp; Glynn (1989)*</td>
<td>4</td>
<td>HS (13)</td>
<td>100</td>
<td>NR</td>
<td>Urban</td>
<td>N/A</td>
</tr>
<tr>
<td>Jacobson &amp; Reid (2010)</td>
<td>3</td>
<td>11 to 12 (NR)</td>
<td>100</td>
<td>CA</td>
<td>NR</td>
<td>ADHD</td>
</tr>
<tr>
<td>Kiuahara et al. (2012)</td>
<td>6</td>
<td>10 (15-16)</td>
<td>67</td>
<td>AA/CA</td>
<td>Suburban</td>
<td>OHI/ADHD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(n = 1)</td>
<td></td>
<td>(n = 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AA/HA</td>
<td></td>
<td>EBD (n = 1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(n = 1)</td>
<td></td>
<td>DD/SLI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 3)</td>
<td></td>
<td>(n = 1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HA (n = 1)</td>
<td></td>
<td>SLD (n = 2)</td>
</tr>
<tr>
<td>McNaughton et al. (1997)*</td>
<td>3</td>
<td>10 &amp; 12 (16-18)</td>
<td>33</td>
<td>NR</td>
<td>Rural</td>
<td>SLD</td>
</tr>
<tr>
<td>Moran et al. (1981, study 1)*</td>
<td>3</td>
<td>8 to 9 (14-16)</td>
<td>67</td>
<td>NR</td>
<td>Suburban</td>
<td>SLD</td>
</tr>
<tr>
<td>Moran et al. (1981, study 2)*</td>
<td>5</td>
<td>8 to 9 (13-16)</td>
<td>80</td>
<td>NR</td>
<td>Suburban</td>
<td>SLD</td>
</tr>
<tr>
<td>Newstrom et al. (1999)*</td>
<td>1</td>
<td>9 (NR)</td>
<td>100</td>
<td>NR</td>
<td>NR</td>
<td>EBD</td>
</tr>
<tr>
<td>Seabaugh &amp; Schumaker (1994)*</td>
<td>3</td>
<td>9 to 12 (14-16)</td>
<td>33</td>
<td>NR</td>
<td>NR</td>
<td>SLD (n = 2)</td>
</tr>
<tr>
<td>Schmidt (1983)*</td>
<td>7</td>
<td>10 to 12 (14-16)</td>
<td>86</td>
<td>CA</td>
<td>Suburban</td>
<td>SLD</td>
</tr>
<tr>
<td>Schmidt et al. (1988)*</td>
<td>7</td>
<td>10 to 12 (14-16)</td>
<td>86</td>
<td>CA</td>
<td>Suburban</td>
<td>SLD</td>
</tr>
</tbody>
</table>
Table 5 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age Range</th>
<th>Grade</th>
<th>Race Code</th>
<th>Ethnicity</th>
<th>Rural</th>
<th>Suburban</th>
<th>Demographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schumaker et al. (1982)*</td>
<td>9</td>
<td>8 to 12 (12-18)</td>
<td>78</td>
<td>NR</td>
<td>NR</td>
<td>SLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonntag &amp; McLaughlin (1984)*</td>
<td>6</td>
<td>8 to 9 (14-16)</td>
<td>33</td>
<td>NR</td>
<td>Rural</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thanhouser (1994)**</td>
<td>6</td>
<td>12 (17-18)</td>
<td>UN</td>
<td>AA (n = 1)</td>
<td>Suburban</td>
<td>SLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Houten &amp; McKillop (1977)*</td>
<td>38</td>
<td>10 &amp; 11 (NR)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viel-Ruma et al. (2010)</td>
<td>6</td>
<td>9 to 11 (14-17)</td>
<td>66</td>
<td>AA (n = 1)</td>
<td>Urban</td>
<td>SLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HA (n = 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walker et al. (2005)</td>
<td>3</td>
<td>NR (15-16)</td>
<td>33</td>
<td>AA (n = 2)</td>
<td>Urban</td>
<td>SLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HA (n = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walker et al. (2007)</td>
<td>3</td>
<td>9 (15)</td>
<td>33</td>
<td>AA (n = 2)</td>
<td>Urban</td>
<td>SLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CA (n = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. NR = not reported; N/A = not applicable; HS = High School (no grade reported); AA = African American; CA = Caucasian American; HA = Hispanic American; SLD = Specific Learning Disability (for consistency, students reported as Learning Disability are referred to as SLD); EBD = Emotional Behavioral Disorder (i.e., emotional disturbance or severe behavior handicap); CD = Cognitive Disability; ADHD = Attention-Deficit Hyperactivity Disorder (for consistency, students reported as Attention Deficit Disorder are referred to as ADHD); DD/SLI = Developmental Disability and Speech Language Impairment; UN = Demographic unknown due to dissertation being unavailable for loan from home university.

*Information from studies included in Rogers & Graham (2008).
**Dissertation unavailable from home university; information reported here in modified format from a similar table in Rogers & Graham (2008), p. 888-891.
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Design</th>
<th>Quality Score (0-11)</th>
<th>PND (phase: outcome measure)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beals (1983, study 1)*</td>
<td>Strategy instruction/ SRSD (editing &amp; sentence writing)</td>
<td>Multiple-baseline: across skills</td>
<td>7.75</td>
<td>71% (post: errors corrected)</td>
</tr>
<tr>
<td>Beals (1983, study 2)*</td>
<td>Strategy instruction/ SRSD (sentence writing)</td>
<td>Multiple-baseline: across skills</td>
<td>7.75</td>
<td>100% (int: complete sentences)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>88% (post: complete sentences)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>89% (int: complicated sentences)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80% (post: complicated sentences)</td>
</tr>
<tr>
<td>Bording et al.</td>
<td>Reinforcement contingent on skill performance</td>
<td>Multiple-baseline: across paired participants</td>
<td>8.83</td>
<td>57% (int: grammar)</td>
</tr>
<tr>
<td>(1984)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dowell et al.</td>
<td>Strategy instruction/ SRSD (paragraph construction)</td>
<td>Multiple-baseline: across participants, combined with alternating treatments</td>
<td>5.58</td>
<td>34% (int: % descriptive words)</td>
</tr>
<tr>
<td>(1994)*</td>
<td></td>
<td></td>
<td></td>
<td>4% (int: quality)</td>
</tr>
<tr>
<td>Hopman &amp; Glynn (1989)*</td>
<td>Goal setting for productivity</td>
<td>Changing criterion</td>
<td>6.4</td>
<td>26% (int: TWW)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63% (post: TWW)</td>
</tr>
<tr>
<td>Jacobson &amp; Reid (2010)</td>
<td>Strategy instruction/ SRSD (planning/ drafting)</td>
<td>Multiple-baseline: across participants</td>
<td>9.43</td>
<td>100% (int: elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (maint: elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (int: TWW)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67% (maint: TWW)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (int: holistic)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67% (maint: holistic)</td>
</tr>
<tr>
<td>Study (Year)</td>
<td>Intervention Type</td>
<td>Baseline Design</td>
<td>General &amp; Specific Measures</td>
<td>Percent Correct in Int.</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Kiuhara et al. (2012)</td>
<td>Strategy instruction/ SRSD (planning/drafting)</td>
<td>Multiple-baseline: across paired participants and multiple probe</td>
<td>9.58</td>
<td>90% (int: essential elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>86% (int: functional elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>71% (int: holistic)</td>
</tr>
<tr>
<td>McNaughton et al. (1997)*</td>
<td>Strategy instruction/ SRSD (editing)</td>
<td>Multiple-baseline: across participants and multiple probe</td>
<td>7.67</td>
<td>95% (int: errors corrected)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (maint: errors corrected)</td>
</tr>
<tr>
<td>Moran et al. (1981, study 1)*</td>
<td>Strategy instruction/ SRSD (paragraph construction)</td>
<td>Multiple-baseline: across skills</td>
<td>6.67</td>
<td>100% (int: elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (int: elements)</td>
</tr>
<tr>
<td>Moran et al. (1981, study 2)*</td>
<td>Strategy instruction/ SRSD (paragraph construction)</td>
<td>Multiple-baseline: across participants</td>
<td>7.7</td>
<td>100% (int: elements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100% (int: elements)</td>
</tr>
<tr>
<td>Newstrom et al. (1999)*</td>
<td>Reinforcement</td>
<td>Multiple-baseline: across skills</td>
<td>5.5</td>
<td>100% (int: grammar)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60% (int: lessons completed)</td>
</tr>
<tr>
<td>Seabaugh &amp; Schumaker (1994)*</td>
<td>Goal setting for productivity</td>
<td>Multiple-baseline: across participants and skills</td>
<td>10.83</td>
<td>50% (post: errors corrected)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29% (int: elements)</td>
</tr>
<tr>
<td>Schmidt (1983)*</td>
<td>Strategy instruction/ SRSD (editing)</td>
<td>Multiple-baseline: across skills</td>
<td>7</td>
<td>78% (int: complete sentences)</td>
</tr>
<tr>
<td>Schmidt et al. (1988)*</td>
<td>Sentence construction</td>
<td>Multiple-baseline: across strategies</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Study (Year)</td>
<td>Intervention</td>
<td>Design</td>
<td>Strategy Instruction</td>
<td>Post Baseline Measure</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>--------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Schumaker et al. (1982)*</td>
<td>Strategy instruction/ SRSD (editing)</td>
<td>Multiple-baseline: across participants</td>
<td>6.34</td>
<td>100% (post: errors corrected)</td>
</tr>
<tr>
<td>Sonntag &amp; McLaughlin (1984)*</td>
<td>Strategy instruction/SRSD (paragraph construction)</td>
<td>Multiple-baseline: across groups</td>
<td>5.40</td>
<td>100% (post: elements)</td>
</tr>
<tr>
<td>Thanhouser (1994)**</td>
<td>Prewriting activities</td>
<td>Multiple-baseline: UN</td>
<td>8.83</td>
<td>44% (int: quality)</td>
</tr>
<tr>
<td>Van Houten &amp; McKillop (1977)*</td>
<td>Goal setting for productivity</td>
<td>ABAB</td>
<td>6.4</td>
<td>91% (int: TWW)</td>
</tr>
<tr>
<td>Viel-Ruma et al. (2010)</td>
<td>Direct Instruction: Expressive Writing program</td>
<td>Multiple-baseline: Multiple probe across participants</td>
<td>7.83</td>
<td>60% (int: %CWS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53% (int: TWW)</td>
</tr>
<tr>
<td>Walker et al. (2005)</td>
<td>Direct instruction: Expressive Writing program</td>
<td>Multiple-baseline: multiple probe across participants</td>
<td>9.33</td>
<td>95% (int: CWS)</td>
</tr>
<tr>
<td>Walker et al. (2007)</td>
<td>Direct instruction: Expressive Writing program</td>
<td>Multiple-baseline: multiple probe across participants</td>
<td>9.17</td>
<td>100% (int: CWS)</td>
</tr>
</tbody>
</table>

Note. PND = Percent Nonoverlapping Data; SRSD = self-regulated strategy development; post = post baseline measurements, can include data from intervention and/or maintenance phases and/or another post-intervention phase; int = data from intervention period only; maint = data from maintenance period only; TWW = Total Words Written; %CWS = percent Correct Writing Sequences; CWS = Correct Writing Sequences; UN = Demographic unknown due to dissertation being unavailable for loan from home university; NR = not reported; ABAB = baseline-intervention-baseline-intervention reversal design.

*Information from studies included in Rogers & Graham (2008), p. 888-891.

**Dissertation unavailable from home university; information reported here in modified format from a similar table in Rogers & Graham (2008), p. 888-891.
Table 7

*Six Stages of Self-Regulated Strategy Development*

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Develop Background Knowledge</td>
<td>The instructor introduces the background knowledge necessary for learning the strategies. The instructor explains the purpose for the strategies and engages in a discussion with the students about the topic. Some introduce the specifics of the strategy in the first stage while others introduce the specifics in the second stage.</td>
</tr>
<tr>
<td>2. Discuss It</td>
<td>The instructor and the students discuss the purpose of the strategy again, are introduced to devices that will assist them in learning the strategy, and are taught self-regulation strategies. Students are often given mnemonic devices, graphic organizers, and goal planning sheets. They evaluate previous examples of their own work and set goals for their next work.</td>
</tr>
<tr>
<td>3. Model It</td>
<td>The instructor explicitly models the strategy and demonstrates appropriate self-instruction though thinking aloud. The instructor reviews the strategy, sets goals, makes a plan for what he or she is going to write, and then writes, all with assistance from the students. The instructor models self-statements, planning, self-evaluation, self-reinforcement, and coping skills throughout this process.</td>
</tr>
<tr>
<td>4. Memorize It</td>
<td>The students memorize each step of the strategies they learned. They practice together and with the teacher until all the students are able to demonstrate that they can repeat the strategies.</td>
</tr>
<tr>
<td>5. Support It</td>
<td>Student independence is increased. Students practice using the strategies by planning, writing, setting goals, and evaluating success on their own. They may start out working together in peer groups, move to independent work with teacher scaffolding, and finally move to completely independent use of the strategies.</td>
</tr>
<tr>
<td>6. Independent Performance</td>
<td>Students practice independently using the strategies in different settings and with various tasks over time.</td>
</tr>
</tbody>
</table>

*Note.* Compiled from descriptions in Harris and Graham (1996), Jacobson and Reid (2010), and Kiuhara et al. (2012).

*Self-regulated strategy development.* All of the strategy instruction interventions followed methods for self-regulated strategy development (SRSD). Within SRSD,
students are taught skill-specific strategies, self-regulation strategies, and declarative, procedural, and conditional knowledge about the strategies (Harris & Graham, 1996). The students actively engage with the teachers and their peers. SRSD incorporates goal setting and uses these goals to provide individualized instruction to each student. Progress through the six stages of SRSD is criterion-based, meaning that students only progress to the next stage if they have met certain criteria, not simply because the time has come to move onto the next stage. Throughout each stage, the teacher scaffolds student learning, gradually helping to increase student independence and mastery of the strategy.

Editing. Four studies emerged within this review that taught students how to edit their writing (Beals, 1983; McNaughton, Hughes, & Ofiesh, 1997; Schmidt, 1983; Schumaker et al., 1982). The outcome measures for all four studies included calculating the number of errors corrected. Additional data including complete and complicated sentences was collected by the Beals (1983) study, as were essay elements (i.e., the number of essay components that a writing contained, such as a topic sentence) in the Schmidt (1993) study. The outcomes of all the studies ranged from ineffective (PND = 29%) to large (PND = 100%) impacts on student performance among post-baseline conditions for all outcome measures. The median PND among post-baseline conditions for errors corrected was 98%, indicating a general positive effect. While strategy instruction for editing interventions overall evidenced large beneficial effects, the low quality scores these studies received raises concerns about the integrity of the studies’ designs and impact of their findings. The quality scores of the interventions ranged from
6.34 to 7.75, with a median score of 7.34. Rogers and Graham (2008) noted concerns about the poor descriptions of participants and settings in two of the studies (e.g., McNaughton et al., 1997; Schumaker et al., 1982), failure to establish experimental control by some of the studies (e.g., Beals, 1983; Schmidt, 1983; and Schumaker et al., 1982, reported only post-treatment outcomes, and no data during the interventions), and uncertainty about treatment fidelity in some of the studies (which studies were not specified).

Paragraph construction. Four of the SRSD studies taught students strategies for constructing paragraphs and were reported in three sources (Dowell, Storey, & Gleason, 1994; Moran, Schumaker, & Vetter, 1981; Sonntag & McLaughlin, 1984). These studies taught students how to write and/or organize expository or descriptive paragraphs and measured vocabulary, errors, and the number of writing elements for basic parts of a paragraph. Dowell et al.’s (1994) intervention demonstrated no effect on significant performance, with PNDs under 50% for both productivity and vocabulary. Moran et al. (1981) investigated paragraph elements in both of the studies in their report, revealing all PNDs of 100% (large effects). Sonntag and McLaughlin (1984) examined errors corrected and errors per word, demonstrating the intervention was highly effective as evidenced by all PNDs of 100%. Despite the positive impacts of these studies, the quality scores for all three studies ranged from 5.40 to 7.70, with a median score of 6.13. Taken together, results of these four studies indicated that strategy instruction for paragraph construction interventions are effective at improving the number of paragraph elements included, improving the number of errors corrected, and diminishing total errors
per word, though the poor quality scores highlight concerns about the integrity of the studies’ designs. Considerable caution must be used when interpreting the results because none of these studies established experimental control and the descriptions of participants and settings, selection procedures, treatment integrity, and social validity were poor or unclear (Horner et al., 2005; Rogers & Graham, 2008).

Planning/drafting. Two studies taught students strategies for planning/drafting (Jacobson & Reid, 2010; Kiuhara et al., 2012). Specifically, these studies focused on teaching students strategies for planning and drafting persuasive essays. The outcome measures for both studies concentrated on number of essay elements (i.e., parts of an essay, such as topic sentence, supporting evidence) and holistic quality (i.e., rating on a holistic rubric). The median PND among the post-baseline conditions for number of elements was 100% (range = 86% to 100%), demonstrating a large effect. For holistic quality, the median PND among post-baseline conditions was 76% (range = 67% to 100%), indicative of a moderate effect. The Jacobson and Reid (2010) study also included a measure of TWW (median PND = 84%). Overall the quality scores of the interventions in both studies were high (9.43 and 9.58). Both studies failed to describe the socioeconomic status of the participants and provide inter-rater reliability for one or more dependent variables; the descriptions of the settings in which the interventions took place were lacking in detail. Taken together, strategy instruction for planning/drafting interventions were of high quality and were effective at improving the number of essay elements included, improving scores on holistic rubric measures, and increasing TWW.
Direct instruction: Expressive Writing. Three of the studies directly taught writing skills to students through the Expressive Writing program (Viel-Ruma et al., 2010; Walker et al., 2005; Walker et al., 2007). In Direct Instruction (DI) programs, students receive fast-paced, progressive, and focused lessons in small group settings (Adams & Engelmann, 1996). Furthermore, students are provided with multiple opportunities to respond, and immediate feedback and correction procedures are provided. Within DI programs, there are three stages of instruction: (a) model the correct response, (b) lead students in choral responding, and (c) test student knowledge through immediate and delayed probes. In a typical DI program, instructors question the students, correct mistakes, review and practice skills, and give examples and demonstrations with an overall goal of student mastery of skills. In the Expressive Writing DI program by Engelmann and Silbert (as cited in Viel-Ruma et al., 2010), instruction follows the DI model and focuses on grammar, usage, punctuation, the stages of writing, and narrative writing. Viel-Ruma et al. (2010) used %CWS as an outcome measure in their study, revealing a median PND of 65% among post-baseline phases, which is small. They also investigated TWW, which had a 52% median PND among post-baseline phases, which is small and very close to ineffective. In the subsequent Walker et al. (2005, 2007) studies, the outcome measure was CWS. The median PND across post-baseline phases in both studies was 100%, representing a large effect. Overall, the quality scores of the direct instruction studies were high, ranging from 7.83 to 9.33 (median = 9.17). Although selection procedures were not detailed, the socioeconomic status of the participants was not described, and there was no inter-rater reliability reported for some dependent
variables, the *Expressive Writing* DI program appears to be a relatively strong approach for teaching writing.

**Goal setting.** Three of the studies had the students engage in setting goals for writing productivity (Hopman & Glynn, 1989; Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977). In goal setting writing interventions, students set goals for their written products (Graham & Perin, 2007a). Goals are set based on previous performance and are compared to current performance, as in behavior interventions (Bandura, 1977). Writing goals often include curriculum-based assessment outcome measures (e.g., TWW, CWS), but they also can focus on adding certain information, writing certain types of paragraphs, or including certain genre elements (Graham & Perin, 2007a). Participants set goals alone, set goals in conjunction with the researcher, or are assigned a goal (Locke, Shaw, Saari, & Latham, 1981). Goals are most effective when they are combined with reinforcement for goal attainment and feedback is provided as to the participant’s progress toward the goal (Locke et al., 1981). In fact, feedback combined with student-set goals and self-graphing has been demonstrated to improve academic performance (Codding, Lewandowski, & Eckert, 2005). Perhaps unsurprisingly, then, goal setting writing interventions often include other components from effective behavior change interventions, such as performance feedback, self-graphing, and reinforcement.

Two of the goal setting studies used TWW (Hopman & Glynn, 1989; Van Houten & McKillop, 1977) and one used total number of writing lessons completed as the outcome measure (Seabaugh & Schumaker, 1994). The median PND for TWW among
the post-baseline phases was 63% (range = 26% to 91%), a small effect. The PND for total number of writing lessons completed was 60%, a small effect. The quality scores of these studies ranged from 6.40 to 10.83, with a mean score of 7.88. The Seabaugh and Schumaker (1994) study received the fourth highest quality rating out of the 76 studies reported in Rogers and Graham’s (2008) meta-analysis. Overall, these effect sizes suggest that goal setting had ineffective to large effects on student writing products. The studies with lower quality scores generally contained limited data on participants, setting, fidelity, and validity, while the Seabaugh and Schumaker (1994) study contained most of these data (Rogers & Graham, 2008).

**Sentence construction.** Two of the studies taught students how to write different types of sentences (Beals, 1983; Schmidt, Deshler, Schumaker, & Alley, 1988). Sentence construction interventions can teach students methods for combining sentences or strategies for constructing sentences. These two studies taught students strategies for constructing sentences. Both studies used percent complete or complicated sentences as the outcome measure. The median PND for percent complete or complicated sentences among post-baseline conditions in these studies was 84% (range = 50% to 100%), indicating a moderate effect. The quality scores of these studies were 7.00 and 7.75. Rogers and Graham (2008) cautioned that these studies generally did not meet criteria for participant selection, fidelity of treatment, and testing procedure description.

**Reinforcement.** Two studies reinforced writing performance through external means (Bording, McLaughlin, & Williams, 1984; Newstrom, McLaughlin, & Sweeney, 1999). Students could earn free time or other rewards based on performance with a
contingency contract. Outcome measures focused on improving grammar by measuring percent correct capitalization and percent correct punctuation. These studies had small (PND = 57%) and large (PND = 100%) effects on improving grammar, with an overall median PND of 79%, which is moderate. The quality scores were 5.50 and 8.83, with a median score of 7.17. These studies lacked some demographic data and one study was of low quality but had a large effect (Newstrom et al., 1999), while the other was of high quality and had a small effect (Bording et al., 1984), so the findings should be interpreted with caution (Rogers & Graham, 2008).

**Prewriting activities.** One dissertation by Thanhouser (as cited in Rogers and Graham, 2008) had students engage in prewriting activities. The home university for this dissertation did not release the dissertation through an interlibrary loan request. As such, the data reported here is the data that was reported by Rogers and Graham (2008). In this study, students were taught to use a graphic organizer to assist with idea generation when engaging in persuasive writing. The outcome measure for writing quality, using a holistic rubric, was similar to other studies with younger students reported in Rogers and Graham’s study. Thanhouser also evaluated writing product. The PND revealed that the intervention was ineffective (PND = 44%) for improving writing quality and had a small effect (PND = 64%) on improving the writing product, even though the study was of high quality (score = 8.83).

**Critique and Analysis of the Research**

The studies in this literature review have been ranked following from Rogers and Graham’s (2008) method of ranking of the findings in their meta-analysis. First, studies
were ranked individually according to effect size. The median effect size of all outcome measures reported in the study was used to rank the study; the mean effect size is also reported. The median rather than the mean effect size was used to reduce the influence of outlying effect sizes. When studies had identical effect sizes, they were then ranked according to quality score. The rankings of all studies are provided in Table 8.

**Effective studies.** Ten studies had large effects (Jacobson & Reid, 2010; McNaughton et al., 1997; both studies in Moran et al., 1981; Newstrom et al., 1999; Schumaker et al., 1982; Sonntag & McLaughlin, 1984; Van Houten & McKillop, 1977; Walker et al., 2005, 2007). These studies investigated the effectiveness of strategy instruction, direct instruction, goal setting, and reinforcement interventions in improving writing performance of high school students. Less than one third of these studies, however, had quality scores above 8.00 (Jacobson & Reid, 2010; Walker et al., 2005, 2007). Three studies had moderate effects (both interventions in Beals, 1983; Kiuhara et al., 2012). These studies investigated the use of strategy instruction and sentence construction interventions in improving the writing performance of high school students. Only one of these studies had quality a score above 8.00 (Kiuhara, 2010).
<table>
<thead>
<tr>
<th>Rank</th>
<th>Study</th>
<th>Intervention</th>
<th>Quality Score</th>
<th>Median PND for All DV</th>
<th>Average PND for All DV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jacobson &amp; Reid (2010)</td>
<td>Strategy instruction/SRSD (planning/drafting)</td>
<td>9.43</td>
<td>100</td>
<td>89</td>
</tr>
<tr>
<td>2</td>
<td>Walker et al. (2007)</td>
<td>Direct Instruction: <em>Expressive Writing</em> program</td>
<td>9.17</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>Moran et al. (1981, study 2)</td>
<td>Strategy instruction/SRSD (paragraph construction)</td>
<td>7.7</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>Moran et al. (1981, study 1)</td>
<td>Strategy instruction/SRSD (paragraph construction)</td>
<td>6.67</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>Schumaker et al. (1982)</td>
<td>Strategy instruction/SRSD (editing)</td>
<td>6.34</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>Newstrom et al. (1999)</td>
<td>Reinforcement</td>
<td>5.5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>Sonntag &amp; McLaughlin (1984)</td>
<td>Strategy instruction/SRSD (paragraph construction)</td>
<td>5.4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>Walker et al. (2005)</td>
<td>Direct Instruction: <em>Expressive Writing</em> program</td>
<td>9.33</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td>9</td>
<td>McNaughton et al. (1997)</td>
<td>Strategy instruction/SRSD (editing)</td>
<td>7.67</td>
<td>95</td>
<td>72</td>
</tr>
<tr>
<td>10</td>
<td>Van Houten &amp; McKillop (1977)</td>
<td>Goal setting</td>
<td>6.4</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>11</td>
<td>Beals (1983, intervention 2)</td>
<td>Sentence construction</td>
<td>7.75</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>12</td>
<td>Kiuhara et al. (2012)</td>
<td>Strategy instruction/SRSD (planning/drafting)</td>
<td>9.58</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>13</td>
<td>Beals (1983, intervention 1)</td>
<td>Strategy instruction/SRSD (editing)</td>
<td>7.75</td>
<td>88</td>
<td>86</td>
</tr>
<tr>
<td>14</td>
<td>Schmidt et al. (1988)</td>
<td>Sentence construction</td>
<td>7</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>15</td>
<td>Seabaugh &amp; Schumaker (1994)</td>
<td>Goal setting</td>
<td>10.83</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>16</td>
<td>Bording et al. (1984)</td>
<td>Reinforcement</td>
<td>8.83</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>17</td>
<td>Viel-Ruma et al. (2010)</td>
<td>Direct Instruction: <em>Expressive Writing</em> program</td>
<td>7.83</td>
<td>57</td>
<td>58</td>
</tr>
<tr>
<td>18</td>
<td>Thanhouzer (1994)</td>
<td>Prewriting</td>
<td>8.83</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>19</td>
<td>Hopman &amp; Glynn (1989)</td>
<td>Goal setting</td>
<td>6.4</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>20</td>
<td>Schmidt (1983)</td>
<td>Strategy instruction/SRSD (editing)</td>
<td>7</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>21</td>
<td>Dowell et al. (1994)</td>
<td>Strategy instruction/SRSD (paragraph construction)</td>
<td>5.58</td>
<td>38</td>
<td>38</td>
</tr>
</tbody>
</table>
Five studies had small effects (Bording et al., 1984; Seabaugh & Schumaker, 1994; Schmidt et al., 1988; Thanhouser, as cited in Rogers & Graham, 2008; Viel-Ruma et al., 2010). These studies investigated the use of sentence construction, goal setting, direct instruction, and prewriting interventions in improving the writing performance of high school students. In contrast to the more effective studies, more than half of the studies with small effect sizes were of high quality (Bording et al., 1984; Seabaugh & Schumaker, 1994; Thanhouser, as cited in Rogers & Graham, 2008).

**Ineffective studies.** Three studies were ineffective at improving student performance (Dowell et al., 1994; Hopman & Glynn, 1989; Schmidt, 1983). These studies used strategy instruction and goal setting interventions. Their poor quality scores indicated that there were several concerns about the quality of their research designs.

**Effectiveness by intervention type.** The most effective interventions overall were strategy instruction (Beals, 1983; Dowell et al., 1984; Jacobson & Reid, 2010; Kiuhara et al., 2012; McNaughton et al., 1997; Moran et al., 1981; Schmidt, 1983; Schumaker et al., 1982; Sonntag & McLaughlin, 1984) and direct instruction (Viel-Ruma et al., 2010; Walker et al., 2005; Walker et al., 2007). The median PNDs for post-baseline errors corrected for strategy instruction on both planning/drafting and paragraph writing were 100%. The median post-baseline PND for errors corrected for strategy instruction on editing was 98%. The median post-baseline PND for CWS for the direct instruction interventions was 100%. Additionally, the strategy instruction on planning/drafting and the direct instruction interventions were both of high quality (median quality scores = 9.51 and 9.17, respectively).
The sentence construction (Beals, 1983; Schmidt et al., 1988) and reinforcement interventions (Bording et al., 1984; Newstrom et al., 1999) were moderately effective overall. The median post-baseline PND for complete and complicated sentences was 84%, and the median post-baseline PND for grammar for the reinforcement studies was 79%. Caution should be used when interpreting the results of these studies, however, based on median quality scores of 7.38 and 7.17, respectively.

The prewriting (Thanhouser, as cited in Rogers & Graham, 2008) and goal setting (Hopman & Glynn, 1989; Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977) interventions had small effect sizes overall. The median post-baseline PND for the writing quality outcome measure on the prewriting intervention was 44% (ineffective), while the median post-baseline PND for the writing product outcome measure for the prewriting intervention was 64% (small effect). The prewriting intervention was of high quality. The median post-baseline production PND for the goal setting interventions was 62%. Caution should be used when interpreting the results of the goal setting interventions overall, however, because of the median quality score of 6.40 for these interventions. Two of the goal setting studies had quality scores of 6.40 and one study had a quality score of 10.83, indicating a large difference in the quality of the studies’ designs.

**Overall Recommendations**

From the extant literature, the best interventions for high school students who are struggling with writing appear to be strategy instruction on planning/drafting (e.g., Jacobson & Reid, 2010) and direct instruction through the *Expressive Writing* program
(e.g., Walker et al., 2007). These studies not only demonstrated strong magnitude of effect, but also were of high quality. The two most well-designed interventions were a goal setting intervention (Seabaugh & Schumaker, 1994) and strategy instruction focused on planning/drafting (Kiuhara et al., 2012). Other effective and well-designed studies also used the *Expressive Writing* program (Walker et al., 2005), implemented a reinforcement intervention (Bording et al., 1984), and focused on prewriting skills (Thanhouser, as cited in Rogers & Graham, 2008). As such, effective interventions for struggling high school students would include strategy instruction, direct instruction, goal setting, reinforcement, and prewriting activities. Interestingly, the strategy instruction interventions included all of these components through the use of the self-regulated strategy development model. This model incorporates goal setting, self-reinforcement, prewriting, and techniques similar to the direct instruction model (e.g., mnemonic devices, practice, feedback, correction, modeling, leading, assessing, or skill mastery).

As mentioned earlier, however, some high schools may not have the resources to devote enough time in their schedules to provide strategy instruction interventions to struggling students. The strategy instruction and direct instruction interventions require multiple hours of training, full-length class periods, and many weeks in order to be implemented with fidelity. Moreover, these interventions appear to be most appropriate for use in the universal or targeted levels of intervention support, as opposed to more individualized supports necessary to rapidly increase a student’s repertoire of skills. For high schools interested in finding quick, easy to implement, and effective interventions at
the individualized level of intervention support, goal setting, reinforcement, or prewriting interventions emerge as the most appropriate.

Although these studies examining goal setting, reinforcement, or prewriting interventions demonstrated small effects, their limited outcomes may be due in large part to the outcome measures which they employed. For example, the outcome measure used for the Seabaugh and Schumaker’s (1994) goal setting intervention likely was not applicable to a majority of high school students because the outcome was dependent upon the setting in which the intervention took place. The school was described by Seabaugh and Schumaker as a private alternative school setting, where students were able to complete lessons at their own pace, including coming to and going from school as they pleased. Most high school students are not afforded those options; most are required to attend classes on time and complete work within the pace of the peers in their classrooms. The intervention taught the students self-reinforcement, self-recording, self-evaluation, and behavior contracting, which are also skills that may not routinely be taught to typical high school students with no behavioral concerns. As part of behavior contracting, students were to set goals for how many lessons to complete, so the intervention used the total number of lessons completed as the outcome measure. Setting goals for the number of lessons completed may not be appropriate in a more traditional school setting because students typically follow the teacher and their classmates in completing lessons on a set schedule. Most students do not usually have control over when a lesson is completed. Setting goals for number of lessons completed also does not directly measure writing performance, making it a less valid goal to set for an academic intervention. When
intervening with a specific academic skill, an outcome measure that directly reflects that skill should be used. Simply completing a certain number of lessons completed is not a direct measure of writing ability, which calls into question the validity of this outcome measure.

Similarly, the Hopman and Glynn (1989) and Van Houten and McKillop (1977) studies also used an outcome measure that is less valid for high school students. Their goal setting intervention studies utilized TWW as their outcome measure, which is a less reliable and valid measure of writing performance for high school students (Amato & Watkins, 2011; Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). Goal setting interventions for high school students should focus on the most reliable and valid progress monitoring writing CBA outcome measures, such as CWS and CIWS [or promising outcome measures such as MLCWS, incorrect writing sequences (IWS), correct punctuation marks (CPM), characters/word, and total sentences]. To date, no such studies have been conducted.

With regard to reinforcement interventions, most prior studies utilized percent correct capitalization and percent correct punctuation as outcome measures (e.g., Bording et al., 1984; Newstrom et al., 1999). Outcome measures that utilize percentages are not as sensitive to change over time (Tindal & Parker, 1991). Counting correct capitalization and punctuation, however, may be reliable and valid measures for assessing writing performance of high school students, though few studies exist that use CPM and no studies exist that use correct capitalization (Diercks-Gransee et al., 2009).
Thanhouser (as cited in Rogers & Graham, 2008) used a holistic rubric as an outcome measure for a prewriting intervention. Holistic rubrics are effective at providing pre/post intervention data, but not for progress monitoring. Rubrics have a fixed set of scores and are rated subjectively, so they will not show growth over time because they cannot measure small changes in performance.

An effective individualized or small group intervention for struggling high school students would have the following qualities: (a) incorporate techniques from effective and well-designed strategy instruction, direct instruction, goal setting, reinforcement, and prewriting interventions; (b) fit effortlessly into a student’s daily schedule; (c) be quick, inexpensive, easy to implement, and effective; and (d) measure skills appropriate for high school students (L. S. Fuchs, Fuchs, & Compton, 2010; Samuels, 2009; Vaughn & Fletcher, 2010; Windram, Scierka, & Silberglitt, 2010).

**Future Directions for Writing Interventions for High School Students**

To summarize, a review of the literature indicates that in addition to a call for more single-subject research to inform the use of interventions in the schools (e.g., Horner et al., 2005; Kratochwill & Stoiber, 2000a, 2000b, 2000c, 2002), there exists a need to further explore the effectiveness and utility of writing interventions for high school students struggling with writing (e.g., Graham & Perin, 2007a, 2007b; Rogers & Gram, 2008). First, only 21 studies were found that addressed this specific topic, which leaves little room for determining evidence-based practices for high school students struggling with writing. All areas of research on writing interventions for struggling high school students are in need of more investigation. While interventions utilizing SRSD
and strategy instruction are the most numerous (43% of studies in this review), the breakdown of how many studies focused on a different strategy is still limited. There were four studies on editing, four on paragraph-writing, and two on planning/drafting. The other interventions discussed in this review were on direct instruction (three studies specifically on the *Expressive Writing* program), goal setting (three studies), sentence construction (two studies), reinforcement (two studies), and prewriting activities (one study). With a range of only one to four studies per type of intervention (some of which were ineffective or had poor quality scores), research in the area of writing interventions for struggling high school students needs to be expanded.

Second, the Rogers and Graham (2008) meta-analysis included other intervention techniques for which no studies using high school students were found during this literature review. Such strategies include teaching grammar/usage, word processing, and self-monitoring. Coker and Lewis (2008) pointed out that the *Writing Next* research report and meta-analysis by Graham and Perin (2007a, 2007b), which was a predecessor to the Rogers and Graham meta-analysis, failed to include many other possibly effective interventions, specifically those written by educators that were published within books (e.g., Atwell, 1998). Because research has not been conducted to support these interventions, they were not included in the meta-analysis or this literature review. A lack of research studies does not mean that these interventions are ineffective. Rather, studies should be conducted that seek to demonstrate the effectiveness of these interventions. Unfortunately, no research studies on these interventions were found by
Rogers and Graham (2008), through the search for the present study, or through the WWC website.

Third, high quality effective studies from this literature review revealed that: (a) strategy instruction and direct instruction interventions are effective at improving writing skills for high school students on a universal or targeted level of intervention support, and (b) goal setting, reinforcement, and prewriting interventions are effective at improving writing skills for high school students on an individualized level of intervention support. As such, an intervention designed to improve the writing skills of high school students should include techniques from the strategy instruction, direct instruction, goal setting, reinforcement, and prewriting interventions. In order to address the needs of students at an individualized level of support, the interventions and progress monitoring should fit effortlessly into a student’s daily schedule and be quick, inexpensive, and easy to implement (Deno, 1985; L. S. Fuchs et al., 2010; Samuels, 2009; Vaughn & Fletcher, 2010; Windram et al., 2010).

Fourth, many of these studies used outcome measures that do not have research supporting their validity or reliability for progress monitoring high school students’ responses to writing interventions. These measures include TWW, %CWS, holistic ratings, errors corrected, complete and complicated sentences, elements, vocabulary, and grammar. Many of the studies in this literature review included one or more of these less valid and less reliable outcome measures (e.g., Beals, 1983; Dowell et al., 1994; Jacobson & Reid, 2010; Kiuhara et al., 2012; Viel-Ruma et al., 2010; Thanhouse, as cited in Rogers & Graham, 2008). Future studies should focus on outcome measures that are
reliable and valid for assessing progress over time for high school students. CWS and CIWS are writing CBM measures that have been shown to be the most reliable and valid outcome measures for progress monitoring for high school students. MLCWS, IWS, CPM, characters/word, and total sentences written also are promising outcome measures for high school students, but these measures need more research to support their reliability and validity.

Fifth, writing interventions should be conducted with more diverse groups of students. Of the 10 studies that reported on the race of the participants, 67% were Caucasian American, 24% were African American, 10% were Hispanic American, and 4% were of multiple races. While this is a diverse sample, there have been no studies including Asian Americans, the fourth largest ethnic group in the United States (United States Census Bureau, n.d.). All but two of the studies in this review included students with disabilities. Of those students, 81% had been identified as having a learning disability. Fewer students from other disability categories were included; 8% were identified as emotionally or behaviorally disturbed, 6% were identified as attention-deficit hyperactivity disorder (i.e., other health impaired), 3% were identified as cognitively disabled, and 1% were identified as speech-language impairment. Two studies did not include students with disabilities. Furthermore, only one study included English Language Learners. Sixty-seven percent of the study participants were male. The settings were almost evenly split between urban \((n = 7)\) and suburban \((n = 6)\), with only one study having been conducted in a rural setting (six studies did not report the
setting). The studies were spread out evenly among the four grade levels associated with high school: 9, 10, 11, and 12.

Sixth, in terms of study design, 18 of the studies utilized a multiple-baseline design, one used changing criterion, one used ABAB, and one combined multiple-baseline and multiple treatments. A greater variety of single-subject study designs would enhance the external validity of single-subject findings and the findings of experimental and group-based designs. A greater variety would demonstrate effectiveness of the intervention in different settings and with different populations (Horner et al., 2005).

Seventh, all of the studies included in this review used PND as the effect size outcome measure. Although PND is currently regarded as a useful effect size measure for single subject research by some research reviews (Scruggs & Mastropieri, 2013), it has been criticized by some researchers as an ineffective statistic for effect size analysis (Parker, & Vannest, 2009). One of the major criticisms is that confidence intervals are unable to be calculated for PND. Without a confidence interval, it is difficult to say with what degree of certainty a result is valid (Parker & Vannest, 2009). The WWC also identifies concerns about PND. Specifically, WWC is concerned that PND does not have sufficient statistical assumptions to be able to compare it to parametric tests, and the attempted use of PND to inform parametric tests, such as the phi coefficient, which creates invalid assumptions about independence of errors (IES, 2011). These critics of PND propose a new statistic that is similar to PND in terms of rationale and in the way that it is calculated but that also provides a means of calculating a confidence interval. This statistic is called the Nonoverlap of All Pairs (NAP) and is similar to Wilcoxon
Rank-Sum tests and Area Under the Curve (AUC) analyses from Receiver Operator Characteristics (ROC) tests in group-based research (Parker & Vannest, 2009; Parker, Vannest, & Davis, 2011). An AUC analysis reveals the relationship between the sensitivity and specificity of a measure (Petersen-Brown, Karich, & Symons, 2012). That is, it reveals a measure’s accuracy in identifying true positives and true negatives from false positives and false negatives. A score of 0.50 would suggest an accuracy no greater than chance. A higher AUC would indicate increased accuracy. NAP involves comparing each data point in the baseline condition to each data point in the intervention condition. This process creates pairs of data points (e.g., one baseline data point and one intervention data point is one pair, that same baseline data point and a different intervention point are another pair, etc.). One point is earned for each pair where the intervention condition exceeds the baseline condition. Half a point is earned for each “tie”—that is, where the baseline data point is the same as the intervention data point. Zero points are earned for each baseline data point that exceeds the intervention data point pair. NAP is then calculated by dividing the total number of points by the total number of pairs. Parker and Vannest (2009) explain NAP well; it is “the probability that a score drawn at random from a treatment phase will exceed (overlap) that of a score drawn at random from a baseline phase,” and is the same definition as the definition for PND (p. 359). NAP, however, compares each score in baseline to each score in intervention, while PND compares all of the intervention data to the “greatest” (i.e., the data point that is closest to the desired direction) baseline data point. In NAP, a non-overlap pair receives one point and a tie receives half a point. NAP is calculated as a
value between 0.00 and 1.00. Qualitative descriptors for NAP results are, “very [tentatively],” described as: 0.00-0.65 weak, 0.66-0.92 medium, 0.93-1.00 large (Parker & Vannest, 2009; p. 364). When using the NAP approach, an NAP score of 0.50 (or 50.0% NAP) would suggest that the result gained was due to chance. NAP represents the probability, then, that “scores from one group will exceed those of the other” (Parker & Vannest, 2009; p. 359).

Parker and Vannest (2009) list many improvements for NAP over PND. First, NAP can discriminate better among results from many studies, which would allow for greater power and validity when conducting literature reviews and meta-analyses. Second, NAP contains fewer human errors in calculation because it can be calculated using statistical programs. Third, NAP is more valid when compared to $R^2$ (i.e., a correlation measure in experimental analyses). This greater validity would allow for more comparison between studies using single-subject research methods and studies using experimental methods. Fourth, NAP is more valid when compared to visual judgments of data; that is, NAP provides results that better align with visual analyses of data (a method commonly used in single-subject research studies to analyze results). Finally, NAP contains more comparisons, leading to narrow confidence intervals (something that PND lacks), which in turn leads to greater score precision. Greater score precision means that researchers will see more valid and more reliable results.

One concern about NAP is that there is little research behind its utility (Petersen-Brown et al., 2012). PND is currently the standard for effect size analysis in single-subject research. There are only a handful of studies that call for NAP to be used (e.g.,
Parker & Vannest, 2009), some of which caution that more research needs to be conducted on NAP before it can be effectively included in research (Petersen-Brown et al., 2012). PND is the only outcome measure that was common among the studies included in this review. As such, it is recommended that future research calculate both PND and NAP in order to maintain the ability to compare outcomes to prior research and to move the field of single-subject research forward in terms of effect size calculation.

Summary

To summarize, only a handful of studies on writing interventions at the individualized level of support for high school students have been conducted since the year 2000. Future studies on writing interventions for high school students should: (a) utilize reliable and valid outcome measures for progress-monitoring with high school students, such as CWS and CIWS; (b) investigate outcome measures with promising reliability and validity, such as MLCWS, IWS, CPM, characters/word, and total sentences; (c) expand diversity in terms of ethnicity, disability status, sex, and socioeconomic status/setting; and (d) include single-subject research designs. More research is needed on all interventions described in this review in order to develop a database of evidence-based practices for high school students. Strategy instruction, direct instruction, goal setting, reinforcement, and prewriting interventions currently have the most research support, but more support is still needed for all of these interventions to meet guidelines for being considered evidence-based. Research should expand into other intervention techniques, as well, especially for popular writing interventions for which little or no research has been conducted. Targeted and individualized levels of
intervention support for high school students should effortlessly fit into a student’s schedule and be quick, inexpensive, easy to implement, and effective. Finally, while PND is currently considered the gold standard effect size for single-subject research, recent discussions in the literature point to other calculations that may be better, such as NAP. It would be useful for future single-subject studies to calculate both PND and NAP.

**Conclusion**

A review of the literature focusing on writing interventions for high school students allows one to draw some conclusions about the components of effective interventions. Based on meta-analyses conducted by a number of prominent researchers in the field of writing instruction, twelve interventions have been demonstrated to improve writing at all grade levels. These interventions include strategy instruction, summarization, peer-assisted writing, goal setting, handwriting/spelling/typing, word processing, sentence combining, inquiry, prewriting, process writing, writing-to-lean, and model study (Rogers & Graham, 2008). In order to determine effective interventions for struggling high school students, the findings of Rogers and Graham’s (2008) meta-analysis were expanded to include more recent research. Their findings also were expanded by including interventions on which they did not report due to fewer than four studies having been conducted on those interventions.

Overall, the present literature review found that effective interventions for high school students involved strategy instruction, direct instruction, goal setting, reinforcement, and prewriting. The present literature review focused on providing a
general description of all meta-analyses that have been conducted on writing interventions and a detailed description of studies that employed single-subject research methods. The review focused generally on all studies in order to provide a description of the overall literature base on writing interventions, especially since studies involving high school populations currently are underrepresented. As such, a more in-depth description of studies with similar research methods and subject pools was desired in order to be able to directly compare the results of previous studies with the results of the present study. As such, this review contained an overview of writing interventions and an in-depth discussion of 21 studies that involved single-subject research methods and high school students.

There were a number of concerns about the interventions that were reported in the context of this review. First, only one to four studies had been conducted on each specific intervention that was included in this review, which is not enough studies to meet standards to be considered evidence-based (while 10 studies have been published on SRSD in general, the number of studies behind a specific skill targeted by SRSD varied from one to four). More studies need to be conducted on all interventions that have been found to have a meaningful impact. Second, some of the interventions were poorly designed and/or were ineffective, which would subsequently negate some interventions’ chances at being considered evidence-based. Third, many of the studies used outcome measures that have been demonstrated to be less valid and less reliable for progress monitoring high school students’ responses to writing interventions (e.g., TWW, %CWS, holistic ratings, errors corrected, complete and complicated sentences, elements,
vocabulary, and grammar), which calls into question the reliability and validity of these studies’ findings. Instead of using these less appropriate measures, researchers should use CWS and CIWS in order to progress monitor with high school students. Fourth, a majority of the participants were male (67%), were Caucasian American (67%), and were receiving special education services (89%). The diversity of the student populations should be expanded to include more females, more Asian Americans (of whom none were included in these studies), and more struggling students who have not received special education services. Fifth, very few studies have been conducted outside of the urban and suburban settings. More studies should be conducted in rural settings. Sixth, only 5 of the 21 single-subject studies were published in the last 10 years. More current research is desired. Finally, all of the studies used PND as the effect size outcome measure. While PND is currently the standard for effect size in single-subject research, some authors suggest that NAP (or other potential measures) should replace PND because it has more statistical soundness (e.g., Parker & Vannest, 2009).
CHAPTER III

METHODOLOGY

Method

Setting

The study took place in an alternative urban high school in Ohio. The district had a total student population of about 22,000 and was comprised of seven traditional high schools and four nontraditional high schools (e.g., specialty Science, Technology, Engineering, and Mathematics school; early college program; evening program; alternative program). The district served students from all socioeconomic backgrounds, but the majority of students (85%) were considered economically disadvantaged (Ohio Department of Education [ODE], 2011). If one were to look at more recent data from ODE (2013), one would note that 100% of students in the district are considered economically disadvantaged, but this data is misleading. The criteria for considering a student as “economically disadvantaged” comes from how many students receive free or reduced lunch. The 2013 data from ODE reflects a change in the school district’s nutritional services program that allows for all students, regardless of economic need, to receive free lunches, and does not reflect an increase in the number of economically disadvantaged students in the district. The district student population was diverse; 46% of students were Black, 39% were White, 7% were Multi Racial, 5% were Asian or Pacific Islander, and 3% were Hispanic (ODE, 2013). Additionally, 19% of students received special education services and 6% of students were limited English proficient. In the state of Ohio, 73% of students were White, 16% were Black, 4% were Multi
Racial, 2% were Asian or Pacific Islander, and 4% were Hispanic (ODE, 2013). Students in the state of Ohio who were considered economically disadvantaged made up 48% of the Ohio school-aged population. In the state of Ohio, students with disabilities made up 14% of the population and students who were limited English proficient made up 3% of the population. For a side-by-side comparison of the demographics of students within the district in which this study took place and the state, see Table 9.

Table 9

*State and District Demographics*

<table>
<thead>
<tr>
<th>Demographic</th>
<th>State</th>
<th>District</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>73%</td>
<td>39%</td>
</tr>
<tr>
<td>Black</td>
<td>16%</td>
<td>46%</td>
</tr>
<tr>
<td>Multi Racial</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>48%</td>
<td>85%*</td>
</tr>
<tr>
<td>Students with Disabilities</td>
<td>14%</td>
<td>19%</td>
</tr>
<tr>
<td>Limited English Proficient</td>
<td>3%</td>
<td>6%</td>
</tr>
</tbody>
</table>

*Note. *All data from the Ohio Department of Education (2013), with the exception of the district’s economically disadvantaged number, which is from 2011.*

Baseline, intervention, and maintenance data were collected by the researcher in a one-on-one setting. The participant sat in a desk in a classroom and the researcher sat in a desk to the left of the participant, with the desk at approximately a 45 degree angle facing toward the participant. The classroom in which the intervention and data
collection took place changed frequently due to the availability of different classrooms on different days. The rooms and hallways generally were quiet. Even though the door was left open during the intervention, the intervention was brief enough so that students only passed by in the hallway a few times over the course of the intervention. A few intervention sessions were conducted with the participant sitting in a desk and the researcher sitting across from the participant in another desk, instead of next to the participant.

Classes were taught as direct instruction, online self-guided coursework (with teachers present to discuss questions), or a combination of the two. Students were able to take English, science, social studies, math, and some elective classes (e.g., physical education, art) at the high school. Other elective classes (e.g., music) and vocational classes were not available at the high school. Students would graduate and earn a diploma from their home high school instead of from the alternative high school. Students would apply and interview with the principal and guidance counselor in order to attend the school. Reasons for attending include recovery of failed coursework, recovery of credits in order to graduate on time, necessity to work during the day, pregnancy, or inability to perform well in a traditional high school setting. About 306 students attended the high school; 69% were Black, 22% were White, 6% were Multi-racial, and 24% had a disability (ODE, 2013). Data on other races and statuses (e.g., limited English proficient) were not available due to fewer than 10 students being reported as part of that race or status. From 2011 data, 99% of students were economically disadvantaged (ODE, 2011).
Participants

A total of six participants were included in this study. Participants attended school for a half-day session, either in the morning or in the afternoon. Three students attended in the morning and three attended in the afternoon. An overview of each participant’s demographic information is available in Table 10. Participants were given pseudonyms, which are reported in the table.

Table 10

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Grade</th>
<th>Sex</th>
<th>Race/Ethnicity</th>
<th>Disability</th>
<th>Session</th>
<th>Winter CWS</th>
<th>State Test</th>
<th>1st Semester English Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>16</td>
<td>10</td>
<td>M</td>
<td>Black</td>
<td>NA</td>
<td>Morn</td>
<td>26</td>
<td>OGT: 368</td>
<td>B</td>
</tr>
<tr>
<td>Alonzo</td>
<td>17</td>
<td>9</td>
<td>M</td>
<td>Multiple</td>
<td>NA</td>
<td>Morn</td>
<td>27</td>
<td>OAT: 393</td>
<td>NE</td>
</tr>
<tr>
<td>Omri</td>
<td>16</td>
<td>9</td>
<td>M</td>
<td>Black</td>
<td>SLD</td>
<td>Morn</td>
<td>26</td>
<td>OAT: 397</td>
<td>NE</td>
</tr>
<tr>
<td>Anderson</td>
<td>18</td>
<td>11</td>
<td>M</td>
<td>White</td>
<td>NA</td>
<td>Aft</td>
<td>29</td>
<td>OGT: 362</td>
<td>NE</td>
</tr>
<tr>
<td>Tyler</td>
<td>17</td>
<td>10</td>
<td>M</td>
<td>Black</td>
<td>NA</td>
<td>Aft</td>
<td>21</td>
<td>OGT: 390</td>
<td>NE</td>
</tr>
<tr>
<td>CJ</td>
<td>16</td>
<td>10</td>
<td>M</td>
<td>Black</td>
<td>NA</td>
<td>Aft</td>
<td>17</td>
<td>OGT: 374</td>
<td>NE</td>
</tr>
</tbody>
</table>

Note. Winter CWS = Winter benchmark correct writing sequences score; NA = not applicable; Morn = Student attended morning session; Aft = Student attended afternoon session; OGT = Ohio Graduation Test, taken starting in 10th grade; OAT = Ohio Achievement Test, taken in multiple years prior to high school, most recent score is reported; NE = student not enrolled in an English course during the first semester. Participant names are pseudonyms.

Kyrie was a 16-year-old Black male in the 10th grade. He had attended a traditional high school in the district for two years before enrolling at the alternative high school. Kyrie failed a few core academic (i.e., English, math, science, and social studies)
classes during his ninth grade year. Although he was able to advance to the 10th-grade version of courses he had passed, he did not possess enough credit hours to be considered a 10th grade student. Following his second year in ninth grade, Kyrie participated in summer school to make up one class. He then enrolled in the alternative high school and, simultaneously, participated in the evening high school programs. Throughout his high school years, Kyrie’s attendance was regular, missing only a few days of school. Kyrie had met the graduation requirements for the Ohio Graduation Test (OGT) in all areas (i.e., reading, math, writing, science, and social studies) except for the writing test.

Alonzo was a 17-year-old male in the ninth grade and was from a multiple race background. He attended at a traditional high school in ninth grade and failed all but one semester of his core academic classes. As a result, he repeated ninth grade at his traditional high school and took a half-credit English and quarter-credit physical education classes at the alternative high school that same year. The next year, he still had not earned enough credits to be considered a 10th grade student, and he enrolled at the alternative high school. Alonzo’s attendance record throughout his high school years was regular. He had not met graduation requirements for any of the OGT subject tests.

Omri was a 16-year-old Black male in the ninth grade who was identified as having a specific learning disability. He attended a traditional high school for ninth grade but only earned a half-credit total. As a result, he enrolled at the alternative high school. Omri’s attendance was irregular throughout his high school years. Specifically, he missed 42 days of school (29.5 excused) at the traditional high school and 42 days of
school (24 excused) at the alternative high school. Omri had not met graduation requirements for any of the OGT subject tests.

Anderson was an 18-year-old White male in the 11th grade who attended a traditional high school for ninth grade but passed only one core academic class. Subsequently, he repeated ninth grade and earned only 1.5 credit hours of core academic classes. He then enrolled at the alternative high school, passed all courses taken, and earned enough credits to start as a 10th grade student. The next year (i.e., the year during which the study took place), Anderson entered 11th grade at the alternative high school. In the mornings, he attended a vocational program at a traditional high school and attended the alternative high school in the afternoon. Throughout his academic years in high school, Anderson’s attendance was regular, missing only a few days of school. Anderson had met the graduation requirements for the OGT in all areas except for the writing test.

Tyler was a 17-year-old Black male in the 10th grade. He had attended a traditional high school for ninth grade and passed all of his coursework. He started his 10th grade year at the traditional high school but failed all of his first semester of coursework. As a result of his academic difficulties, Tyler enrolled in the alternative high school. While Tyler’s attendance was regular in ninth grade, he missed 42 days of school (only 2 excused) in 10th grade. Tyler met all graduation requirements for the OGT except for the math, writing, and social studies tests.

CJ was a 16-year-old Black male in the 10th grade who attended a traditional high school for ninth grade. However, CJ failed several core academic classes but had earned
enough credits to matriculate as a 10th grade student the following year. CJ transferred to a different traditional high school within the district for 10th grade where he failed the majority of his classes that year. As such, he enrolled in the alternative high school the following year. CJ’s attendance was regular throughout his high school years. He had met graduation requirements only for the reading OGT.

**Researcher and data collectors.** The researcher and interventionist in the study was in the fourth year of the Ph.D. track of the school psychology program at Kent State University. The researcher also was a school psychology intern in the school district in which the intervention took place. Data were collected or analyzed by three other school psychology interns in the school district. The intern who provided baseline probe administration also was a fourth year doctoral student in the school psychology program at Kent State University. The intern who provided interobserver reliability agreement data and intervention probe administration was in the third and final year of the Ed.S. (i.e., educational specialist) track of the school psychology program at Kent State University. The intern who provided inter-rater reliability agreement data was a Psy.S. (i.e., psychology specialist) student in the school psychology program at Cleveland State University (i.e., she was in her third and final year of study in the program).

**Materials**

The following materials were required to complete this study: lined paper, pencils, timer, desks, data collection sheets, self-graphing/goal setting chart, reinforcements chosen by participants, colored writing utensils for self-graphing and goal setting, and AIMSweb’s guide for administering writing probes and scoring written
expression (Powell-Smith & Shinn, 2004). Additional procedural details are provided below.

**Dependent Variables**

**Curriculum-based assessments.** The outcome measure for participant performance included multiple measures taken from writing samples, but correct writing sequences (CWS) drove the intervention. The writing samples were obtained using standardized AIMSweb writing fluency administration and scoring procedures (Powell-Smith & Shinn, 2004).

**Primary outcome measure and goal setting measure.** CWS was utilized as the primary outcome measure and the measure from which goals were set because it is a reliable and valid measure for progress monitoring the performance of high school students (Espin, Scierka, Skare, & Halverson, 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). CWS currently is the most valid outcome measure for high school students because it takes into account spelling, grammar, and punctuation. The AIMSweb directions for scoring CWS were followed (Powell-Smith & Shinn, 2004). To summarize, a pair of words makes one correct writing sequence if the words (a) are spelled correctly, (b) are capitalized and punctuated correctly, and (c) make sense grammatically (i.e., semantically, syntactically, etc.) within the context of the passage. An example set of sentences scored using the CWS protocol is presented in Table 1.

**Additional outcome measures.** Correct minus incorrect writing sequences (CIWS), incorrect writing sequences (IWS), and correct punctuation marks (CPM) are
promising candidates for valid and reliable measures of the performance of high school students (Diercks-Gransee, Weissenberger, Johnson, & Christensen, 2009; Espin et al., 2000; McMaster & Campbell, 2008). As such, these measures were calculated for each participant. CIWS is calculated by finding the total number of writing sequences and subtracting the number of CWS from that total. IWS is calculated by counting the total number of writing sequences that are incorrect. CPM is calculated by counting the total number of punctuation marks (e.g., periods, commas, quotation marks, dashes, exclamation points, question marks) that are located in the correct place and used correctly (Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002). Commas that are used but are not mandatory are counted as correct so long as they are used properly.

In order to further discussion of the validity of total words written (TWW; Deno, 1985; Shinn, 1989) for high school students, TWW was calculated, as well. A word is counted in TWW so long as it is a group of one or more letters that are separated from other groups of one or more letters by a space. The word does not have to be spelled correctly, make sense grammatically, or be punctuated correctly to count in the TWW score. All such groupings are counted and this is the TWW score.

Each of these additional outcome measures (e.g., CIWS, IWS, CPM, & TWW) were calculated and reported to provide information pertaining to the utility of using these measures as progress-monitoring tools. Such additional information was viewed as a way to expand the extant literature and to assist other researchers; they were not used to set goals (CWS was used for this purpose).
**Generalization measures.** The effect of the intervention in other settings also was measured. Specifically, performance in English/Language Arts class, writing samples from other courses, and cover letters for job applications were requested from every participant’s teachers and served as generalization measures. Percentage curriculum-based assessment (CBA) measures were calculated in order to investigate the intervention’s generalizability to these tasks. Percentage CBA measures are appropriate for pre/post comparisons of performance because they are reliable and valid measures of the performance of older students (Amato & Watkins, 2011; Jewell & Malecki, 2005; Tindal & Parker, 1989). Percentage measures are production-independent, meaning that the measure is not affected by how much a student is able to write within a specified time limit (Amato & Watkins, 2011; Jewell & Malecki, 2005; Tindal & Parker, 1989). As such, the pre/post writing samples offer reliable and valid data on participant performance, even though data were not taken on the time it will take the participants to write the samples. The writing samples gathered for generalization purposes were assessed using percent correct writing sequences (%CWS; Tindal & Parker, 1989). Percent CWS is calculated first by counting the number of CWS. The CWS score is divided by the total number of writing sequences and multiplied by 100.

**Classroom performance.** Participant performance on in-class writing assessments was monitored. In-class writing performance was monitored using a data sheet completed by every participant’s teacher and included information on grades for writing-related assignments (see Appendix A). These data were not used to set goals, but
were used to evaluate the effectiveness of the intervention in the general education setting.

**Participant samples.** The participants’ teachers in English/Language Arts, Science, and Social Studies were asked to provide writing samples, if available, for every participant involved in the intervention. Two samples were requested: one that was completed before the intervention and one completed after the intervention. Percent CWS was calculated on these samples. Again, these data were not used to set goals, but were used to evaluate the effectiveness of the intervention in the regular education setting.

**Cover letter sample.** As a third way to measure generalization, participants were asked to write a cover letter for an application for a job: once during the baseline phase and once more after the intervention was completed. The cover letter prompt was provided and administered by the school’s guidance counselor, using local job postings (selected at random) and the simple direction to write a cover letter for the job (see Appendix B). The cover letter sample was assessed using %CWS and compared pre/post-test for all participants. These data were not used to set goals, but were used to evaluate the effectiveness of the intervention in an applied setting.

**Independent Variable**

A goal setting intervention that included components of performance feedback, self-graphing, and reinforcement, was utilized. This intervention is described in detail in the procedure section below. The intervention took place three times per week, with the exception of days in which the school was not in session (e.g., holidays) or the researcher
was out of the district. It was anticipated that the intervention period would be a minimum of six weeks in duration. During the intervention phase, participants continued to meet with the researcher during free time, as they did during the baseline phase.

Data Collection

**Inter-rater reliability.** In order to confirm that the AIMSweb probes would be scored correctly using the scoring procedures outlined in AIMSweb (Powell-Smith & Shinn, 2004), the researcher and a school psychology intern were trained in scoring CWS by a school psychologist in the district familiar with AIMSweb prior to the intervention. Specifically, the school psychologist conducted a 45-minute session with the researcher and the school psychology intern. At the beginning of this session, the researcher, school psychology intern, and school psychologist reviewed the scoring manual together. Once a common level of understanding was verified between all three practitioners, the school psychologist provided the researcher and the school psychology intern with five sample responses to score. The researcher and intern scored their own copies of the same responses, and, later calculated the reliability of their scoring. Inter-rater reliability of at least 80% was desired, as Kazdin (2010) suggests this as the traditional level of desired agreement (though Kazdin also notes that different circumstances may warrant different levels). After scoring the five sample responses, a reliability of 91.3% was attained following the 45-minute training session.

For the study, CWS was calculated across 33.3% of the writing samples (about one sample per week) from the baseline and intervention phases. Because the researcher and the participants scored the writing samples immediately after completing them, each
writing sample was photocopied after being scored. Any markings on the writing samples were covered with white out by the researcher and photocopied again to ensure a clean copy from which the school psychology intern could score the sample. To maintain confidentiality, the participant’s name also was concealed on the copy that the school psychology intern scored. The scores from the researcher and intern were compared and a percentage generated on the degree to which the researcher and school psychology intern agreed on the CWS score for those writing samples. Agreement was calculated as the number of CWS agreements divided by the total number of agreements plus disagreements, divided by 100. Mean agreement was 93% (range, 81% to 100%).

Additional writing probe administration. A school psychology intern administered writing probes at randomly selected times during baseline (one time) and during intervention (two times) for each participant as a further check on the administration and scoring of CWS. The researcher reviewed the administration procedures with the intern. One probe was administered a couple days after the third baseline measure was administered, one probe was administered the same day as the seventh intervention session, and one probe was administered the same day as the 15th intervention session, each time using a different AIMSweb writing probe (Powell-Smith & Shinn, 2004). The probes that were administered the same day as the intervention sessions were administered immediately following the intervention sessions. Agreement between the researcher-administered probe and the intern-administered probe was calculated as the number of CWS agreements divided by the total number of agreements plus disagreements, multiplied by 100. The third baseline probe was compared to the
baseline intern-administered probes that were completed in the few days following administration of the third baseline probe. The seventh and 15th intervention probes were compared to the intern-administered probes that were completed on the same day. Mean agreement was 85% (range, 73% to 97%).

**Experimental Design**

The experimental design followed a criterion-referenced single-subject research design. Specifically, this study utilized a changing-criterion design. In changing-criterion designs, there is one baseline phase and one intervention phase for every participant (Kazdin, 2010). Changing criterion designs allow the subject to set goals (i.e., criteria) alone or in conjunction with the investigator, or for the investigator to set goals.

Initial student performance is assessed during baseline conditions. This initial performance is then followed by a series of interventions subphases. When the intervention begins, a goal is set as a level of performance that the subject must meet. When the subject’s performance meets the goal in two of three consecutive trials, the goal is changed and made more challenging (Alberto & Troutman, 2009). Thus, a subphase consists of at least two data points. This process continues throughout the subphases, during which the goal is continually changed. In some changing-criterion designs, subjects are rewarded for meeting each goal as another level of reinforcement (Kazdin, 2010). To demonstrate a change in performance, the rate of performance during baseline is compared to the overall intervention performance. A change in performance is also demonstrated from one intervention subphase to the next, when each subphase acts as a kind of baseline for the subphase following it. Essentially, each subphase is a mini-
intervention to improve on the conditions in the phase before it, which (if the intervention is successful) leads to the repeated demonstration of experimental control (Kazdin, 2010). Additionally, the mean levels of performance during each subphase are compared to demonstrate improvement from one subphase to the next. Ideally, the change in the criterion or goal between subphases is large enough to demonstrate a change in the rate of the performance.

**Procedure**

**Participant recruitment.** Participants were selected to participate based on a nomination process. Specifically, the school’s existing interdisciplinary decision-making team, called the Intervention Assistance Team (IAT), discussed students who were in need of intervention services. The existing IAT included the researcher (who was a school psychology intern in the school district), a school psychologist, regular and special education teachers, a counselor, and an administrator. The IAT examined student performance on classroom-based assessments, curriculum-based assessments, and state assessments. In order to document these scores, the team completed the Intervention Nomination Form (see Appendix C). The Intervention Nomination Form requested basic demographic information, such as age, sex, race/ethnicity, grade level, and status as a student with or without a disability. Students of all races and ethnicities were considered for participation. Students did not need to be receiving special education services in order to participate. Likewise, students with disabilities were allowed to participate. The purpose of collecting the demographic information was to report on the demographics of the participating students and in order to abide by standards for high-quality research.
Such standards require that researchers report on the demographic characteristics of study participants (Cook, Tankersley, & Landrum, 2009; Horner et al., 2005).

Inclusion criteria. The IAT nominated the six lowest-performing students based on the data provided on the Intervention Nomination Form. Specifically, data for each student was available through AIMSweb. Because the intervention focused on improving writing using CWS, performance on the AIMSweb outcome measure of CWS was the primary data source used to identify the six lowest-performing students. Classroom-based assessments and state assessments also were used to confirm poor writing performance. A student needed to have received one of the six lowest CWS scores in the school and also performed poorly on classroom-based assessments and/or state assessments. For example, if a student received the lowest AIMSweb CWS score but had performed within the average range on both classroom-based assessments and state assessments, that student was not considered to be in need of a writing intervention.

The six students who were nominated (and their parents or guardians for students under 19 years of age) were contacted about possible participation in the study. The researcher described the informed consent process through the informed consent form (see Appendix D). If interested in participating, students less than 19 years of age and their parent or legal guardian provided informed consent. If a student was 19 years of age or older, the student was able to provide his or her own informed consent alone. If a student and/or their parent or guardian did not give consent to participate, the next lowest-performing student on the AIMSweb writing fluency measure was contacted to participate. This process continued until six students and/or their parents or legal
guardians had provided informed consent to participate in the study, or until four students who had already provided consent had begun the intervention phase of the study, whichever came first. All students who participated and their parents gave consent to participate in the study.

**Exclusion criterion.** Because the high school was an alternative high school, some students in the school may have had only a few days or weeks remaining at the school before they completed their credits required for graduation. As such, the only exclusion criterion for this study was that if the student had fewer than eight weeks remaining at the school, that student was excluded from participation. This exclusion criterion was necessary in order to ensure that the student would continue to be a high school student for the entire duration of the intervention. No students who met inclusion criteria for the present study also met this exclusion criterion.

**Response rate.** Initially, six students (and their parents or guardians) who met inclusion criteria and had the lowest CWS scores were contacted to participate in the study. Two of the students and their parents/guardians provided consent to participate and four parents did not respond to requests to discuss the study and consent. Following the same nomination procedures, four more students and their parents/guardians were contacted, of which three students and their parents/guardians gave consent and one student declined to participate. Two additional students and parents/guardians were contacted. Of these, one declined to participate and one elected to participate. In total, 12 students and their parents/guardians were contacted and 6 responded with consent to participate, yielding a 50% response rate. The students who either declined to participate
or whose parents did not respond to requests to discuss the study and consent were similar in age, grade level, sex, race/ethnicity, and disability status to the participants. All students on the nomination list who were enrolled in an English course were passing their English courses, so this data point did not factor into the nomination process.

**Teacher consent.** Because teachers were asked to provide information on participating students’ grades during the study as a method of measuring generalizability of the intervention, teacher consent was sought. Consent was sought for teachers and staff on the IAT and teachers of students involved in the study. The researcher described the informed consent process through the informed consent form (Appendix D) to these teachers (including the administrator and counselor). All teachers gave consent to participate in the study.

**Institutional research approval.** Permission to conduct the research study was obtained from the school’s principal, the research board of the school district, and the Institutional Review Board of Kent State University. Participation was voluntary and students and teachers were able to terminate their participation at any time. The district’s review board required the following disclaimer to be added as a footnote to all materials: *Completion of this form is completely voluntary and the findings of this study in no way represent the philosophy and beliefs of the school district.*

**Baseline.** Before the intervention began, the researcher asked the teachers to report on what currently was being taught in the classroom in order to describe baseline academic instruction conditions. Teachers also were asked to provide the number of students in each class, how each class was instructed (e.g., direct instruction, self-directed...
computer-based modules, blended), and the number of years of teaching experienced they possessed at the beginning of the study. The teachers, who were given pseudonyms, did not provide direct data related to some of these questions, however, so a summary of what the researcher observed is provided in Table 11.

Table 11

**Teacher Demographics**

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Participant(s) in Class</th>
<th>Years Teaching</th>
<th>Sex</th>
<th>Race/Ethnicity</th>
<th>Instructor Type</th>
<th>Method of Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Heller</td>
<td>Kyrie, Alonzo</td>
<td>1</td>
<td>M</td>
<td>White</td>
<td>RegEd, Homeroom</td>
<td>Direct Instruction for English and Self-Directed Instruction</td>
</tr>
<tr>
<td>Ms. Alvarez</td>
<td>Omri</td>
<td>NR</td>
<td>F</td>
<td>White</td>
<td>IntSpec, Homeroom</td>
<td>Self-Directed Instruction</td>
</tr>
<tr>
<td>Mr. Bradbury</td>
<td>Tyler, CJ</td>
<td>NR</td>
<td>M</td>
<td>White</td>
<td>RegEd, Homeroom</td>
<td>Self-Directed Instruction</td>
</tr>
<tr>
<td>Mr. Golding</td>
<td>Anderson</td>
<td>NR</td>
<td>M</td>
<td>White</td>
<td>RegEd, Homeroom</td>
<td>Self-Directed Instruction</td>
</tr>
</tbody>
</table>

*Note. RegEd = Regular Education Teacher; IntSpec = Intervention Specialist; NR = Data not reported by teacher and researcher unable summarize on own. Teacher names are pseudonyms.*

For all instruction methods where “self-directed instruction” is listed as the only method of instruction, the students took the courses at their own pace on computers in the classroom with the teacher present to answer questions and to guide learning. Teachers are listed as homeroom teachers because they were in charge of a class of students in multiple grade levels taking a variety of courses. Students could go to another teacher for direct instruction in another subject or to another room for an elective (e.g., physical education, art), so classroom size varied throughout the morning and afternoon sessions.
All of the teachers generally had 10 students in their classroom at the largest point during a day, but could have had as few as one or two students in their classroom if most of their students also were enrolled in the same elective course. Ms. Alvarez, Mr. Bradbury, and Mr. Golding had classes set up in such a way. Mr. Heller, on the other hand, taught English to his students using direct instruction, but once English instruction was completed, the students began self-directed work on the computers, went to another direct instruction class, or went to an elective. The teachers were aware only of the general intervention procedures described in the consent form. Specific details of the writing intervention were not shared with the teachers during the study in order to avoid possible spoiling of results due to teacher familiarity with scoring procedures.

During baseline, participants met with the researcher during free time (e.g., before or after the morning or afternoon session began, during a self-directed instruction time when students could take a break, or during an elective class with permission of the teacher) and were presented with a pencil and a sheet of paper (e.g., college-ruled or wide-ruled or whatever the participant is accustomed to using in school). Next, participants were told the writing prompt following AIMSweb directions (Powell-Smith & Shinn, 2004). Specifically, participants were told to think about what they would write for one minute (participants will not be permitted to write during this minute). After 30 seconds had passed, participants were reminded of the prompt. After one minute of think time elapsed, the researcher instructed the participants to begin writing. Following AIMSweb protocol, participants were provided a full three minutes to write their response. Mid-way (one minute and thirty seconds) through the writing probe, the
researcher reminded participants of the writing prompt. After three minutes of writing, participants were told to stop writing and to put their pencils down.

The writing probes were counterbalanced across the participants in order to offset the possibility of some probes being easier than others. Probes were chosen from AIMSweb’s list of writing probes, starting with the probes for intermediate students (Powell-Smith & Shinn, 2004). The counterbalancing for the AIMSweb probes administered to each student may be viewed in Table 12. If a participant was absent on any given day, the probe that the participant should have received that day was skipped in order to maintain the counterbalancing across the participants.

Table 12

*Student Probe Sequences*

<table>
<thead>
<tr>
<th>Student</th>
<th>Probe Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omri</td>
<td>2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20</td>
</tr>
<tr>
<td>Kyrie</td>
<td>3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20</td>
</tr>
<tr>
<td>Alonzo</td>
<td>4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21</td>
</tr>
<tr>
<td>Tyler</td>
<td>5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22</td>
</tr>
<tr>
<td>CJ</td>
<td>6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23</td>
</tr>
<tr>
<td>Anderson</td>
<td>7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24</td>
</tr>
</tbody>
</table>

*Note.* Italicized numbers reflect a session that was missed due to absences or unavailability.

In order to ensure that probes administered by people other than the researcher at random times (e.g., the inter-rater reliability probes) were not repeated later by the
researcher, these probes were chosen at random from the list of probes for all participants 
(Powell-Smith & Shinn, 2004), and not from the probes in Table 12.

Participants were to remain in baseline until they achieved at least three stable 
baseline points. Stability was defined as 100% of baseline data falling within 50% of the 
mean of the baseline data (Alberto & Troutman, 2009). Alberto and Troutman describe 
this method, which was followed in the present study. First, the mean of the three 
baseline data points was calculated. The mean number minus half of the mean provided 
the lower limit within which the baseline data must fall. The mean number plus half of 
the mean provided the upper limit within which the data must fall. For example, if a 
participant’s mean baseline CWS score was 50, then all baseline data should fall between 
25 CWS and 75 CWS, in order for the baseline data to be considered stable. In this 
scenario, 50% of 50 is 25. By subtracting 25 from 50, the lower limit is 25 CWS. By 
adding 25 to 50, the upper limit is 75 CWS.

All six participants achieved stable baseline data within the first three baseline 
sessions. Although Kyrie’s first three baseline points alone were stable, these three data 
points just barely fell within 50% of the mean of baseline. His first three data points were 
29, 15, and 36 CWS, respectively (mean = 26.67, lower = 13.33, upper = 40.00). While 
these three data points fell within 50% of the mean of baseline, the lowest performance 
only missed being outside the limit by 1.67 CWS. A decision was made to gather a 
fourth baseline data point for Kyrie in order to determine if his baseline performance 
would continue (a) to be stable and (b) to increase in the steep upward trend that he 
demonstrated from the second baseline data point to the third data point. This decision
followed Kazdin’s (2010) suggestion of variability being a relative notion and of the
desire to demonstrate less variability in order to show changes in the magnitude and rate
of change from one phase to the next. A fourth baseline data point was gathered that
demonstrated a decrease in performance. While Kyrie’s baseline data still fell just within
the bounds of stability (his highest CWS was 36.00 and his upper bound was 36.75), his
increase in performance became less steep. As such, five of the six participants entered
the intervention phase following three stable baseline data points and Kyrie entered the
intervention phase after four baseline data points. All six participants began the
intervention during the same week. Table 13 demonstrates the baseline stability for all
six participants.

Table 13

<table>
<thead>
<tr>
<th>Student</th>
<th>Baseline Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>24.50</td>
<td>12.25</td>
<td>36.75</td>
<td>15</td>
<td>36</td>
<td>100</td>
</tr>
<tr>
<td>Alonzo</td>
<td>49.33</td>
<td>24.67</td>
<td>74.00</td>
<td>45</td>
<td>53</td>
<td>100</td>
</tr>
<tr>
<td>Omri</td>
<td>34.00</td>
<td>17.00</td>
<td>51.00</td>
<td>27</td>
<td>41</td>
<td>100</td>
</tr>
<tr>
<td>Anderson</td>
<td>28.67</td>
<td>14.33</td>
<td>43.00</td>
<td>27</td>
<td>30</td>
<td>100</td>
</tr>
<tr>
<td>Tyler</td>
<td>43.00</td>
<td>21.50</td>
<td>64.50</td>
<td>36</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>CJ</td>
<td>38.33</td>
<td>19.17</td>
<td>57.50</td>
<td>31</td>
<td>44</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. CWS = Correct Writing Sequences, NA = Not Applicable.

**Introductory session.** At the beginning of the first intervention session, the
researcher explained to the participant how the intervention procedure would progress.
The researcher followed the steps outlined in the intervention integrity checklist (see
Appendix E). First, the researcher explained that they would score the writing samples together. During this brief explanation, the researcher also informed the participant that these scores would be used to develop goals for the participant’s writing. Second, the researcher explained to the participant that he would be able to earn points for meeting goals over the intervention period. At this time, the researcher asked the participant to identify several reinforcers towards which he would like to earn points (additional information provided below). Third, the researcher administered an AIMSweb probe using the same procedures as during baseline.

**Performance feedback.** Immediately following the administration of the AIMSweb writing probe, the researcher worked with the participant on scoring the CWS of the probe. The researcher briefly explained the scoring procedures for CWS, saying that a correct sequence occurs when two adjacent words are both spelled correctly, capitalized correctly, and make sense in the grammatical context of the sentence (e.g., punctuation, subject-verb agreement, consistent verb tense). While scoring each probe, direct verbal feedback was provided immediately to the participant in both positive and negative forms. Both forms of feedback were provided because prior research has demonstrated that positive feedback (i.e., informing a participant what he or she did correctly) improves participant outcomes and negative feedback (i.e., informing a participant what he or she did incorrectly) can lead to increased effort (Kluger & DeNisi, 1996). In order to provide positive feedback in a reliable manner, the researcher provided verbal praise for every CWS. Negative feedback was provided by first asking the participant if he knew why every incorrect writing sequence (IWS) was incorrect. If the
participant was able to explain why the sequence was incorrect, the researcher provided verbal praise. If the participant was unable to explain why the sequence was incorrect, the researcher offered an explanation. Manual self-correction of errors was not a component of this intervention (and no students asked to correct their errors on paper), but some students sometimes verbally self-corrected their sentences while explaining why their incorrect sequences were incorrect.

The feedback provided was direct and explicit (Li, 2010). For example, if a participant wrote, “The dog than run to the tree,” in the middle of a paragraph that had been using past tense, this sentence was to be scored as: “^The^dog-than-run-to^the^tree^.” The researcher would have said something similar to, “You wrote five correct writing sequences in this sentence. Good job! You also wrote three sequences incorrectly. Do you know why they are incorrect? [Pause for answer.] No? Here, you spelled the word ‘then’ with an a instead of an e. Than is used for comparisons, like this dog is larger than that dog. Then is used for time, like first the dog saw a squirrel, then the dog started to run. And here, you used the wrong form of the verb ‘to run.’ You wrote ‘run’ instead of ‘ran.’ You had been writing in the past tense, so the correct third person past tense of ‘to run’ is ‘ran.’” See Appendix F for a guiding sample script for error correction.

**Self-graphing.** Following performance feedback, the participant recorded his performance on a goal setting chart (see Appendix G) with a dot to indicate his current CWS score. Self-graphing is a method of self-monitoring that requires a participant to
monitor and record his own behavior (Mace & Kratochwill, 1988), whereby a participant
charts performance on a graph (Figarola et al., 2008).

**Goal setting.** After self-graphing CWS, the researcher and participant worked
together to establish an appropriate goal. The researcher set both an end of intervention
goal and an initial goal, but the participant was able to modify them, if requested
(Johnson & Graham, 1990).

**End-of-intervention goal.** The end-of-intervention goal was established to
improve the participant’s baseline performance to at least the average range of
performance (e.g., Barry & Messer, 2003). Specifically, AIMSweb (2012) normative
comparison charts were used to determine what the average range of performance at the
50th percentile would be six weeks from the start of the intervention. Six weeks
following the start of the intervention was the end of the school year. The average
student in eighth grade performing at the 50th percentile at the end of the school year
(i.e., at the spring benchmark period) earns a score of 56 CWS (AIMSweb, 2012). Eighth
grade rates were selected because AIMSweb does not calculate separate norms for high
school students, and eighth grade norms are used to measure the performance of high
school students through AIMSweb measures. As such, 56 CWS was set at the end-of-
intervention goal for all participants.

**Initial goal.** The initial goal suggested by the researcher was 10% of the
difference between the end-of-intervention goal and the mean baseline performance. The
criterion of 10% was selected because the intervention was going to last six weeks. If the
first goal were set as one-sixth of the difference between the end-of-intervention goal and
the mean baseline performance, then the criterion of about 17% would have been chosen. Because research on goal setting suggests that goals should be ambitious but attainable (e.g., L. S. Fuchs, Fuchs, & Deno 1985; L. S. Fuchs, Fuchs, & Hamlett, 1989), so that students feel encouraged by meeting their goals while still being challenged at the same time, the criterion of 10% was chosen. Kazdin (2010) does not offer specific guidelines for setting the first criterion; he simply suggests that when setting the first criterion, the researcher should use baseline data to inform the first criterion and that the initial criterion “may need to be negotiated with the client” (p. 187). As such, the researcher suggested the initial goal as being 10% of the difference between the participant’s mean baseline performance and 56 CWS, but the participant was able to collaborate with the researcher on changing this criterion, if he so desired (Johnson & Graham, 1990).

**Remaining goals.** All remaining goals were set by the participant primarily, but with input from the researcher in case the participant set a goal too low or unrealistically high, following procedures in many changing-criterion studies (Kazdin, 2010). These procedures represent participative goal setting, which involves a collaborative interaction between the participant and the interventionist while setting goals (i.e., a self-selected goal with interventionist feedback), and has been demonstrated to be an effective method of goal setting (Johnson & Graham, 1990). In participative goal setting, the researcher assists the participant in determining an ambitious but achievable goal. The researcher in the present study recorded session notes on the interaction between the participant and researcher in setting the goals in order to report on how each goal was determined. These notes can be seen in Appendix H.
In order to establish ambitious but achievable goals for writing performance in the present study, an expected rate of improvement was used in order to ensure positive growth. AIMSweb describes the typical rate of improvement of CWS in eighth grade students as 0.19 CWS per week, with a standard deviation of 0.00 (AIMSweb, 2012). A rate of 0.19 CWS per week would result in an increase of only 1.14 CWS by the end of the intervention period. The researcher and participant set goals for each subphase as at least 1.00 CWS higher. If a participant was able to meet goals in two out of three sessions within a subphase, and there were three sessions per week, that participant could have improved his performance at a rate of at least 1.00 CWS per week, leading to an increase of at least 6.00 CWS by the end of the intervention period. The improvement of a rate of 1.00 CWS per week could help participants in approaching average levels of performance. As mentioned earlier, the participant’s baseline level of performance was compared to the average level of performance six weeks from the start of the intervention, using AIMSweb charts. The rate of improvement necessary to reach the average level of performance in six weeks of intervention was determined by subtracting the participant’s mean baseline level of performance from the goal average level of performance of 56 CWS six weeks after baseline, and the difference was then divided by six. This rate of improvement needed to meet the end goal was used by the researcher when collaborating on setting goals with the participant and is reported in the results chapter.

The researcher encouraged the participant to set the goal at least one CWS higher than the previous goal. Kazdin (2010) suggests setting subsequent goals at a large
enough difference from the previous goal that a steeper improvement slope may be demonstrated. By setting a goal of at least one more CWS per subphase, the participants could have improved at a faster rate than an eighth grade student scoring at the 50th percentile on the AIMSweb writing fluency measure. After the participant had achieved his goal in at least two out of three consecutive sessions, the participant was then encouraged to raise the goal (Alberto & Troutman, 2009).

In order to demonstrate greater intervention effects and a functional relationship between the intervention and the improvement in CWS, the researcher encouraged the participant to increase the goal in a varied manner (Kazdin, 2010). That is, the researcher in the present study discouraged the participant from increasing the goal by the same amount each time (e.g., always setting the goal as 3 CWS higher). Additionally, the nature of the changing criterion design allowed for variation in the length of time spent in each subphase, which Kazdin (2010) suggests as a way to better demonstrate correspondence between the goal change and the performance change. If a participant easily met or exceeded a goal within the first two sessions during that goal’s subphase, the researcher and participant set a new, higher goal in order to encourage the participant in his rapid upward performance. Kazdin suggests that as few as two sessions are appropriate for a subphase if the participant is able to meet the criterion that quickly.

In order to prevent a participant from becoming discouraged by staying in one subphase (i.e., keeping an unachievable goal) for too long, if a participant did not meet a goal three times in a row, he was encouraged to start a new subphase with a new goal at the highest goal previously attained. If a participant achieved the goal one time during a
subphase, he was encouraged to keep the same goal until he either met that goal one more
time (i.e., two times total) or failed to meet the same goal again three times in a row. If
he met the goal one more time, he was encouraged to set a higher new goal and enter a
new subphase; if he failed to meet the goal again three times in a row, he was encouraged
to set a new goal at the highest level previously attained and enter a new subphase. These
procedures follow Alberto and Troutman (2009) and Kazdin’s (2010) descriptions of
changing criterion designs and goal setting.

**Reinforcement.** Following goal setting, the researcher reminded the participant
that he could earn one point for meeting or exceeding a goal, and zero points for not
attaining a goal. The researcher explained that the participant would be able to earn a
small reinforcer for earning three points and a large reinforcer for earning six points. The
researcher and participant created a list of small reinforcers (e.g., small pieces of candy,
pencils) and large reinforcers (e.g., pizza lunch, $5.00 gift certificate) during the first
session (see Appendix I for student-identified reinforcers). Suggestions for reinforcers
were offered from Young, West, Smith, and Morgan’s (1991) book on teaching self-
management techniques to adolescents. Points earned were recorded on the goal setting
chart (see Appendix G), which was kept by the researcher and viewed by the participant
only during the intervention. Reinforcement in this intervention was contingent on the
goal being met, as reinforcement contingent upon goal achievement has been
demonstrated to be an effective way of improving performance (e.g., Locke, Shaw, Saari,
& Latham, 1981). As such, reinforcement was not given for any behavior other than the
behavior on which the reinforcement was contingent. When a student earned three or six
points toward a reinforcer, the student was able to select his small or large reinforcer immediately following self-graphing.

**Generalization prompt.** Alberto and Troutman (2009) described mediating generalization as a way to promote the generalization of intervention effects into other settings. Mediating generalization involves teaching participants to monitor and report on their own generalization of the intervention effects. As a method of mediating generalization, the researcher offered a verbal prompt to the participant as a reminder that he may use the skills he learned during the intervention in other settings. Specifically, at the end of every intervention session, the researcher said something similar to, “Remember, you can use what you have learned here in other places. Where else have you been able to use what we learn here?” This question served the purpose of asking the participant to report on his own generalization of the intervention, as a method of mediating generalization (Alberto & Troutman, 2009). If the participant could not come up with an answer, the researcher said something similar to, “I think you can tell me at least one place. Here’s a hint—we’re in this place right now,” with the answer to the hint being “school.” The researcher praised answers that involve writing and added more suggestions for the participant, such as college entrance letters and tests, state tests, job applications and cover letters, and letters or emails to people the participant would like to impress. By asking the participants about where else he had been able to use what he had learned during the intervention, the researcher intended to promote the use of self-analysis of CWS in other settings.
**Intervention.** For all sessions after the introductory session, the researcher began the session by reviewing the participant’s goal setting chart with the participant (Conte & Hintze, 2000). Together, the researcher and the participant reviewed the goal set for that session and the number of points currently earned for reinforcers. The researcher then administered the AIMSweb probe using the same procedures as described in the Baseline phase. Once the time for writing expired, both the researcher and the participant scored CWS. While scoring, the researcher provided immediate performance feedback in both positive and negative forms (see Appendix F for a sample script). After the probe had been scored and all feedback had been provided, the participant graphed his performance by plotting the current CWS on the chart. If the participant met or exceeded his goal, the researcher offered verbal praise and provided the participant one point toward earning a reinforcer. The participant and researcher then reviewed or set a goal for the next session. The same goal was reviewed if the participant had not yet achieved the goal in two of the three previous sessions. A new higher goal was set if the previous goal had been achieved in at least two of the three previous sessions. If the participant failed to meet the previous goal in three consecutive sessions, a new lower goal was set. These criteria for raising, maintaining, or lowering goals followed the method for shifting criteria in the changing criterion research design (Kazdin, 2010). If a student had met the three or six point criterion for earning a small or large reinforcement, the reinforcement was provided immediately following self-graphing. The prompt for generalization was provided at the end of each session. In order to maintain participant motivation at the end of the intervention period, when the participant may not have enough sessions left in
order to earn a reinforcer, the contingency for earning a reinforcer was changed during the final two intervention sessions. If the participant met the goal during one of the final two intervention sessions, the participant was able to earn a small reinforcer. If the participant met the goal during both of the final two intervention sessions, the participant was able to earn a large reinforcer. The participant was notified of this change at the end of the third to last intervention session.

**Maintenance and generalization.** Kazdin (2010) states that a maintenance phase should contain enough data points to evaluate trend. He does not, however, set forth specific requirements as to how long a maintenance period should last, how soon or long after an intervention has ended should a maintenance period begin, or how many data points should be included in a maintenance period. Researchers have evaluated maintenance using a variety of assessment schedules. As an example, in Codding, Feinberg, et al. (2005), some participants received three maintenance probes, some received two maintenance probes, and some only received one maintenance probe. Sometimes, the maintenance period immediately follows the intervention period (e.g., Miller, Skinner, Gibby, Galyon, & Meadows-Allen, 2011), and sometimes it is a few weeks after (e.g., Viel-Ruma, Houchins, Jolivette, Fredrick, & Gama, 2010). Maintenance periods are sometimes eliminated or cut short for one to all participants due to the end of the school year occurring simultaneously with the end of the intervention. In the present study, an attempt was made to gather three maintenance data points.

**Maintenance.** Due to the intervention starting seven weeks prior to the end of the school year, the final week of school (i.e., the week immediately following the
intervention) was used to gather maintenance data. Because the schedule on the participants’ last day of school, a Thursday, was different from the other days of the last week, the maintenance probes were administered on the Monday, Tuesday, and Wednesday of the final week. Student absences and availability due to class schedules on those days resulted in two probes being administered to Kyrie, Alonzo, and Anderson, one probe being administered to CJ and Tyler, and zero probes being administered to Omri.

**Generalization.** Generalization data were collected throughout the course of the intervention and throughout the maintenance period. English class grades, a cover letter writing sample, classroom writing samples, and state testing were used to gather generalization data. First, the teachers were asked to complete the Data Sheet for Classroom Teachers (see Appendix A) in order to record English grades. They were asked to turn in the data sheet at the end of the school year. First and second semester grades in English for the 2012 to 2013 school year were gathered, as well. Second, the school counselor administered one cover letter writing prompt to each participant during the baseline phase and one cover letter writing prompt during the maintenance phase. Third, the teachers were asked during the baseline phase to provide the researcher with one in-class writing sample. During the maintenance phase, teachers were asked again to provide another in-class writing sample from that last week of school. Finally, student performance on the writing Ohio Achievement Test (OAT; for students in ninth grade) and writing Ohio Graduation Test (OGT; for students in 10th grade or higher) was
gathered. Data from these state tests given before the intervention and after the intervention were collected.

**Treatment Integrity and Interobserver Agreement**

In order to foster internal validity and external validity, treatment integrity measures were administered (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000). Treatment integrity was assessed by a checklist of steps that had been reviewed and approved by a panel of researchers familiar with the interventions procedures and design of the study (i.e., the dissertation committee). This checklist stated whether or not the researcher completed each step of the intervention (see Appendix E). A school psychology intern observed 40% the intervention sessions and completed the same checklist (Gresham et al., 2000). The results of the researcher’s and the school psychology intern’s checklists were compared to determine inter-observer agreement about the level of treatment integrity. Before the first observation session, the researcher and school psychology intern reviewed the treatment integrity checklist together and created operational definitions of each step to ensure that they were looking for the same behaviors (Gresham et al., 2000). If inter-observer agreement dropped below 80% at any point during the study, the researcher and school psychology intern were to discuss the operational definitions and review the inconsistencies. The researcher and intern agreed on which steps were completed and which steps were not completed 100% of the time. The researcher and intern agreed that an average of 98% of steps were completed for each student across the six intervention sessions during which observations took place. Across the entire intervention, the researcher’s self-administered checklist of treatment
integrity revealed an average of 99% treatment integrity. The steps that were sometimes missed involved recording points on the chart at the end of the session, but the researcher remedied the missed steps during the next intervention session. If, for example, the researcher realized that she and the student forgot to record the student’s points on the chart, the researcher would note the missed step and ensure that the points were reviewed and recorded at the beginning of the subsequent session. Additionally, one participant and the researcher once forgot to record the goal line on the chart. The student and researcher remembered the goal the next session and plotted the goal line before beginning the intervention session. And once, on the second to last intervention session, the researcher forgot to offer a prompt for generalization for two participants. At the beginning of the final intervention session, the researcher asked the participants to complete these steps (i.e., steps 15 to 17 on the Treatment Integrity Checklist, Appendix E); these steps were completed again at the end of the intervention session, following the checklist. All other steps in the intervention were followed 100% of the time.

**Social Validity**

A social validity scale was given to the participants’ main teachers and to the participants at the conclusion of the intervention (Kazdin, 2010). The scale contained 15 items rated on a six-point Likert-type scale (with a one indicating “strongly disagree” to a six indicating “strongly agree”) that focused on the degree to which the participants enjoyed or liked the intervention and whether or not they would recommend this intervention to/for other students, among other things. It has been argued that interventions that are perceived by treatment agents as having social validity—including
high levels of treatment acceptability—are more likely to be implemented with fidelity by
teachers and to be responded to more favorably by students (Gresham et al., 2000; Kazdin, 2010). Appendix J contains the social validity scale for the teachers and Appendix K contains the social validity scale for the participants. The social validity scales are versions of the Intervention Rating Profile for Teachers (IRP-15; Martens, Witt, Elliot, & Darveaux, 1985), with language modified to reflect the present study. The IRP-15 has been demonstrated to be a reliable measure of treatment acceptability, with a Cronbach’s alpha of .98 (Martens et al., 1985). In previous research, the IRP-15 has undergone modifications by other researchers and given to teachers who implemented a behavioral intervention (e.g., Wright & McCurdy, 2012) and to teachers to rate their perceptions of a home-school intervention implemented by another clinician (e.g., Power et al., 2012). Even though the teachers in the present study did not administer the intervention, the IRP-15 was regarded as an appropriate way for them to assess the intervention’s perceived effectiveness, as the IRP-15 was developed by asking teachers to rate possible interventions for a hypothetical situation involving a behavioral concern (Martens et al., 1985). The IRP-15 has also frequently been given to parents (e.g., Kelly & McCain, 1995; Resetar, Noell, & Pellegrin, 2006; Scattone, 2008), even though it was not originally designed for parent use. As such, it is a fairly flexible measure.

Additionally, even though the IRP-15 has not previously been used to measure treatment acceptability in students before, the researcher desired to have the teachers’ and the participants’ social validity measures assess the same perceptions. A children’s version of the IRP-15 is available, called the Children’s Intervention Rating Profile
(CIRP; Witt & Elliott, 1985), but is a more simplified version of the IRP-15. Students rate their agreement with each statement based on a scale from “I agree” to “I do not agree” (e.g., Turco & Elliott, 1986). The CIRP was developed described by Witt and Elliot (1985) in a study where they asked children to rate their perceptions of possible interventions that a hypothetical student could receive for correcting a classroom behavior problem. The students completing the CIRP did not receive the intervention themselves. Like the IRP-15, however, the CIRP has been used to assess the perceptions of students who actually received the intervention (e.g., Kelly & McCain, 1995; Wood, Umbreit, Liaupsin, & Gresham, 2007). Similarly, the IRP-15 has been used to assess the treatment acceptability of participants in a study who received an intervention (e.g., Fan, Sidani, Cooper-Brathwaite, & Metcalfe, 2013). Due to the flexibility in the ability of researchers to change wording of the items on the IRP-15 and CIRP, and the flexibility of administering the IRP-15 to parents and to recipients of an intervention, the present study developed a teacher and a student version of the IRP-15 to assess the social validity of the present intervention.

Data Analysis

Research Questions. With regard to the first research question, Does a goal setting intervention that includes performance feedback, self-graphing, and reinforcement improve the writing ability of high school students?, descriptive visual analyses were conducted and effect sizes were calculated. First, the outcome measure depicted in the graphs for each participant was CWS, as this was the measure that was used to set goals. Second, four visual inspection methods were employed for analyzing
changes in the frequency of targeted reading performance for each participant: variability, magnitude of change, rate of change, and overall pattern. All visual inspection analyses considered both the changes in individual participant’s data, as well as comparing changes across each of the participants. Third, in-depth descriptive analyses provide further examination of the data behind the visual analyses. Finally, the single-subject effect sizes of PND and NAP were calculated to describe effect sizes for all participants.

First, all participant data was examined visually for variability. According to Kazdin (2010), variability is difficult to define in single-subject research; variability is a “relative notion” (p. 126). Kazdin simply suggests that “excessive” variability is not desired (p. 126). If the data within a phase is variable instead of stable, it is more difficult to determine a functional relationship between the intervention and the behavior. A researcher can determine variability by comparing baseline performance to the “magnitude of behavior change when the intervention is implemented” (Kazdin, 2010; p. 126). In this study, each participant graph was examined visually and the means of the data within the baseline, intervention, and maintenance phases were calculated using Microsoft Excel. The means were then divided in half. The resulting number was subtracted from the mean to gauge the lower limit of variability and then also added to the mean to gauge the upper limit of the variability (Alberto & Troutman, 2009). Alberto and Troutman (2009) state that all data points should be within 50% of the mean of each phase to demonstrate a lack of variability.

Second, data were visually analyzed to examine magnitude of change. Specifically, the researcher calculated the mean and median levels of performance in all
phases in order to describe the magnitude of change between phases (Kazdin, 2010). The mean level describes the average level at which the behavior is exhibited within a phase. The median level describes the probability distribution of the data; half of the data fall below and half of the data fall above that level (median levels are influenced less by extreme or outlying data than mean levels). These mean and median levels were compared between the phases and subphases to describe the degree to which the behavior changed. A leap or a jump in the data from the end of one phase to the beginning of the next phase indicates a discontinuity in performance and immediacy of change. This leap or jump lends credibility to the assumption that the change in the phase was responsible for the change in the behavior from the baseline phase to the intervention phase (Horner et al., 2005; Kazdin, 2010). In the present study, a leap in performance from baseline to intervention and a continuous increase in performance during the intervention subphases were desired.

Third, in order to examine the rate of change among phases, the researcher examined trend lines of the data within each phase (Kazdin, 2010). A change in the rate (i.e., a change in the direction or steepness of the trend line) among phases would lend support to the conclusion that there is a functional relationship between the change in the phase and the change in the rate of performance. The slopes of the data within each phase were computed using Microsoft Excel. For each phase, celeration lines were drawn. In addition to this visual analysis of data trends, rate of improvement in CWS per week was calculated by subtracting the first baseline data point from the last intervention data point and dividing this difference by seven (i.e., one week of baseline plus six weeks
of intervention). This number was compared to the average rate of improvement in CWS per week, 0.19 CWS, as determined by AIMSweb (2012). There are two key points to keep in mind when reviewing the trend lines for the present study. First, the intervention phase contained changes in criterion performance, so each participant’s performance was somewhat constrained by the criterion set for each subphase during the intervention. Therefore, the trend line during the intervention phase could possibly be positively or negatively influenced by the criteria set during each subphase during the intervention. A participant who chose and was able to meet large changes in criteria might demonstrate a steeper trend line during intervention than a participant who chose and was able to meet small changes in criteria. A participant who chose changes in criteria and was not able to meet them might demonstrate a flatter trend line than participants who were able to meet their criteria. Second, the maintenance period only contained two data points for those participants who were present for more than one maintenance session. A trend line can be drawn for two data points, but may not be as reliable as a trend line containing three or more data points (Kazdin, 2010). A third data point would serve to confirm the direction in which and steepness of the maintenance trend line. Student absences and class schedules where the students were unable to leave for the intervention some times during the last week of school prevented the participants from attending all three desired maintenance session.

Finally, the overall pattern of data was explored (Kazdin, 2010). A common approach to analyzing all of the data and determine an effect size in single-subject research is to calculate percent nonoverlapping data (PND; Scruggs et al., 1986). PND
measures the percentage of data points in one condition that do not overlap with the data points from another phase (Scruggs et al., 1986). In order to calculate PND, the researcher found the highest CWS score in the baseline phase for every participant. The researcher then calculated the percentage of data points in the intervention phase that fell above this CWS score. The maintenance CWS scores also were compared to the baseline phase. Scruggs et al. (1986) provide qualitative descriptors for the use of PND as an effect size: 50.0% or below is ineffective, 50.1% to 70.0% is small, 70.1% to 90.0% is moderate, and above 90.0% is large.

Because of recent concerns about the utility of PND as an effect size measure, nonoverlap of all pairs (NAP) was also investigated (Parker, & Vannest, 2009). NAP involves comparing each data point in the baseline condition to each data point in the intervention condition. The researcher created pairs of data points until all the data points from baseline and intervention were paired with each other (e.g., one baseline data point and one intervention data point make one pair, that same baseline data point and a different intervention point make another pair, etc.). The researcher followed the procedures described by Parker and Vannest (2009) in calculating NAP for every participant. A pair where the intervention CWS was higher than the baseline CWS received one point. A pair where the baseline CWS was higher than the intervention CWS received zero points. A pair of CWS scores from baseline and intervention that were the same score received 0.5 points. The points were totaled and divided by the total number of data pairs. NAP is calculated as a value between 0.00 and 1.00. Qualitative
descriptors for NAP results are, “very [tentatively],” described as: 0.00-0.65 weak, 0.66-0.92 medium, 0.93-1.00 large (Parker & Vannest, 2009; p. 364).

With regard to the second research question: Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by participants?, information from the student version of the IRP-15 (Appendix K) was evaluated. Specifically, judgments on the acceptability of the intervention package were evaluated using the standard scoring protocol on the IRP-15. Student ratings were measured based on the scaled score obtained. This information was examined descriptively. Similarly, data from the teacher version of the IRP-15 (Appendix J) was used to answer the third research question: Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by the participants’ teachers? Teacher ratings were measured and evaluated in the same manner as the student ratings.

For the final research question, Will intervention effects maintain following the intervention?, several types of data were analyzed descriptively. The data from the maintenance period was analyzed visually in the same way that the intervention data was analyzed, using the same visual and descriptive analyses described above. First, variability of the data was analyzed following the same methods outlined above for the intervention data. Second, the magnitude of change between the baseline and maintenance phases was analyzed by comparing the mean and median data from both phases. Third, the celeration lines and the rates of improvement from the baseline and maintenance phases were compared to analyze the rate of change. Finally, the overall
pattern of data was analyzed by calculating the PND and NAP from baseline to maintenance.

**Generalization.** Additional data was also analyzed in order to report on the generalization of skills learned in the intervention to other settings. The participants’ English class grades, abilities to write a cover letter, performances on in-class writing samples, and performances on state testing were used to evaluate the possible impact of the intervention on writing in other settings. First, classroom grades were compared pre- and post-intervention. The participants’ records were searched to obtain data from English class grades during the first and second semesters of the 2012 to 2013 school year. The intervention took place during the second half of the second semester. If the intervention had an impact on participant performance in English class, the second semester English grades would be higher than the first semester English grades. Second, the cover letters written by the participants during the baseline period and the maintenance period were evaluated pre- and post-intervention by calculating the %CWS for the cover letter writing samples. If the intervention had an impact on the participants’ abilities to write a cover letter, then the %CWS would improve from baseline to maintenance. Third, participants’ teachers provided samples from in-class writing activities across subject matters from the baseline and maintenance periods. These writing samples were also evaluated pre- and post-intervention for %CWS. If the intervention had an impact on the participants’ in-class writing abilities, the %CWS would improve from baseline to maintenance. Finally, participant records were searched to determine performance on statewide tests of academic proficiency. The only test that
was administered both before and after the intervention was the writing OGT. The writing OGT was administered to the students in the March preceding the intervention (which lasted from April to June) and the October following the intervention (which was also the following academic year). If the intervention had an impact on the participants’ OGT scores, their OGT scores would improve from baseline to maintenance.

**Other data.** Finally, for purely scholastic interest, TWW, CIWS, and CPM were calculated for the baseline, intervention, and maintenance writing samples. These data were compiled into graphs along with the CWS data. These data outcomes were calculated and graphed in order to provide the academic community with a graphical comparison of the possible utility of TWW, CIWS, and CPM as progress-monitoring tools for high school students.
CHAPTER IV

RESULTS

Participants’ Improvement in Written Expression

The first research question was, *Does a goal setting intervention that includes performance feedback, self-graphing, and reinforcement improve the writing ability of high school students?* Results were examined through visual and descriptive analysis of the writing probe data. Figures 1 through 6 provide a graphic display of the correct writing sequences (CWS) for each participant across baseline, intervention, and maintenance phases. For each of these graphs, solid vertical lines represent a change in phase; dotted vertical lines represent a change in criterion (i.e., a change to a new subphase); and dotted horizontal lines represent the criterion set for every subphase. Breaks in connections between points within the same phase or subphase indicate that session was missed due to participant absence or unavailability, with the exception of intervention session 13, which was missed due to researcher absence. Detailed summaries of the data in terms of variability, changes in magnitude (including immediacy), rate, and overall changes are reported.
Figure 1. Kyria's results. Intro = Introductory session, where student begins session by writing a sample and for the first time receives feedback on his writing. Maint. = Maintenance. A = baseline session, I = introduction session, B = intervention session, M = maintenance session.

Figure 2. Alonzo's results.
Figure 3. Omri’s results.

Figure 4. Anderson’s results.
Figure 5. Tyler's results.

Figure 6. CJ's results.
Variability

The mean data within the baseline and intervention phases were calculated using Microsoft Excel to examine variability. Variability was calculated as the percentage of data falling within 50% of the mean of each phase (Alberto & Troutman, 2009). To calculate variability, the mean of each phase was found and 50% of the mean number was subtracted from the mean to gauge the lower level of accepted variability and 50% of the mean number was added to the mean to gauge the upper level of accepted variability. In the baseline phase, 100% of data was found to be within 50% of the mean for all participants (see Table 14). In each subphase (i.e., each criterion change) within the intervention phase, 100% of data was found to be within 50% of the mean of those phases (see Table 15) for all participants.

Table 14

Variability in the Baseline Phase

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>24.50</td>
<td>12.25</td>
<td>36.75</td>
<td>15.00</td>
<td>36.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Alonzo</td>
<td>49.33</td>
<td>24.67</td>
<td>74.00</td>
<td>45.00</td>
<td>53.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>34.00</td>
<td>17.00</td>
<td>51.00</td>
<td>27.00</td>
<td>41.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Anderson</td>
<td>28.67</td>
<td>14.33</td>
<td>43.00</td>
<td>27.00</td>
<td>30.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>43.00</td>
<td>21.50</td>
<td>64.50</td>
<td>36.00</td>
<td>49.00</td>
<td>100.00</td>
</tr>
<tr>
<td>CJ</td>
<td>38.33</td>
<td>19.17</td>
<td>57.50</td>
<td>31.00</td>
<td>44.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note. CWS = Correct Writing Sequences, NA = Not Applicable.
Table 15

Variability in Intervention Criterion Subphases

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 1 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>34.00</td>
<td>17.00</td>
<td>51.00</td>
<td>32.00</td>
<td>36.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Alonzo</td>
<td>62.00</td>
<td>31.00</td>
<td>93.00</td>
<td>58.00</td>
<td>66.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>39.00</td>
<td>19.50</td>
<td>58.50</td>
<td>29.00</td>
<td>49.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Anderson</td>
<td>29.50</td>
<td>14.75</td>
<td>44.25</td>
<td>29.00</td>
<td>37.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>52.00</td>
<td>26.00</td>
<td>78.00</td>
<td>48.00</td>
<td>56.00</td>
<td>100.00</td>
</tr>
<tr>
<td>CJ</td>
<td>32.50</td>
<td>16.25</td>
<td>48.75</td>
<td>31.00</td>
<td>34.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 2 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>34.50</td>
<td>17.25</td>
<td>51.75</td>
<td>22.00</td>
<td>43.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Alonzo</td>
<td>64.50</td>
<td>32.25</td>
<td>96.75</td>
<td>56.00</td>
<td>65.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>40.00</td>
<td>20.00</td>
<td>60.00</td>
<td>33.00</td>
<td>58.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Anderson</td>
<td>34.00</td>
<td>17.00</td>
<td>51.00</td>
<td>29.00</td>
<td>48.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>54.00</td>
<td>27.00</td>
<td>81.00</td>
<td>44.00</td>
<td>54.00</td>
<td>100.00</td>
</tr>
<tr>
<td>CJ</td>
<td>38.50</td>
<td>19.25</td>
<td>57.75</td>
<td>34.00</td>
<td>44.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 3 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>31.00</td>
<td>15.50</td>
<td>46.50</td>
<td>29.00</td>
<td>33.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Alonzo</td>
<td>61.00</td>
<td>30.50</td>
<td>91.50</td>
<td>61.00</td>
<td>61.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>46.00</td>
<td>23.00</td>
<td>69.00</td>
<td>42.00</td>
<td>50.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Anderson</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tyler</td>
<td>57.00</td>
<td>28.50</td>
<td>85.50</td>
<td>54.00</td>
<td>60.00</td>
<td>100.00</td>
</tr>
<tr>
<td>CJ</td>
<td>44.00</td>
<td>22.00</td>
<td>66.00</td>
<td>42.00</td>
<td>46.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note. CWS = Correct Writing Sequences. NA = Participant’s criteria did not change this frequently.
Table 15 continued

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 4 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>39.50</td>
<td>19.75</td>
<td>59.25</td>
<td>38.00</td>
<td>41.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Alonzo</td>
<td>59.00</td>
<td>29.50</td>
<td>88.50</td>
<td>57.00</td>
<td>67.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Anderson</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tylera</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CJ</td>
<td>48.50</td>
<td>24.25</td>
<td>72.75</td>
<td>48.00</td>
<td>49.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 5 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Alonzo</td>
<td>69.50</td>
<td>34.75</td>
<td>104.25</td>
<td>67.00</td>
<td>72.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Omri</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Anderson</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tyler</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CJ</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Criterion 6 Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Alonzoa</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Omri</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Anderson</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tyler</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CJ</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note. CWS = Correct Writing Sequences. NA = Participant’s criteria did not change this frequently.

a = Participant only reached one data point during this subphase before the intervention phase ended.
Magnitude and Immediacy of Change

The analysis of the magnitude of change for every participant included a comparison of the mean and median data in both the baseline and intervention phases and the presence or absence of a leap or jump from one phase to the next (i.e., the immediacy of the change). The means and medians for the baseline and intervention phases are included in Table 16.

Table 16

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline Mean</th>
<th>Baseline Median</th>
<th>Intervention Mean</th>
<th>Intervention Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>24.50</td>
<td>23.50</td>
<td>32.25</td>
<td>32.50</td>
</tr>
<tr>
<td>Alonzo</td>
<td>49.33</td>
<td>50.00</td>
<td>63.46</td>
<td>64.00</td>
</tr>
<tr>
<td>Omri</td>
<td>34.00</td>
<td>34.00</td>
<td>44.86</td>
<td>47.00</td>
</tr>
<tr>
<td>Anderson</td>
<td>28.67</td>
<td>29.00</td>
<td>35.27</td>
<td>35.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>43.00</td>
<td>44.00</td>
<td>52.78</td>
<td>54.00</td>
</tr>
<tr>
<td>CJ</td>
<td>38.33</td>
<td>40.00</td>
<td>40.50</td>
<td>42.50</td>
</tr>
</tbody>
</table>

Descriptive analysis of mean and median data revealed increases in CWS scores from baseline to intervention for all participants. Additionally, upon implementation of the intervention, relatively immediate (i.e., within the first two intervention sessions) increases in CWS emerged for Kyrie, Alonzo, Tyler, and Omri. Kyrie demonstrated a leap of 14 CWS from the last baseline session to the first intervention session (see Figure 1), suggesting an abrupt change. Similarly, Alonzo and Tyler demonstrated a leap of 5 CWS and 4 CWS, respectively, from the last baseline to the first intervention session (see
Figures 2 and 5), suggesting a rapid behavior change. Interestingly, both Alonzo and Tyler also demonstrated continued increases in behavior following implementation of the intervention (a large leap of 13 CWS and 12 CWS from the last baseline to the second intervention session for both participants). Omri demonstrated a dip of 5 CWS from the last baseline to the first intervention session, but a leap of 5 CWS from the last baseline to the second intervention session (see Figure 3), suggesting relatively robust change within the first two intervention sessions. Despite these promising results, Anderson and CJ did not demonstrate immediacy of change. Anderson demonstrated equal performance from the last baseline session to the first intervention session and a dip of 1 CWS from the last baseline to the second intervention session (see Figure 4), suggesting no immediacy of change. CJ demonstrated a decrease of 10 to 13 CWS from the last baseline session to the first four intervention sessions and only a slight increase in performance throughout the remaining intervention sessions (see Figure 6), suggesting no immediacy of change.

**Criterion changes.** Following procedures outlined in the Methods chapter, the researcher calculated a guide from which each participant should set his initial goal. The initial goal that the researcher recommended was for the participant to make up ten percent of the difference between the end-of-intervention goal of 56 CWS and the participant’s mean baseline performance. The participants were invited to offer feedback on this initial goal and change it, if desired. Following these procedures: (a) Kyrie’s initial goal was suggested to be 28 CWS and he raised it slightly to 30 CWS; (b) Alonzo’s initial goal was recommended to be 45 CWS and he raised it to 55 CWS; (c) Omri’s initial goal was suggested to be 34 CWS and he raised it to 36 CWS; (d)
Anderson’s initial goal was suggested to be 32 CWS and he raised it slightly to 34 CWS; (e) Tyler’s initial goal was suggested to be 39 CWS, but he raised the initial goal to 60 CWS because he had been able to write 56 CWS during the introductory intervention session and he wanted to challenge himself; and (f) CJ’s initial goal was suggested to be 34 CWS and he raised it slightly to 38 CWS.

Kyrie met his first goal of 30 CWS easily and, subsequently, set his next goal to 39 CWS, which was a few CWS higher than his highest recent performance (see Figure 1). He failed to meet this goal the first time he attempted, but achieved it the second time. Following this initial success, Kyrie failed to meet the goal the next three sessions. As a result, Kyrie and the researcher decided to lower his goal to the highest one previously met, which was 30 CWS. After this decision, Kyrie was successful two out of three times, and, subsequently, he raised his goal to 32 CWS for the fourth subphase, which he was able to meet. Overall, Kyrie met or exceeded his goals 7 out of 12 times, and he typically performed within 1 to 6 CWS of his goals. Kyrie’s performance immediately increased in response to each criterion change within the first attempt at the goal, with the exception of the second criterion.

Alonzo met his first goal of 55 CWS (see Figure 2). Following this success, he set a highly ambitious goal of 70 CWS, which prompted the researcher to suggest selecting a more realistic goal. Despite this attempt, Alonzo was adamant that he could reach 70 CWS. Alonzo was unsuccessful in reaching this second goal, and his performance dropped during the third attempt. As a result, the goal was lowered, though Alonzo indicated that he did not want to lower the criterion all the way back to his
previously achieved goal since he had consistently performed above that range. Consequently, the goal was set to 60 CWS. Following the change, Alonzo was successful and decided to raise his goal again by only 1 CWS so he could earn more points. At first, Alonzo was unsuccessful, but on his second and third attempts, he achieved his goal. Subsequently, he raised his goal to 63 and was successful two times in a row. Finally, he raised his goal to 65, missing the goal by just 1 CWS for the last intervention session. Overall, Alonzo met or exceeded his goals 8 out of 13 times, and he typically performed within 1 to 6 CWS of his goal.

Omri easily met his first goal of 36 CWS (see Figure 3). Consequently, he set an ambitious goal of 45 CWS for the next session, and, at first, was unable to reach the goal. After missing a few intervention sessions between his first and second attempts at the second goal, Omri was able to meet his goal. Again, after missing a few intervention sessions, Omri easily met his goal during his third attempt and, subsequently, decided to raise his goal for the next session to 50 CWS. Omri was unsuccessful during his first attempt at this goal and was absent during the final scheduled intervention session. Since Omri had missed several intervention sessions and opportunities for earning points for meeting his goals, the researcher decided to allow Omri one more intervention session during the final intervention week. During this last session, Omri was successful in meeting his final goal. Overall, Omri met or exceeded his goals 5 out of 7 attempts. He typically performed within about 10 CWS of his goal. Omri responded to each criterion change by meeting or exceeding the goal by the second attempt.
Anderson was unsuccessful the first two times he tried to meet his first goal of 34 CWS, but was successful during his third and fourth attempts (see Figure 4). He cautiously raised his goal to 37 CWS and was unsuccessful the first two times he attempted but was successful the third time. As a result, the goal was maintained. Initially, Anderson was unsuccessful in meeting the goal two times in a row, but then performed well past the goal during his following two attempts. Following a criterion change, he again met his goal twice in a row on the last two intervention days, which would have permitted him to raise his goal had the intervention continued. Overall, Anderson met or exceeded his goals 5 out of 11 times, though he typically performed within 2 to 4 CWS of his goals. Anderson’s performance tended to respond to criterion changes by staying within a few CWS of each goal, gradually increasing during his further attempts at each goal.

Tyler was unsuccessful in reaching his first goal of 60 CWS, which he had set as four CWS higher than his highest previous performance (see Figure 5). After attempting the goal a second time, missing a few days of school, and attempting the goal a third time, the researcher and Tyler decided to lower his goal to 57 CWS, which was only 1 CWS higher than his highest previous performance. Tyler again attempted the goal three times and was unsuccessful, dropping sharply during his third attempt. Tyler and the researcher then decided to set his next goal to 54 CWS, which was a level at which he had been able to perform consistently in the past. Tyler was successful in meeting this goal two times in a row. The researcher and Tyler then decided to raise his goal by only 1 CWS in order to increase his chances at earning points. Unfortunately, Tyler was
unsuccessful in meeting this goal the one time he was able to attempt it. Overall, Tyler met or exceeded his goals 2 out of 9 times, though he typically performed within about 3 to 5 CWS of his goals. Tyler tended to respond to each criterion change by staying within a few CWS of each goal.

CJ was not successful in reaching his first goal of 38 CWS (see Figure 6). CJ and the researcher decided to lower his goal to 36 CWS. At first, CJ was unsuccessful, but then he met the goal two times in a row. Consequently, CJ and the researcher decided to raise his goal back up to 38 CWS. CJ was absent for several sessions, but was able to reach this goal two times in a row. As a result, CJ raised his goal to 40 CWS, of which he was successful in reaching. Overall, CJ met or exceeded his goal 6 out of 10 times. He typically performed within 7 to 8 CWS of his goals. CJ responded to each criterion change by improving his performance from his performance during previous goals.

**End-of-intervention goal.** For all six participants, the ambitious goal of improving their performances to the average range (50th percentile) was monitored by the end-of-intervention goal of 56 CWS. Specific details related to the percentiles of each participant’s winter benchmark CWS score (used to nominate the participants for the intervention), last intervention CWS score, and highest intervention CWS score are provided in Table 17. All participants improved their performances from below the 10th percentile to above the 12th percentile by the last intervention session. When considering the highest CWS scores obtained by each participant, their performances all improved to above the 22nd percentile.
Table 17

Comparison of CWS Percentiles from Winter Benchmark to Intervention

<table>
<thead>
<tr>
<th>Participant</th>
<th>Winter Benchmark CWS</th>
<th>Winter Benchmark Percentile</th>
<th>Last Intervention CWS</th>
<th>Last Intervention Percentile</th>
<th>Highest Intervention CWS</th>
<th>Highest Intervention Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>26</td>
<td>6th</td>
<td>38</td>
<td>14th</td>
<td>43</td>
<td>23rd</td>
</tr>
<tr>
<td>Alonzo</td>
<td>27</td>
<td>7th</td>
<td>69</td>
<td>75th</td>
<td>69</td>
<td>75th</td>
</tr>
<tr>
<td>Omri</td>
<td>26</td>
<td>6th</td>
<td>50</td>
<td>37th</td>
<td>58</td>
<td>53rd</td>
</tr>
<tr>
<td>Anderson</td>
<td>29</td>
<td>9th</td>
<td>37</td>
<td>13th</td>
<td>48</td>
<td>32nd</td>
</tr>
<tr>
<td>Tyler</td>
<td>17</td>
<td>2nd</td>
<td>50</td>
<td>37th</td>
<td>60</td>
<td>55th</td>
</tr>
<tr>
<td>CJ</td>
<td>21</td>
<td>4th</td>
<td>49</td>
<td>35th</td>
<td>49</td>
<td>35th</td>
</tr>
</tbody>
</table>

*Note.* Winter benchmark percentiles from AIMSweb (2012) and were used to nominate the participants for the intervention. Last intervention percentile and highest intervention percentile taken from the spring benchmark percentiles from AIMSweb, as the intervention occurred between the end of April through May, which is within the spring benchmark period for AIMSweb.

**Rate of Change**

The analysis of the rate of change for every participant included an examination of the trend lines in all phases and a calculation of the change in rate of performance (i.e., the rate of improvement) from one phase to the next.

**Trend lines.** Trend lines during the baseline and intervention phases were plotted using Microsoft Excel’s graphing features. Figures 7 through 12 show the trend line plots for each participant in all phases.
Kyrie’s baseline data demonstrated a downward trend and his intervention data demonstrated an upward trend (see Figure 7). This trend was influenced by his need to drop his criterion from the second intervention subphase to the third.
Alonzo’s baseline data demonstrated an upward trend (see Figure 8). His intervention data demonstrated a slightly less steep upward trend, though the average level of performance was higher than baseline. This trend was influenced by the need to drop his criterion from the second intervention subphase to the third subphase.

Figure 8. Alonzo’s trends.
Omri’s baseline data demonstrated a downward trend (see Figure 9). Omri’s intervention data demonstrated an upward trend. This data was influenced by multiple absences and tendency to dip in performance the first time a criterion was attempted.

Figure 9. Omni’s trends.
Anderson’s baseline data demonstrated an upward trend (see Figure 10). Anderson’s intervention data demonstrated a similar upward trend, almost continuing the same trend from baseline. This trend was constrained by his limited ability to meet his second criterion and the extended time he thus spent within the second subphase of the intervention.

![Figure 10. Anderson's trends.](image)
Tyler’s baseline data demonstrated a steep upward trend (see Figure 11). His intervention data demonstrated an upward trend of less magnitude, though this was influenced by the need to decrease his criterion from the first subphase to the second subphase and from the second subphase to the third subphase of intervention.
CJ’s baseline data demonstrated a steep upward trend (see Figure 12). Similarly, his intervention data demonstrated a steep upward trend, though this was less magnitude than during the baseline phase. This trend was influenced by his sharp decrease in performance from the last baseline to the first intervention session and the need to lower his criterion from the first intervention subphase to the second.

**Rate of improvement.** All participants increased their performances from the first baseline session to the last intervention session. Rate of improvement was calculated
by subtracting the first baseline session CWS from the last intervention session CWS, and then dividing that difference by the 7 weeks that took place from the first baseline to the last intervention session. The quotient of that calculation represents the rate of improvement in the number of CWS gained per week. The goal was to demonstrate a rate of improvement of at least 1.00 CWS per week, which would be above the AIMSweb average rate of 0.19 CWS per week for high school students (AIMSweb, 2012). AIMSweb lists the standard deviation for the rate of improvement as 0.00. All participants achieved this goal by improving upon their baseline performance by more than 1.00 CWS per week (see Table 18).

Table 18

Rate of Improvement of Correct Writing Sequences from Baseline to Intervention

<table>
<thead>
<tr>
<th>Participant</th>
<th>First Baseline</th>
<th>Last Intervention</th>
<th>Rate of Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>29.00</td>
<td>38.00</td>
<td>1.29</td>
</tr>
<tr>
<td>Alonzo</td>
<td>50.00</td>
<td>64.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Omri</td>
<td>41.00</td>
<td>50.00</td>
<td>1.29</td>
</tr>
<tr>
<td>Anderson</td>
<td>27.00</td>
<td>37.00</td>
<td>1.43</td>
</tr>
<tr>
<td>Tyler</td>
<td>36.00</td>
<td>50.00</td>
<td>2.00</td>
</tr>
<tr>
<td>CJ</td>
<td>31.00</td>
<td>49.00</td>
<td>2.57</td>
</tr>
</tbody>
</table>

*Note.* Rate of improvement for baseline to intervention calculated over 7 weeks.

Overall Pattern of Change

The analysis of the overall pattern of change included the calculation of effect size, including the percent nonoverlapping data (PND) and nonoverlap of all pairs (NAP) from baseline to intervention for every participant.
**PND.** PND was calculated by first finding the highest baseline data point. Second, all intervention data points that were above that data point were counted. This number was divided by the total number of intervention data points, and then multiplied by 100 to get a percentage. The final number represents the percent of intervention data that fell above the baseline data (i.e., that did not overlap with baseline data). Kyrie’s highest baseline performance was 36 CWS, and only 3 of 12 of his intervention data points fell above this level, yielding a PND of only 25.00%, or ineffective. Alonzo’s highest baseline performance was 53 CWS, and all 13 of his intervention data points fell above this level, yielding a PND of 100.00% (large effect). Omri’s highest baseline performance was 41 CWS, and 5 of his 7 intervention data points fell above this level, yielding a PND of 71.43% (moderate effect). Anderson’s highest baseline performance was 30 CWS, and 8 of his 11 intervention sessions fell above this point, yielding a PND of 72.73% (moderate effect). Tyler’s highest baseline performance was 49 CWS, and 7 of his 9 intervention data points fell above this level, yielding a PND of 77.78% (moderate effect). CJ’s highest baseline performance was 44 CWS, and 3 of his 10 intervention sessions fell above this point, yielding a PND of 30.00% (ineffective).

**NAP.** NAP was calculated following procedures outlined by Parker and Vannest (2009). For each participant, each baseline data point was paired up with each intervention data point. The total number of pairs was counted. Pairs that contained an intervention CWS that was higher than the baseline CWS were counted as a “positive” pair and were worth 1.00 points. Pairs that contained an intervention CWS and baseline CWS that were equal were considered a “tied” pair and were worth 0.50 points. Pairs
that contained an intervention CWS that was lower than the baseline CWS were considered a “negative” pair and were worth 0.00 points. Parker et al.’s (2011) equation for calculation NAP was used to calculate NAP for each participant: (number of positive pairs + 0.50 x number of ties) / total number of pairs. Microsoft Excel was used to count the numbers of each type of pair and to calculate NAP. Kyrie’s NAP score for the comparison of baseline to intervention was 0.75 (medium effect). Alonzo’s NAP score for the comparison of baseline to intervention was 1.00 (large effect). Omri’s NAP score for the comparison of baseline to intervention was 0.86 (medium effect). Anderson’s NAP score for the comparison of baseline to intervention was 0.91 (medium effect). Tyler’s NAP score for the comparison of baseline to intervention was 0.91 (medium effect). CJ’s NAP score for the comparison of baseline to intervention was 0.67 (medium effect).

### Participant Acceptability of Intervention: Social Validity

The second research question was, *Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by participants?* A descriptive analysis of the data from the modified versions of the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985) is provided. Total scores above 52.50 using the IRP-15 indicate that the intervention was acceptable (Von Brock & Elliott, 1987).
Participant Ratings

Five of the six participants anonymously completed the social validity scale (see Appendix K). The sixth participant, Omri, was absent when the scale was distributed and did not return to school before the end of the school year. Overall, participants found the intervention to be socially valid as rated on the modified IRP-15 (Martens et al., 1985; see Appendix K). The five participants who completed the scale each rated the intervention as acceptable, with total ratings of 84, 83, 73, 64, and 58, respectively. Of note is that all of the participants either agreed (N = 4) or strongly agreed (N = 1) that this would be an acceptable intervention for students who have writing problems. Moreover, three of five participants strongly agreed that the intervention was effective at improving their writing. Participant ratings for each item on the IRP-15 are provided in Table 19.

Teacher Acceptability of Intervention: Social Validity

The third research question was, *Is a goal setting intervention that includes performance feedback, self-graphing, and reinforcement valued by the participants’ teachers?* It was anticipated that teachers would complete the teacher version of the IRP-15 (see Appendix J). Unfortunately, none of the teachers returned the scale. As a result, no data on teacher acceptability of the intervention was available for analysis.
Table 19

*Ratings for Each Item on the Social Validity Scale for Students*

<table>
<thead>
<tr>
<th>Item</th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. This would be an acceptable intervention for students who have writing problems.</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5.20</td>
</tr>
<tr>
<td>2. Most students would think this intervention is appropriate for students who have writing problems.</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4.80</td>
</tr>
<tr>
<td>3. This intervention was effective in improving my writing.</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>4.80</td>
</tr>
<tr>
<td>4. I would tell other students to use this intervention.</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5.00</td>
</tr>
<tr>
<td>5. My writing problems were bad enough to necessitate the use of this intervention.</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>4.00</td>
</tr>
<tr>
<td>6. Most students would think this intervention is suitable for students who have writing problems.</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>4.60</td>
</tr>
<tr>
<td>7. I would be willing to use this intervention again.</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>5.00</td>
</tr>
<tr>
<td>8. This intervention did <em>not</em> result in negative side effects for my writing.</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5.40</td>
</tr>
<tr>
<td>9. This intervention would be appropriate for a variety of students.</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5.60</td>
</tr>
<tr>
<td>10. This intervention is consistent with what I am learning in class.</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>4.20</td>
</tr>
<tr>
<td>11. The intervention was a fair way to handle my writing problems.</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4.60</td>
</tr>
<tr>
<td>12. This intervention is reasonable for improving students' writing problems.</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4.60</td>
</tr>
<tr>
<td>13. I liked the procedures used in this intervention.</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5.00</td>
</tr>
<tr>
<td>14. This intervention was a good way to improve my writing.</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>5.00</td>
</tr>
<tr>
<td>15. Overall, this intervention was beneficial for me.</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>4.60</td>
</tr>
</tbody>
</table>

*Note.* The following scale was used on the rating scale: 1 = *Strongly Disagree*, 2 = *Disagree*, 3 = *Slightly Disagree*, 4 = *Slightly Agree*, 5 = *Agree*, 6 = *Strongly Agree.*
Maintenance

The final research question was, *Will intervention effects maintain following the intervention?* Results were examined through visual and descriptive analysis of the writing probe data. Figures 1 through 6 provide a graphic display of the CWS for each participant across baseline, intervention, and maintenance phases. Detailed summaries of the data in terms of variability, changes in magnitude (including immediacy), rate, and overall changes are reported.

Variability

Mean within the maintenance phase was calculated using Microsoft Excel to examine variability. In the maintenance phase, 100% of data was found to be within 50% of the mean of those phases for all participants (see Table 20). Results indicated that all participants met the standard for demonstrating stability set by Alberto and Troutman (2009).

Table 20

*Variability in the Maintenance Phase*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Maintenance Mean CWS</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Lowest CWS</th>
<th>Highest CWS</th>
<th>% within Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>34.50</td>
<td>17.25</td>
<td>51.75</td>
<td>31</td>
<td>38</td>
<td>100</td>
</tr>
<tr>
<td>Alonzo</td>
<td>67.00</td>
<td>33.50</td>
<td>100.50</td>
<td>61</td>
<td>73</td>
<td>100</td>
</tr>
<tr>
<td>Omri(^a)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Anderson</td>
<td>39.00</td>
<td>19.50</td>
<td>58.50</td>
<td>37</td>
<td>41</td>
<td>100</td>
</tr>
<tr>
<td>Tyler(^b)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CJ(^b)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Note.* CWS = Correct Writing Sequences, NA = Not Applicable.

\(^a\) = Participant absent both days of maintenance data collection.

\(^b\) = Participant absent one day of maintenance data collection, so variability and means unable to be calculated.
Magnitude and Immediacy of Change

The analysis of the magnitude of change for every participant included a comparison of the mean and median in the baseline and maintenance phases. The means and medians for the baseline and intervention phases are provided in Table 21.

Table 21

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline Mean</th>
<th>Baseline Median</th>
<th>Maintenance Mean</th>
<th>Maintenance Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>24.50</td>
<td>23.50</td>
<td>34.50</td>
<td>34.50</td>
</tr>
<tr>
<td>Alonzo</td>
<td>49.33</td>
<td>50.00</td>
<td>67.00</td>
<td>67.00</td>
</tr>
<tr>
<td>Omri</td>
<td>34.00</td>
<td>34.00</td>
<td>NA(^a)</td>
<td>NA(^a)</td>
</tr>
<tr>
<td>Anderson</td>
<td>28.67</td>
<td>29.00</td>
<td>39.00</td>
<td>39.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>43.00</td>
<td>44.00</td>
<td>NA(^b)</td>
<td>NA(^b)</td>
</tr>
<tr>
<td>CJ</td>
<td>38.33</td>
<td>40.00</td>
<td>NA(^b)</td>
<td>NA(^b)</td>
</tr>
</tbody>
</table>

*Note.  \(^a\) Participant was not present for any maintenance data collection session.  
\(^b\) Participant only present for one maintenance data collection session.*

Descriptive analysis of mean and median data revealed increases in CWS scores from baseline to maintenance for three participants. These three participants (Kyrie, Alonzo, and Anderson) who completed two maintenance sessions increased their overall mean and median CWS scores from intervention to maintenance. For Kyrie and Anderson, data demonstrated adequate maintenance following the implementation of the intervention. Data for Alonzo demonstrated significant increases in the overall mean and median CWS scores, indicating that skills maintained following implementation of the intervention. Because Tyler and CJ were only present for one maintenance session, and
Omri was absent for both maintenance sessions, no comparison of means and medians can be made about their maintenance sessions.

**Rate of Change**

The analysis of the rate of change for every participant included an examination of the trend lines in all phases and a calculation of the change in rate of performance (i.e., the rate of improvement) from one phase to the next.

**Trend lines.** Trend lines during the baseline and maintenance phases were plotted using Microsoft Excel’s graphing features. Figures 7 through 12 show the trend line plots for each participant in all phases. Data for Kyrie and Anderson demonstrated accelerating trends during the brief maintenance period (see Figures 7 and 10). For both participants, only two opportunities for data collection existed. As such, it is difficult to ascertain the continued direction of trends. Alonzo’s maintenance data demonstrated a rapidly decelerating trend (see Figure 8), making it difficult to ascertain his continued performance (a third data point may have confirmed or denied such a steep downward trend). Data trends for Omri, Tyler, and CJ could not be evaluated due to limited data points.

**Rate of improvement.** Rate of improvement was calculated by subtracting the first baseline session CWS from the last maintenance session CWS, and then dividing that difference by the 8 weeks that took place from the first baseline to the last maintenance session. The quotient of that calculation thus represents the rate of improvement in the number of CWS gained per week. Again, the goal was to demonstrate a rate of improvement of at least 1.00 CWS per week, which would be above
the AIMSweb average rate of 0.19 CWS per week for high school students (AIMSweb, 2012). For those participants who completed at least one maintenance session, all achieved this goal by improving upon their baseline performance by more than 1.00 CWS per week (see Table 22).

Table 22

<table>
<thead>
<tr>
<th>Participant</th>
<th>First Baseline</th>
<th>Last Maintenance</th>
<th>Rate of Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kyrie</td>
<td>29.00</td>
<td>38.00</td>
<td>1.29</td>
</tr>
<tr>
<td>Alonzo</td>
<td>50.00</td>
<td>61.00</td>
<td>1.57</td>
</tr>
<tr>
<td>Omri</td>
<td>41.00</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Anderson</td>
<td>27.00</td>
<td>41.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Tyler</td>
<td>36.00</td>
<td>47.00</td>
<td>1.57</td>
</tr>
<tr>
<td>CJ</td>
<td>31.00</td>
<td>40.00</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Note. Rate of improvement from baseline to maintenance calculated of 8 weeks. NA = Not Applicable because participant was not present for any maintenance sessions.

Overall Pattern of Change and Effect Sizes

The analysis of the overall pattern of change included the calculation of effect size, including the PND and NAP from baseline to maintenance for every participant. PND and NAP were calculated following the same methods described above for the baseline to intervention PND and NAP calculations. Kyrie’s highest baseline performance was 36 CWS, and only one of his two maintenance data points fell above 36 CWS, yielding a PND of 50.00% (ineffective). Alonzo’s highest baseline performance was 53 CWS, and both of his maintenance data points fell above 53 CWS, yielding a maintenance PND of 100.00% (large effect). Omri was not present for either
maintenance session, and, subsequently, PND was not calculated. Anderson’s highest baseline performance was 30 CWS, and both of his maintenance data points fell above 30 CWS, yielding a maintenance PND of 100.00% (large effect). Tyler’s highest baseline performance was 49 CWS, but his one maintenance session was only 47 CWS, yielding a PND of 0.00% (ineffective). CJ’s highest baseline performance was 44 CWS, and his one maintenance session was only 40 CWS, yielding a PND of 0.00% (ineffective).

NAP was calculated following procedures outlined in the presentation of the baseline to intervention NAP data. Kyrie’s NAP score for the comparison of baseline to maintenance was 0.88 (medium effect). Alonzo’s NAP score for the comparison of baseline to maintenance was also 1.00 (large effect). Again, Omri was not present for any maintenance session. Anderson’s NAP score for the comparison of baseline to maintenance was 1.00 (large effect). Tyler’s NAP score for the comparison of baseline to maintenance was 0.67 (medium effect). CJ’s NAP score for the comparison of baseline to maintenance was 0.50 (weak effect).

Generalization

The participants’ classroom grades, ability to write a cover letter, performances on in-class writing samples, and performances on state testing also was evaluated.

English Class Grades

Three of the four teachers did not return the Data Sheet for Classroom Teacher (Appendix A) that they had been given to monitor grades. Mrs. Alvarez returned the data sheet for Omri, but it contained only one grade: a C- on a short answer worksheet for a World History assignment. As a result of not obtaining specific grades from teachers, the
researcher obtained semester grade reports. Examination of these reports indicated that Kyrie earned a B in English at the end of the first semester (around January) and raised this to an A in English at the end of the year (June). Although Alonzo was not enrolled in an English course at the beginning of the year, his end of the year English grade was an A-. Similarly, Omri was not enrolled in an English course at the beginning of the year and earned a final, year-end grade of C in English. Neither Anderson nor Tyler were enrolled in an English course at the beginning or end of the year. As such, no grades were recorded. CJ was not enrolled in the alternative high school at the beginning of the year and his record does not indicate that he was enrolled in an English class at the high school he had been attending. CJ’s end of the year English grade at the alternative high school, however, was a C.

**Cover Letter Writing**

All participants were administered a prompt by the school counselor to write a cover letter for a job application before the intervention was implemented. Following the intervention, participants were to complete a cover letter for which percent correct writing sequences (%CWS) scores would be calculated as supplemental data. Unfortunately, the post-intervention cover letters were able to be completed by only two participants (Anderson and Tyler). The counselor responsible for implementing the cover letter writing prompts was unable to give Kyrie and Alonzo the post-intervention probe due to scheduling demands; and, Omri and CJ were absent when the counselor went to give them the post-intervention cover letter probe. For both Anderson and Tyler, %CWS was calculated for their pre and post cover letters by counting the number of CWS and
dividing that number by the total number of writing sequences. Anderson was able to improve his pre-intervention %CWS of 85.71 to a post-intervention %CWS of 90.48. Tyler dipped slightly from his pre-intervention %CWS of 93.88 to a post-intervention %CWS of 91.30.

**Classroom Writing**

Kyrie, Alonzo, Omri, and Anderson’s teachers did not provide pre- or post-intervention writing samples by the conclusion of the study. Omri’s teacher was not able to provide samples for him because she reported that he “skipped all essays in English units recently.” Even though she gave him multiple opportunities to complete the units, he did not, and no writing data for him was available. CJ’s teacher provided only pre-intervention writing samples because the participant changed to a different English class at the very end of the school year. One sample was an essay pertaining to an in-class assignment related to slavery. The other essay was a five-paragraph essay written for the prompt of, “The funniest thing I ever saw.” CJ’s slavery essay yielded a %CWS of 83.10 and his “funniest thing” essay yielded a %CWS of 87.68, resulting in an average %CWS for both essays of 85.39. Tyler’s teacher provided both a pre-intervention sample and a post-intervention sample. Specifically, the pre-intervention sample included a worksheet with three short answer (i.e., 1 to 3 sentences) questions and one essay question (the teacher indicated on the sample that the essay was too short to be considered an essay). Analysis of the %CWS across the short answers and the essay response was 92.62. The post-intervention sample consisted of a worksheet of 24 short answer questions about a play students were reading in class and yielded a %CWS of 95.86. Taken together,
results of this simple analysis indicated that Tyler demonstrated slight improvement pre
to post.

**State Testing**

Alonzo and Omri were in ninth grade, so the most recent state test that they took was the Ohio Achievement Test (OAT) in eighth grade. Kyrie, Anderson, Tyler, and CJ were in 10th grade or beyond so they each had taken the Ohio Graduation Test (OGT). For both the OAT and OGT, a score of 400 is considered a proficient grade. Proficient grades on the OGT are needed for graduation. Before the intervention, Alonzo and Omri received OAT scores of 393 (seventh grade) and 397 (fourth grade), respectively, which are considered to be in the “basic” range. Before the intervention, CJ scored in the basic range on the 10th grade OGT, with a score of 390. Before the intervention, Kyrie, Anderson, and Tyler each received OGT scores of 367 (10th grade), 362 (11th grade), and 374 (10th grade), respectively, all of which are in the “limited” range of proficiency.

In the fall following the intervention, Anderson, Kyrie, and CJ each took the OGT again. During this assessment, Anderson and Kyrie both improved their performances and scored within the “proficient” range, a score of 422 and 402, respectively. CJ’s score on the OGT was a 399, missing the proficient range by just one point. Despite CJ not meeting the proficiency expectation, he demonstrated improved performance. Tyler enrolled in another school district in the fall, and, as a result, no data about subsequent OGT performance was available. Because the OGTs are not administered until the spring of 10th grade, no post-intervention data was available for Omri or Alonzo.
Other Data

Graphs depicting the baseline, intervention, and maintenance data in correct minus incorrect writing sequences (CIWS), correct punctuation marks (CPM), and total words written (TWW) format can be seen in Figures 13 through 18. These data were not used to set goals or monitor the progress of the participants. Rather, they are reported in order to provide a discussion point on the possible utility of these measures for high school students. For comparison purposes, CWS data are included in the graphs.

*Figure 13. Kyrie’s other data. CWS = Correct writing sequences, CIWS = Correct minus incorrect writing sequences, CPM = Correct punctuation marks. TWW = Total words written.*
Figure 14. Alonzo's other data.

Figure 15. Omri's other data.
Figure 16. Anderson’s other data.

Figure 17. Tyler’s other data.
Summary

Overall Relationship among All Participants

Table 23 summarizes the findings for the variability, magnitude of improvement, immediacy of change, criterion change, trend line, rate of improvement, PND, and NAP data for every participant. Taken together, the data mainly supports positive effects on CWS for Kyrie, Alonzo, Omri, Anderson, and Tyler, while a mix of positive and negative data was observed for CJ. The average PND across all six participants from baseline to intervention revealed a small effect on CWS performance (62.85%), while the median PND revealed a moderate effect (71.93%). The mean and median NAP data across all six participants from baseline to intervention revealed a medium effect, 0.84 and .91, respectively. With regard to maintenance, the average and median PNDs revealed an
ineffective result (50.00%), whereas the average NAP (0.68) and median NAP (0.88) revealed medium effects.

While each of the participants viewed the intervention positively, teacher acceptability of the intervention was not reported. Moreover, generalization data were limited due to participant absences and missing data from several of the participant’s teachers. Overall, some positive effects on generalization were noted. Specifically, Kyrie was able to improve his English course grade and Anderson was able to improve his cover letter writing abilities. Although Tyler’s post-intervention writing sample earned a slightly lower score than the pre-intervention sample, he was able to improve his in-class writing ability. Finally, Kyrie, Anderson, and CJ demonstrated improved performances on state tests.
Table 23

Data Summary

<table>
<thead>
<tr>
<th>Data</th>
<th>Kyrie</th>
<th>Alonzo</th>
<th>Omri</th>
<th>Anderson</th>
<th>Tyler</th>
<th>CJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of Variability</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Magnitude (A to B)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Magnitude (A to M)</td>
<td>Yes</td>
<td>Yes</td>
<td>NA\textsuperscript{a}</td>
<td>Yes</td>
<td>NA\textsuperscript{b}</td>
<td>NA\textsuperscript{b}</td>
</tr>
<tr>
<td>Immediacy of Change</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Criterion Change</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Trend Lines (A to B)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ROI (A to B)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ROI (A to M)</td>
<td>Yes</td>
<td>Yes</td>
<td>NA\textsuperscript{a}</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>PND (A to B)</td>
<td>No Effect</td>
<td>Yes, Large</td>
<td>Yes, Moderate</td>
<td>Yes, Moderate</td>
<td>Yes, Moderate</td>
<td>No Effect</td>
</tr>
<tr>
<td>PND (A to M)</td>
<td>No Effect</td>
<td>Yes, Large</td>
<td>NA\textsuperscript{a}</td>
<td>Yes, Large</td>
<td>No Effect</td>
<td>No Effect</td>
</tr>
<tr>
<td>NAP (A to B)</td>
<td>Yes, Medium</td>
<td>Yes, Large</td>
<td>Yes, Medium</td>
<td>Yes, Medium</td>
<td>Yes, Medium</td>
<td>Yes, Medium</td>
</tr>
<tr>
<td>NAP (A to M)</td>
<td>Yes, Medium</td>
<td>Yes, Large</td>
<td>NA\textsuperscript{a}</td>
<td>Yes, Large</td>
<td>Yes, Medium</td>
<td>No, Weak</td>
</tr>
</tbody>
</table>

Note. Lack of variability = Was there no variability in any phase? Magnitude = Did the mean and median correct writing sequence scores demonstrate improvement from one phase to the next? A to B = comparison from baseline to intervention. A to M = comparison from baseline to maintenance. NA\textsuperscript{a} = Not Applicable because participant was not present for any maintenance sessions. NA\textsuperscript{b} = Not Applicable because participant was present for only one maintenance session. Immediacy of change = Did the participant demonstrate an immediacy of change from baseline to intervention within the first or second data point? Criterion change = Did the participant’s data demonstrate a response to criterion change by changing with each new criterion? Trend Lines = Did the participant’s trend line during intervention demonstrate an improvement over the baseline trend line? ROI = Rate of improvement; was the participant’s rate of improvement above the study’s goal of 1.00 correct writing sequences improvement per week? PND = Percent nonoverlapping data; was the intervention effective and by how much? NAP = Nonoverlap of all pairs; was the intervention effective and by how much?
CHAPTER V
DISCUSSION

The purpose of the present study was to evaluate the effectiveness of a goal setting intervention that incorporated performance feedback, self-graphing, and reinforcement on the writing skill development for high school students. Participants included six male students who attended an alternative high school in a large urban Ohio school district; and, who had performed below the average/proficient range on curriculum-based and statewide tests of written expression. Following baseline data collection and a brief introductory teaching session, participants individually met with the researcher two to three times per week for six weeks of intervention. During the intervention period, the participant and the researcher: (a) reviewed previous correct writing sequence (CWS) scores and set goals; (b) administered and scored AIMSweb writing fluency probes (Powell-Smith & Shinn, 2004); (c) exchanged feedback pertaining to each writing sequence; (d) charted participant progress on a graph; (e) selected reinforcers as appropriate; and (f) determined subsequent goals. A brief maintenance period occurred during the week following the intervention to ascertain if the participants’ skills continued post intervention. Moreover, generalization data was gathered pre/post intervention and throughout the intervention to examine transfer of training effects.
Effectiveness of the Intervention

Participant Improvement in Written Expression

Overall, a goal setting intervention that incorporated performance feedback, self-graphing, and reinforcement was effective at improving the writing performance for five of the six participants. The intervention was most effective at improving Kyrie and Alonzo’s writing, as indicated by their improvements in CWS throughout the intervention and the immediacy by which their scores improved at the start of the intervention. Their performances changed with each criteria change (i.e., goal), and they exceeded the goal rate of improvement from the baseline phase to the intervention phase. While both participants’ data trended in a positive direction from baseline to intervention, an interesting pattern emerged. Specifically, Kyrie and Alonzo both experienced difficulty with meeting the second (and ambitious) goals that they set. During the second subphase, Alonzo’s first two attempts were near the goal and only Kyrie’s second attempt was near the goal (actually exceeding the goal), but the third attempt for both participants fell about 10 points from their second attempts. Such sharp declines in performance during the third attempt at ambitious second goals may have indicated that Kyrie and Alonzo lost motivation due to setting too ambitious of goals. Following their limited second subphase performances, Kyrie and Alonzo lowered their goals to more approachable, but still challenging, levels and began increasing subsequent goals at smaller intervals. While Kyrie and Alonzo were successful at reaching remaining goals, their drops in performances from the second subphase to the third subphase may have diminished the rate of improvement anticipated during the intervention. When examining the magnitude
of the intervention over time, both percent nonoverlapping data (PND) and nonoverlap of all pairs (NAP) effect sizes demonstrated a large effect for Alonzo. Data for Kyrie was less consistent, with PND analysis indicating no effect and NAP indicating a medium effect of the intervention. Taken together, the data supported a functional relationship between the intervention and improvements in CWS for Kyrie and Alonzo.

The intervention also was effective at improving Omri’s writing, though he possessed the least amount of data due to excessive absences. Despite his absences, Omri was able to demonstrate large and stable changes throughout the intervention. Specifically, Omri’s performance increased immediately in response to the implementation of the intervention and tended to improve with each increase in criterion. He also exceeded the goal rate of improvement from baseline to intervention. PND and NAP effect sizes demonstrated moderate and medium effects on writing performance, respectively. Overall, the data supported a functional relationship between the intervention and Omri’s improvements in CWS.

Anderson and Tyler demonstrated large and relatively stable improvements in CWS during the intervention. While Anderson’s intervention data demonstrated some variability across criterion changes, his overall performance trended in a positive direction. Interestingly, the large dip in Anderson’s performance from the fourth to fifth attempt likely was due to his struggle with the prompt; he mentioned that he did not know what to write for that prompt. Conversely, for the sixth attempt, Anderson mentioned that his performance improved so much because the prompt was related to a topic of great interest to him. When reviewing data for Tyler, an immediate improvement in his
performance from the last baseline session to the first two intervention sessions was observed. While Tyler did not meet the first two goals he set, when he lowered his goal from the first to second subphase, his performance also lowered, suggesting that his performance fluctuated with the changes in criteria. It is interesting to note that, throughout the intervention, Tyler did not appear to show much interest. He was quiet and reserved, and he responded with brief answers to conversational questions from the examiner. His disinterest could be due in part to his average level of performance from the outset of the intervention. Because his performance during the introductory session was at the end-of-intervention goal level, perhaps his motivation to improve was lessened (it may have been more difficult for Tyler, who already performed within the average range, to improve his performance even further than it would be for someone below the average range). The intervention data did indicate that Tyler’s performance improved, even though he began the intervention around the level of the end-of-goal intervention goal. Aside from their individual differences in performance across the intervention, PND and NAP analyses indicated moderate and medium effects for Anderson and Tyler, respectively. Taken together, the data supported a functional relationship between the intervention and improvements in CWS for Anderson and Tyler.

The data was mixed regarding the relationship between CJ’s performance and the intervention. On the positive side, CJ was able to demonstrate large and stable improvements in CWS during the intervention, and his performance fluctuated in response to each criterion change (with the exception of the first criterion). Moreover, CJ exceeded the goal rate of improvement from baseline to intervention and NAP analysis
suggested that the intervention had an overall medium effect. On the negative side, CJ’s performance did not immediately improve from baseline to intervention. In fact, his performance decreased to a level consistent with his performance during the baseline and introductory sessions; and, his data demonstrated a somewhat flat trajectory over the intervention period. As a result, PND analysis revealed no effect of the intervention on CJ’s writing performance. Overall, then, the data provided mixed results as related to whether or not a functional relationship existed between the implementation of the intervention and CJ’s improvements in CWS. Visual analyses and comparisons of averages in baseline and intervention and his rate of improvement were more positive, while the effect size data was mostly negative.

**Maintenance**

Overall, outcomes associated with maintenance data were mixed. Three of the six participants (Kyrie, Alonzo, & Anderson) maintained the positive effect of the writing intervention following the intervention period. For Kyrie, Alonzo, and Anderson, mean maintenance CWS scores were at least 10 points higher than their mean baseline scores. Moreover, calculation of PND and NAP indicated generally promising effects. Specifically, PND calculations revealed large effects during the maintenance period for Alonzo and Anderson but no effect for Kyrie; whereas NAP calculations revealed large effects for Alonzo and Anderson and a medium effect for Kyrie. For Tyler, CJ, and Omri, limited data precluded the ability to suggest the intervention maintained long term. Only one data point was able to be collected for Tyler and CJ. No maintenance data was collected for Omri. Because of the limited data for Tyler, CJ, and Omri, no definitive
conclusions can be made about whether or not they were able to maintain the positive effects of the intervention.

Overall, the number of opportunities for collecting maintenance data in this study was limited (e.g., two data points for three participants, one data point for two participants, and zero data points for one participant). While this was an unfortunate circumstance, this is not an unusual pattern within the extant literature. Several previous studies have possessed only one to two maintenance data points for their participants (e.g., Jacobson & Reid, 2010) or have run into difficulty with obtaining all intended maintenance data points due to the end of the school year (e.g., Viel-Ruma, Houchins, Jolivette, Fedrick, & Gama, 2010). Although the quantity of maintenance data collected within the current study was similar to that reported in prior investigations, it is a possible limitation. Limited maintenance data does not allow for a trend to be established that confirms that the skills were maintained following the intervention. In the current study, only partial trends were established for Kyrie, Alonzo, and Anderson, and no trend was established for CJ and Tyler (and no data was available for Omri). While this is a consequence experienced by some researchers, future researchers should take care to plan for a more extensive maintenance period.

Social Validity

From a participant perspective, the intervention was rated as acceptable (five of six participants completed the ratings). All of the participants agreed or strongly agreed that the intervention would be acceptable for students with writing problems and that it would be appropriate for a variety of students. Four of five agreed or strongly agreed that
they would tell other students to use the intervention, while the fifth slightly agreed.
Three of five participants strongly agreed that the intervention was effective at improving their writing, although their CWS scores couldn’t be aligned with these ratings since participants completed their ratings anonymously. Importantly, four of five participants agreed or strongly agreed that the intervention did not result in negative side effects on their writing, while the fifth slightly agreed. While the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985) has been used with treatment recipients in the past (Fan, Sidani, Cooper-Brathwaite, & Metcalfe, 2013), it has not been used with students. While a children’s version, the Children’s Intervention Rating Profile (CIRP; Witt & Elliott, 1985) exists, its statements are more limited than the IRP-15. The present study preferred to have more similarity and more directly comparable statements between the students’ social validity scale and the teachers’ social validity scale than the CIRP allows. As such, the IRP-15 was used as the basis for both scales. Future research should investigate the possibility of using the IRP-15 instead of the CIRP with older students.

Unfortunately, no teachers returned the teacher version of the IRP-15. As a result, it is not possible to ascertain teachers’ perceptions of intervention acceptability. The social validity scale was delivered to the participant’s teachers after the intervention had concluded, which corresponded with the final week of school. Specifically, the students’ last day was a Thursday and the teachers’ last day was the following Friday. It is very likely that the teachers did not complete and return the rating scale because of the time
constraint of only having five days to complete the scale and because of the busy, demanding end of the school year.

**Generalization**

The effects of the intervention appeared to demonstrate some skill carryover across a variety of measures. First, data pertaining to class grades revealed that Kyrie was able to raise his English grade from a B in the first semester of the year (i.e., pre-intervention) to an A in the second semester of the year (i.e., post-intervention). Unfortunately, the remaining five participants were not consistently enrolled in English courses during the semester immediately preceding the intervention, so no conclusions can be made for the effect of the intervention on their grades in English following the intervention. In fact, Alonzo and CJ were the only other participants who were enrolled in English class during the second semester of the year. Because neither was enrolled in an English class in the first semester of the year, no pre/post comparison for Alonzo and CJ could be made. The generalization of the intervention on participants’ grades can only suggest, then, that the intervention may have had positive effects on Kyrie’s overall improvement in English.

Second, additional writing skills assessed through the use of a sample cover letter indicated that Anderson was able to improve his cover letter writing performance from baseline to maintenance but Tyler was not. Prompts for Kyrie, Alonzo, Omri, and CJ were not administered at the end of the school year due to participant absences and the unavailability of the school counselor who administered the prompts. This generalization data suggests that Anderson may have been able to generalize what he learned to a job-
related skill. Overall, no conclusions can be made about the ability or inability of the intervention to generalize participants’ performances to writing cover letters because only one positive and one negative result were found.

Third, only Tyler’s teacher provided writing samples from pre and post intervention in-class writing assignments. The samples provided suggested that Tyler was able to generalize what he learned during the intervention to his in-class writing assignments. While CJ had the same teacher as Tyler at the beginning of the year, he switched to a new teacher, who had not been one of the original four teachers, at the end of the year. As such, no conclusions can be made about the ability of the intervention to generalize to the in-class writing performances of Kyrie, Alonzo, Omri, Anderson, or CJ.

Finally, results of high stakes state tests demonstrated that Kyrie, Anderson, and CJ were able to improve their performances on pre- and post-test administrations of the Ohio Graduation Test (OGT). Specifically, Kyrie and Anderson improved their performances from the limited to proficient range. CJ’s performance improved slightly, to just one point below the proficient range. Unfortunately, test performances for Tyler, Alonzo, and Omri were unavailable. Tyler enrolled in another school district following the intervention, making data regarding his OGT performance inaccessible. Because Alonzo and Omri were in ninth grade during the intervention and in 10th grade in the fall following the intervention, they had not yet taken the OGT (which is administered in the spring of 10th grade). Overall, for those participants whom state testing records were available, limited data suggest that they were able to generalize their writing performances to the state testing arena.
Other Data

Data also was collected on additional measures for comparison purposes. Specifically, correct minus incorrect writing sequences (CIWS; Espin et al., 2000), correct punctuation marks (CPM; Diercks-Gransee, Weissenburger, Johnson, & Christensen, 2009; Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002), and total words written (TWW; Deno, Marston, & Mirkin, 1982) data were examined descriptively to identify any trends useful for practitioners.

CIWS. Overall, CIWS tended to yield lower scores than CWS and TWW. On the other hand, CIWS appeared higher than CPM for all participants, with exception to two of Kyrie’s baseline writing samples. Interestingly, when examining trends between CWS and CIWS, correspondence between scores emerged. That is, when CWS increased/decreased, CIWS scores increased/decreased. For participants who made more errors (e.g., Kyrie, Alonzo, Omri, and CJ), CIWS was especially “brutal” and visibly different from CWS. Specifically, CIWS was much lower than CWS and TWW because the participants were penalized for errors. Interestingly, all four of these participants were able to decrease the amount of errors they made from baseline to intervention. For participants who did not make many errors, such as Anderson and Tyler, CIWS scores tended to cluster near CWS scores. The ability of CIWS to capture errors in addition to correct writing is an attribute unique to CIWS that the other CBM progress-monitoring hopefuls (CWS, CPM, and TWW) do not have. While grading Alonzo’s performance during one session, he mentioned, “Oh, but my mistakes don’t count against my total, right?” This comment highlighted one of CWS’s negatives—that mistakes do not count
against a total. It is also interesting to note that Alonzo was the only participant who seemed to try to want to improve his CWS score by simply writing *faster* each time. He sometimes wrote so quickly that he would have to take a few seconds to stretch or shake out his hand, and the legibility of his handwriting suffered as a consequence of his desire to write *more* instead of simply writing *better*. His desire to write faster without regard to making mistakes led him to make errors. For comparison, Anderson would take more time to think and pause while writing, and he was still able to demonstrate improvements without sacrificing the legibility of his handwriting and making more errors. In CIWS, then, a student could write 20 sequences correctly and 10 sequences incorrectly one session, and 25 sequences correctly but 20 sequences incorrectly the next session (in fact, Kyrie even received negative CIWS scores on two baseline measures). CWS would suggest that the student’s performance improved, when the number of errors and content of the sample would suggest no improvement. CIWS takes into account those errors, which is an improvement over CWS. As such, results of the present study would support the use of CIWS as an alternative to CWS, especially for its utility in assessing the writing of students who tend to make many errors. The present study is thus consistent with prior research on CIWS with high school students (e.g., McMaster & Espin, 2007).

**CPM.** Overall, CPM tended to provide participants with scores that were much lower than CWS, CIWS, and TWW. The only instance in which CPM was higher than any score was during two of Kyrie’s baseline samples—when he wrote more incorrect sequences than correct sequences and received negative CIWS scores. CPM did not tend to vary much from one sample to the next and did not demonstrate any correspondence
with the other measures (i.e., tend to increase or decrease when CWS increased and
decreased). It is possible that CPM scores were limited and low due to the fact that
participants wrote for only a limited amount of time (three minutes) and used hardly any
punctuation because few sentences (the main method of earning a CPM) were written.
As a result, students were unable to write many sentences, which could explain the
limited variability within scores. Having such little variation in scores is not an attribute
that is desirable in a CBM progress-monitoring outcome measure because there is little
chance for the students to be able to demonstrate growth. In order to demonstrate
measureable changes in performance, a more sensitive measure is desired. While the
present study would not support the use of CPM as an outcome measure for three-minute
writing samples, perhaps CPM would be appropriate for writing samples with longer time
limits. The present study is consistent with other research on CPM and time limits for
writing fluency measures of high school students (Diercks-Gransee et al., 2009; Gansle,
et al., 2002; McMaster & Campbell, 2008).

**TWW.** Overall, TWW tended to yield higher scores than CWS, CIWS, and TWW. For all participants, TWW scores always were higher than CPM. When
examining such patterns more closely, TWW scores appeared to be much higher than
CWS and CIWS in those participants who made more errors (i.e., Kyrie, Alonzo, Omri,
& CJ). In addition, TWW scores tended to vary in the same direction as CWS. That is,
when CWS and CIWS increased/decreased, TWW typically increased/decreased (though
there were exceptions). While previous studies have questioned the validity of TWW for
high school students, perhaps its only variance from CWS and CIWS is in the magnitude
of performance—that students tend to have higher TWW scores than other scores. However, the validity of TWW still poses a concern since incorrect spelling, capitalization, grammar, syntax, semantics, or other writing conventions do not affect the TWW score. While simply getting students to write more words, regardless of spelling, may be an important factor for younger students, the content (including correct spelling, grammar, syntax, etc.) of what is written is more valuable when assessing the writing of high school students. The findings in the present study are consistent with other recommendations for the use of TWW with older students (Amato & Watkins, 2011; Espin, Scierka, Skare, & Halverson, 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007; Tindal & Parker, 1991).

**Treatment Integrity**

While not a core research question, it is important to note that a very high level of treatment integrity was maintained. Specifically, treatment integrity as reported by the interventionist (i.e., the researcher) throughout the study was 99%. Interobserver reliability checks occurred during 40% of intervention sessions. During these sessions, the interventionist and interobserver reached 100% agreement that 98% of steps were completed. It is possible that such a high level of treatment integrity was created by the requirement of the researcher to self-report during every intervention session. The researcher had all of the intervention’s steps at her fingertips during every intervention session and she had to check off each step as she completed it. This constant self-monitoring likely contributed to the high level of fidelity. The self-monitoring was minimally intrusive and enhanced treatment integrity. Steps may be more likely to be
missed if an interventionist is not checking off steps as they are completed. In fact, the steps that were missed in the present study were missed because the interventionist failed to check the self-monitoring list when completing those steps. While having more than one interventionist may introduce more variables that could lead to lower treatment integrity, having the interventionists check off steps as they complete them could be one way to improve treatment integrity. Additionally, simply knowing that someone else is going to come in and monitor one’s treatment integrity could lead to improved fidelity. The interventionist may want to perform well for the observer to ensure that his or her own self-monitoring agrees with what the observer finds. A final reason as to why treatment integrity was high was that the interventionist designed the intervention and the self-monitoring checklist. By being heavily involved in the intervention’s design and in the writing of the self-monitoring checklist, the interventionist repeatedly reviewed each step of the intervention to the point of nearly memorizing the checklist prior to the start of the intervention. Such familiarity and in-depth knowledge of the intervention likely increased fidelity. In other subsequent studies, researchers may want to involve participants in the creation of the intervention and in the development of treatment integrity tools. In practice, practitioners may want to sit down with interventionists and thoroughly review and practice an intervention’s steps before implementing it, thereby increasing the interventionists’ familiarity with and knowledge of the intervention.

**Significance of Findings**

The present study addressed several major gaps in the literature on writing interventions for high school students. First, it answered the call for more research on
applied, school-based writing intervention strategies for older students (Graham & Perin, 2007a, 2007b; National Commission on Writing, 2003, 2004; Rogers & Graham, 2008) by recruiting six high school students as participants. Second, for its outcome measure, the present study utilized CWS, which is a more robust and valid progress monitoring measure for high school students (Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007) than previous goal setting writing interventions have used (e.g., Hopman & Glynn, 1989; Seabaugh & Schumaker, 1994; Van Houten & McKillip, 1977). Third, the present study reported on two effect size measures for single-subject studies: (a) PND, which is currently the standard and most widely used measure, but which has come under increased scrutiny in the past few years as an inappropriate effect size measure (e.g., Parker & Vannest, 2009; Petersen-Brown, Karich, & Symons, 2012); and (b) NAP (Parker & Vannest, 2009). The present study reported both PND and NAP and offered side-by-side comparisons and discussions of the differences between the sometimes conflicting effect sizes generated by both measures. Fourth, the present study followed Kratochwill and Stoiber’s (2000a, 2000b, 2000c, & 2002) call for academic intervention studies to utilize more single-subject research designs. The present study employed a changing criterion design, which is a methodologically sound and clinically useful design that permits the demonstration of gradual changes in performance over the course of an intervention (Kazdin, 2010). Changing criterion designs fit well with goal setting interventions because the goals that are set make natural changes in criteria throughout the intervention.
Overall, the present study supports goal setting as an effective writing intervention for high school students. First, the present study added to the overall literature on writing interventions. The present study sought to utilize types of interventions that have been demonstrated to improve the writing performance of students within small group or individualized settings and that could fit into the schedules of high school students. Rogers and Graham’s (2008) meta-analysis of writing interventions revealed that interventions that utilized strategy instruction for planning/drafting, grammar, and goal setting interventions not only were the most effective at improving student writing performance, but also were considered to be well-designed and methodologically sound. Goal setting was chosen because it can incorporate other effective techniques, such as performance feedback and self-graphing. Goal setting also is supported within experimental and quasi-experimental studies for students in grades 4 through 12 (Graham & Perin 2007a, 2007b). As such, the present study expanded upon the extant writing intervention literature by contributing to the research on goal setting. To date, only five experimental and quasi-experimental research studies (Graham & Perin, 2007a, 2007b) and seven single-subject research studies (Rogers & Graham, 2008) have utilized goal setting to improve writing.

Second, and more broadly, the present study offered expanded detail on aspects of goal setting interventions that have sometimes gone unspecified in the past. In general, goal setting interventions have tended to incorporate other effective techniques, such as performance feedback (e.g., Chandler, 2003; Krohn, Skinner, & Fuller, 2012; Van Houten & McKillop, 1977), self-graphing (e.g., Figarola et al., 2008), and reinforcement
(e.g., Seabaugh & Schumaker, 1994) and reinforcement (e.g., Kruglanski et al., 1975; Locke, Shaw, Saari, & Latham, 1981; Schunk, 1983). Performance feedback is an effective component of goal setting interventions when it incorporates both positive and negative feedback (Kluger & DeNisi, 1996), is immediate (Codding, Feinberg, Dunn, & Pace, 2005), and is direct and explicit (Li, 2010). Self-graphing, which is a self-monitoring technique, is an effective component of goal setting when students are able to see graphical representations of their performances (DiGangi, Maag, & Rutherford, 1991; L. S. Fuchs & Fuchs, 1986, 1987). Reinforcement is another effective component of goal setting when reinforcement is contingent upon goal attainment, rather than for simple participation (Kruglanski et al., 1975; Locke et al., 1981; Schunk, 1983). One limitation of prior goal setting studies is the limited amount of detail that is provided on these additional components. The present study sought to address this limitation by detailing its performance feedback, self-graphing, and reinforcement procedures. The findings of the present study reinforced the ability of positive, negative, immediate, direct, and explicit performance feedback to improve academic performance. The findings also reinforced the abilities of self-graphing and goal-contingent reinforcement to improve academic performance. Overall, the findings of the present study support the effectiveness of goal setting on improving academic performance when the goal setting incorporates performance feedback, self-graphing, and reinforcement.

Third, and more specifically, the present study added to the limited literature base behind goal setting writing interventions with high school students. Few studies exist that have employed writing interventions with high school students. In fact, only three
previous studies have focused on the use of goal setting interventions with high school students (Hopman & Glynn, 1989; Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977). The primary flaw with these three studies was their use of an outcome measure that is less reliable and less valid for high school students: TWW and number of writing lessons completed. TWW has been identified as less reliable and less valid for high school students because it is more concerned with the quantity of the writing rather than the quality, as it does not account for errors (Amato & Watkins, 2011; Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007; Tindal & Parker, 1991). Words can be spelled incorrectly, make no grammatical sense, and be punctuated improperly, but they would still count toward the total score; TWW is merely a count of how many groups of letters separated by spaces are contained within a writing sample. Setting a goal for the number of writing lessons to complete is not a direct measure of a student’s writing ability, either. It is only measuring the amount of coursework completed. Such a measure may be a useful measure for generalization—that is, if a student’s writing improves, that student may complete more lessons—but it is not appropriate as a direct measure of a student’s writing performance. As such, the present study improved upon these previous studies by utilizing CWS as the goal setting measure and as the measure of the participants’ writing performances. Previous studies also provided limited generalization and maintenance data, and the present study sought to further improve upon the current literature base behind goal setting writing interventions with high school students by reporting on maintenance data and a variety of generalization data.
Taken as a whole, the present study supports goal setting as an effective writing intervention for high school students when the goal setting intervention (a) uses an outcome measure that is reliable and valid for high school students, such as CWS, and (b) incorporates performance feedback, self-graphing, and reinforcement. Because the three prior studies on goal setting writing interventions with high school students used less valid and less reliable outcome measures, and because these studies revealed mixed results as to the effectiveness of goal setting, much more research needs to be conducted on goal setting before it can be established as an evidence-based technique for improving the writing skills of high school students.

**Implications for Practice**

The present study demonstrated a variety of implications for educational practice that relate to: (a) using goal setting writing interventions at the high school level, (b) evaluating the feasibility of various curriculum-based measurement (CBM) and curriculum-based assessment (CBA) writing measures for high school students, and (c) overcoming identified barriers to the implementation of individualized interventions at the high school level.

**Goal Setting**

The present study demonstrated that goal setting is an effective writing intervention at the high school level. Although previous studies have used goal setting with interventions to improve high school students’ writing abilities (e.g., Hopman & Glynn, 1989; Seabaugh & Shumaker, 1994; Van Houten & McKillop, 1997), the results related to effectiveness were mixed. Prior limited findings could have been the result of
utilizing less valid and reliable outcome measures and not including components often found in other goal setting studies. Goal setting interventions often incorporate other features of effective interventions, such as reinforcement, self-monitoring, and performance feedback (e.g., Barry & Messer, 2003; Coddington, Chan-Iannetta, Palmer, & Lukito, 2009; Conte & Hintze, 2000; Eckert, Ardoin, Daly, & Martens, 2002; Frayne & Latham, 1987; Schnoll & Zimmerman, 2001; Seabaugh & Schumaker, 1994; Van Houten & McKillop, 1977). The present study specifically identified those features and demonstrated that when a goal setting intervention is comprehensive, it can produce effective outcomes for participants. Practitioners wishing to use goal setting in the field should consider additional elements such as reinforcement for meeting goals, self-monitoring of performance, and performance feedback, as such elements may create a condition of optimal learning.

It also is important that practitioners maintain treatment integrity when bringing research to practice (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000; Lane, Bocian, MacMillan, & Gresham, 2004). An evidence-based intervention will only produce the results that it is designed to produce if it is implemented in the way in which it was designed—that is, if it maintains internal validity (Lane et al., 2004). Self-monitoring checklists, such as the one used in the present study, and scripted interventions (i.e., manualized treatment protocol) are ways in which teachers may monitor their own treatment integrity (Gresham et al., 2000; Lane et al., 2004). Such checklists and scripts are simple and easy-to-use ways of staying on-track. Perhaps one of the reasons why the treatment integrity in the present study was high stemmed from
the fact that the researcher created an almost scripted method for delivering the intervention. Similarly, teachers could have input into the development and format of treatment integrity checklists that may serve to increase the likelihood of accurate and consistent implementation of the intervention. Other measures of treatment integrity could include direct observations of the teacher, post-intervention interviews, permanent products, and consultant feedback (Gresham et al., 2000; Lane et al., 2004).

**Progress-Monitoring and Outcome Measures for Writing Fluency in High School**

**CWS.** The present study demonstrated that the participants were able to set goals for their writing based on a valid CBM outcome measure, CWS, which was scored by the researcher. When engaging in goal setting writing interventions with high school students, practitioners should consider using CWS. CWS is currently the most reliable and valid measure for assessing high school students’ writings (Espin et al., 1999; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007). While CWS can be reliable and valid when used by practitioners, it is unknown whether students themselves can reliably score their own writing using CWS. Ideally, students would be able to self-monitor their writing at the high school level in order to encourage independence, but CWS requires the rater to be knowledgeable about correct spelling, grammar, syntax, and semantics. Students who struggle with writing may not be able to reliably self-monitor using CWS because they may not know when their spelling, grammar, syntax, or semantics is incorrect. While using CWS with high school students, practitioners should expect to score CWS themselves, but still involve students in the process and provide feedback.
**CPM and CIWS.** The present study reported on two other possibly valid writing CBM outcome measures for high school students: CPM and CIWS. The present study provided a direct comparison of CPM and CIWS with CWS and TWW via graphical presentations. The purpose of the present study was not to compare the reliability or validity of these four measures against each other, but rather to offer data through direct comparison in graphs that could add to the conversation about which of these measures should be utilized at the high school level. The data from the present study revealed CIWS as a promising outcome measure for high school students’ writings, confirming previous research (Espin et al., 2000; Espin et al., 2008; McMaster & Campbell, 2008). CIWS is sensitive to incorrect writing sequences, which is not measured by CWS. Incorrect sequences are important to consider, especially for students who make many errors. By using CIWS instead of just CWS, the interventionist can put equal emphasis on producing correct writing and decreasing writing errors.

The present study questions the validity of CPM for measuring work obtained during brief writing sessions. CPM may have promise as an outcome measure in longer writing sessions (e.g., Diercks-Gransee et al., 2009), but it does not appear to be useful as an outcome measure for high school students during brief writing sessions because few sentences (the main factor in calculating CPM) were written within the three-minute sample. The small range of CPM that was yielded would provide for little sensitivity to change, which is an important quality when studying a student’s response to an intervention. Therefore, practitioners should continue using CWS and CIWS on fluency assessments, but may be interested in using CPM on longer sample.
CBM time limits for high school students. Additionally, practitioners should consider using longer time limits for gathering data about writing fluency at the high school level. A wide variety of other time limits have been investigated by researchers, but five and seven minutes have promise as more valid time limits than three minutes for measuring writing fluency at the high school level (McMaster & Campbell, 2008).

Interventions in High School

Scheduling. The difficulty of implementing interventions at the high school level has been a consistent concern within the extant literature. One reason is that scheduling difficulties arise when students need to earn certain numbers of credits in various classes in order to graduate, and removing students from class for interventions may negatively impact credit accumulation (L. S. Fuchs, Fuchs, & Compton, 2010; Sansosti, Telzrow, & Noltemeyer, 2010; Vaughn et al., 2010; Windram, Scierka, & Silberglitt, 2007). The present study demonstrated that individualized writing interventions could fit into students’ daily schedules. Collaboration with teachers and students as to free time during the day was a key factor in the ability of the present intervention to be included within the participants’ schedules two to three times per week. By working with teachers, staff, and students, and by being flexible about available intervention space, schools should be able to carve out time for individualized interventions.

Use in the classroom. The procedures outlined in this intervention can be incorporated easily into an in-class writing activity to monitor writing performance for all students. Goals and feedback could be produced and provided individually during student writing conferences, similar to the individualized feedback provided in the
present study. In this way, an individualized writing intervention could be conducted with all students on a one-on-one basis in the classroom. In addition to helping low-achieving students improve their writing abilities, average-achieving students could improve their performance to the high-achieving level, and high-achieving students could improve their performance to an even higher level. This type of individualized intervention encourages differentiation in the classroom. Practitioners should differentiate instruction based on student performance and response to interventions; goal setting provides one way to make data-based decisions about differentiating instruction.

Student motivation and participation. Student motivation and willingness to participate in interventions also has been identified as a possible barrier to successful interventions at the secondary level (L. S. Fuchs et al., 2010). The present study incorporated motivation, creativity, and self-monitoring, which have been identified as possible techniques to improve intervention effectiveness in secondary schools (L. S. Fuchs et al., 2010; Vaughn & Fletcher, 2010). Motivation was encouraged through the ability of the participants to earn reinforcers for improved writing. Allowing the students to think of continuations of a story prompt incorporated creativity. Self-monitoring was incorporated through the use of self-graphing. Practitioners interested in implementing interventions at the high school level should also incorporate motivation, creativity, and self-monitoring into their interventions.

Limitations

Although results of this study indicated favorable outcomes, the study does have limitations. These limitations concern: (a) teacher participation in data collection, (b) the
timing of the intervention within the school calendar, (c) the generalizability of the population, (d) the sample size, (e) the genre of the writing prompts, (f) the intervention’s three-minute time limit for writing, (g) the use of the IRP-15 with students, and (h) the limited maintenance data.

First, the ability to draw conclusions about the teachers’ perceptions of the intervention (and the generalization of the intervention to the regular education classroom) is limited due to incomplete participation. Although each of the teachers interacted collegially with the researcher and gave no indication as to why writing samples, weekly logs, and the teacher perception ratings were not returned for all participants, it is difficult to make any conclusions regarding how other educators may view the intervention. A series of personal inquiries/emails before the end of the school year, as well as multiple attempts to follow up with each teacher, produced some of the missing data, but not all. It is possible that the year-end and day-to-day demands on teachers may have made completion of forms for a research study less important than other school-based requirements.

Second, the fact that the end of the intervention coincided so closely with the end of the school year made it difficult to gather maintenance data. While this study proposed to gather three maintenance data points if possible, participant absences and availability at the end of the school year made only two data collections possible for three participants, one data collection possible for two participants, and zero data collections possible for one participant. While collection of maintenance data due to unanticipated issues at the end of the school year is not an uncommon occurrence in school-based
research, it does limit the ability to ascertain if skills taught as part of the intervention continued. Future studies may wish to implement interventions earlier in the school year so that end-of-school-year demands and the call of summer vacation do not interfere with data collection. Longer data collection periods would allow for a more extended investigation of maintenance effects.

Third, all participants in this study were male and attended an alternative high school. Taken together, such limitations may negatively affect the generalizability of the findings to female students and other high school settings. As such, it is possible that female students may respond differently to the intervention. Additionally, the participants received their education in a public, high poverty urban district and had started their secondary education at traditional high schools. Each of the participants in this study applied to “open enroll” in the alternative school in order to recover failed credits and “get back on track” for graduation. As part of the enrollment process, students were interviewed by the school to determine, among other things, whether or not they possessed the motivation to complete the credit recovery programs and whether the school’s half-day setup would fit their needs. This application, interview, and enrollment procedure is different from other high schools in the district and is not common for public high schools in general. It is possible that students enrolled at the alternative high school have some characteristics that separate them from other high school students within the district. That is, they may be applying to the school because they have been unsuccessful in a traditional school setting. Consequently, their academic histories may be more limited than the typical high school student and/or they may demonstrate lower academic
abilities. Moreover, home-based factors (e.g., a need to work, a need to attend school part-time due to childcare or other reasons, or health concerns) could have contributed to the students enrolling at the alternative school. These participants may have taken the initiative to contact the school, fill out an application, and complete an interview, all of which potentially indicates that they have heightened initiative to perform well (e.g., recover coursework in order to graduate). So, the population of students at this high school may be experiencing a different kind of motivation than the typical high school student.

A fourth factor, which also concerns generalizability, is that only six students participated in this study. While six participants is an acceptable number for single-subject research, these six participants demonstrated some variability when comparing their performances. On one hand, such variability may accurately reflect what a teacher is likely to encounter when working with students one-on-one, rather than using data from studies with large participant populations and trying to generalize that data to the small group or individualized level of intervention. On the other hand, such a limited sample requires caution to be taken when generalizing the findings across students, settings, and/or other academic behaviors.

A fifth limitation is that the participants were provided with narrative writing prompts, which may not be reflective of most of the writing assignments given in class or on tests. While some classes, such as English, may require students to respond to narrative prompts, other classes (e.g., science, social studies) may have more expository or persuasive writing demands and fewer narrative writing opportunities. A key step in
improving a student’s writing ability, however, is simply to get them to write more of any kind of writing (Santangelo & Olinghouse, 2009). In fact, student writing performance improves when they are required to write throughout the curricula in many classes, various genres, and several formats (Santangelo & Olinghouse, 2009).

A sixth limitation is that the participants were provided only three minutes to write. This study was meant to be a quick and easy to implement individualized intervention. It was intended to fit easily into the participants’ schedules and to take as little time away from instruction as possible. In that sense, and in the sense that the study consistently followed AIMSweb administration procedures, it was successful. However, some researchers have suggested that such quick writing tasks should be expanded to five or more minutes for older students (e.g., McMaster & Campbell, 2008). While there is little research on the difference between the outcomes of three versus five minute or other time limits for older students, it presents is an intriguing area ripe for further exploration.

A seventh limitation relates to the social validity (i.e., treatment acceptability) measure for the participants. While the IRP-15 is most often given to teachers to gauge their perceptions of the effectiveness of an intervention, it is used just as frequently in a variety of settings to study a variety of scenarios. It was originally developed by asking a group of teachers to read hypothetical scenarios of student misbehavior. The teachers were provided with a possible intervention and were asked to use the IRP-15 (Martens et al., 1985) to rate their perceptions of the intervention. The reliability of this measure was tested and reported using these participants. In practice, the language of the IRP-15 is frequently modified to reflect the current circumstances in which it is being utilized. The
IRP-15 is often used in scenarios that have a rater who is not a teacher, that have a rater who implemented the intervention instead of having just read about the intervention, and/or that are assessing an intervention that is not related to behavior. For example, the IRP-15 has been given to parents instead of teachers (e.g., Scattone, 2008), given to teachers who have implemented an intervention (e.g., Wright & McCurdy, 2012), given to teachers to rate their perceptions of an intervention implemented by someone else (e.g., Power et al., 2012), used to gauge perceptions of academic problems (e.g., Resetar, Noell, & Pellegrin, 2006), and given to intervention participants (e.g., Fan et al., 2013). While few studies exist like the Fan et al. (2013) study where the IRP-15 was provided to the intervention participants, the social validity data is still useful and interpretable.

Some studies do not use statistically sound social validity measures. In fact, some guidelines for assessing social validity discuss the use of informal questionnaires, interviews, or other surveys (e.g., Alberto & Troutman, 2009), or even developing one’s own social validity ratings scale (e.g., Rathvon, 2008) as important components of evaluating interventions. As such, while the applicability of the specific statistics behind the IRP-15 as related to high school students may be a limitation of the present study, the data provided by the social validity measure is still useful as an informal measure of treatment acceptability and still supports that the students, overall, gave more positive than negative ratings to the items on the questionnaire. While a children’s version, the CIRP (Witt & Elliott, 1985) exists, that rating scale was created in a similar method to the creation of the IRP-15. The participants were students in the sixth grade who rated their perceptions about an intervention for a scenario about which they read. Students in sixth
grade may need the simpler CIRP in order to express their opinions about an intervention, but students in high school may be able to provide just as valid ratings using the more “adult-friendly” IRP-15 as when they use the more “child-friendly” CIRP. This is an interesting area for future research, as well.

The final limitation concerns the incomplete sets of maintenance data that were collected. While constricted and truncated maintenance periods are not uncommon in the literature (e.g., Viel-Ruma et al., 2010), the maintenance period in the present study was less than ideal. Ideally, to demonstrate experimental control and to confirm the trend of data within a phase, at least three data points would be desired within each phase of an intervention (Kazdin, 2010). While maintenance periods with only one or two data points also are not unusual, they allow for limited interpretation of the maintenance of skills because they cannot establish a data trend within the maintenance phase. Future studies should strive to establish maintenance periods that evaluate trend (i.e., at least three data points) over time (i.e., in increasing increments of time following the conclusion of the study). Because time limits related to school year calendars can negatively impact researchers in the schools, contingency plans for collecting maintenance data could be developed in the case that the ideal situation for evaluating maintenance is interrupted by school breaks or changes in the school schedule.

**Recommendations for Future Research**

The present study supports multiple areas of future research, many of which stem from the implications for practice. These recommendations stem from: (a) conducting goal setting writing interventions at the high school level, (b) evaluating the feasibility of
various CBM writing measures for high school students, and (c) comparing the use of
percent nonoverlapping data (PND) and nonoverlap of all pairs (NAP) as possible effect
size measures for single-subject studies.

**Goal Setting Writing Interventions in High School**

First, researchers should take care to specify when their goal setting writing
interventions include all, none, or some of the multiple components that goal setting
interventions may include. These components could include self-monitoring,
performance feedback, and reinforcements, as the present study did, or different
combinations of other components. In order to enhance the reliability, validity, and
generalizability of findings related to goal setting writing interventions, researchers
should take care to specify what components the goal setting entails and how such
components were implemented. Moreover, goal setting writing interventions should
include goals that directly relate the students’ writings. Goals such as number of lessons
completed (e.g., Seabaugh & Schumaker, 1994) are not directly tied to a student’s writing
performance, thus questioning the validity of the use of the number of lessons completed.
While the results of the present study are promising for goal setting as an effective
intervention for improving writing at the high school level, many more studies are needed
to establish a solid research base that supports goal setting as an evidence-based multi-
element intervention.

Second, motivation, creativity, and self-monitoring have been identified as
possible techniques to improve intervention effectiveness in secondary schools (L. S.
Fuchs et al., 2010; Vaughn & Fletcher, 2010). The present study incorporated motivation
through reinforcement for meeting goals, creativity through narrative writing, and self-monitoring through self-graphing. Researchers should strive to include similar techniques within writing interventions.

Third, researchers should monitor social validity (i.e., treatment acceptability). The IRP-15 (Martens et al., 1985) has been used with teachers in evaluating a variety of interventions and is just one example of the many ways of evaluating treatment integrity. Many statistically-validated measures exist, but researchers can also use informal measures of social validity (Alberto & Troutman, 2009; Rathvon, 2008). The present study used the IRP-15 with students, and while it has been utilized with intervention participants before (e.g., Fan et al., 2013), future researchers should investigate the possibility of its use with students as intervention participants. Comparing the IRP-15 and the CIRP (Witt & Elliott, 1985) with high school students also may provide needed data on which form may be suitable for different ages of participants.

Fourth, researchers should evaluate generalization when assessing any writing intervention. If an intervention is effective only in the setting in which it takes place, then that intervention’s external validity is questioned. When a student can translate what he or she learns from an intervention to another setting, the intervention has then gained external validity; the intervention’s external validity is decreased when the student cannot demonstrate improvements in performance across settings. In addition to evaluating social validity, then, researchers should evaluate generalization. The present study utilized classroom grades, performance on high-stakes tests, classroom writing samples,
and cover letter writing samples, and these may be useful generalization measures for other researchers.

Finally, treatment integrity also should be monitored in future research on writing interventions. The present study’s use of self-report and inter-observer checks follows evidence-based suggestions for monitoring treatment integrity (e.g., Lane et al., 2004). Effective monitoring of treatment integrity enhances an intervention’s internal validity, external validity, and criterion validity (Gresham et al., 2000).

**Writing Fluency Progress Monitoring Outcome Measures**

There are many avenues of future research available in the investigation of writing CBM measures. Currently, there is no published research on benchmark or normative indicators and expected rates of improvement for CBM writing measures for high school grade levels. Companies that offer data about writing fluency outcome measures, such as AIMSweb, need to expand their data into the high school level. AIMSweb only offers benchmark data for students up through eighth grade (AIMSweb, 2012). Other researchers also should provide benchmark data for writing fluency measures (along with reading and math) at the high school level, in order to increase the reliability and validity of the benchmark data. Additionally, while CWS is currently the gold standard in evaluating the writing fluency of high school students, researchers are encouraged to continue to investigate the utility of CIWS for brief writing sessions and both CIWS and CPM for longer writing sessions. Other promising outcome measures identified in the earlier literature review, such as mean length of correct writing sequences (MLCWS), number of characters per word (characters/word), total sentences
written (total sentences written), and incorrect writing sequences (IWS), should continue to be investigated (Diercks-Gransee et al., 2009; Espin et al., 1999; Gansle et al., 2002; Videen et al., 1982). As mentioned earlier, students ideally would be able to self-evaluate their own writing samples so that they could monitor their writing performance independently. Research also should be conducted on how accurately students can evaluate their own writings using CWS and other measures. Also related to CBM is the issue of time limits. Researchers should continue to investigate the use of longer time limits when gathering data on writing fluency for high school students. The current AIMSweb standard of three minutes may be too short for high school students, but five or seven minute time limits may be more appropriate. More research is needed in this area (McMaster & Campbell, 2008).

**PND versus NAP**

Finally, the present study also provided data on both PND and NAP as effect size measures for single-subject studies. The purpose of the present study was not to delve into an in-depth analysis of which effect size measure is more reliable or valid in single-subject research, but rather to offer a side-by-side comparison of these two effect sizes and add to the conversation about their validity. PND and NAP tended to agree on the outcome for most of the participants. For example, *moderate* PND outcomes tended to coincide with *medium* NAP outcomes. There were a few exceptions. Kyrie’s baseline to intervention and baseline to maintenance effect sizes revealed *no* effect, according to PND, but a *medium* effect, according to NAP. Tyler’s baseline to intervention effect sizes revealed *no* effect, according to PND, but a *medium* effect according to NAP. CJ’s
baseline to maintenance effect sizes also revealed no effect, according to PND, but a medium effect according to NAP. One might conclude that NAP provides more positive outcomes, based on these participants’ results, but that was not always the case. Anderson’s baseline to maintenance effect sizes revealed a large effect, according to PND, but a medium effect according to NAP. Perhaps NAP provides more moderate results, rather than more results at the extreme ends of effectiveness (i.e., no effect and large/high effect). Additionally, perhaps PND would be more appropriate for certain designs where the upper and lower limits of performance during the intervention period are not confined by the study’s design (e.g., AB-based designs), whereas NAP may be more appropriate for designs where performance during intervention is directly impacted by the study’s design and could affect the amount of change seen during intervention versus during baseline (e.g., changing criterion designs). These are areas that future research on PND versus NAP should consider. Practitioners will want to investigate the theories behind PND and NAP when making the decision of which measure to use in the field. NAP took only slightly longer to calculate by hand in the present study than PND, so time may not be a deterrent to using NAP in the field. In practice, PND may be a measure that is more easily explained to and understood by parents, educators, or others to whom a practitioner may be presenting data. Practitioners and researchers may want to wait for more research on NAP, however, before making a complete switch from PND to NAP or before ruling out NAP entirely. Researchers are thus encouraged to use and report on both PND and NAP, and to continue to compare the two measures.
Conclusions

Overall, the present study supports goal setting as an effective writing intervention for high school students. Specifically, this individualized goal setting writing intervention was effective at improving the writing abilities of five of the six participants, with mixed data about the intervention’s effectiveness for the sixth participant. Participants demonstrated marked increases in performance from baseline to intervention. They were generally able to maintain the positive effects of the intervention on their writing performances in the week following the intervention’s conclusion. The students also were able to generalize their improved writing performances to other settings, though only limited quantities of generalization data was available. All of the participants found the intervention acceptable. Unfortunately, none of the participants’ teachers returned the acceptability rating scale, so it is unknown whether or not the teachers also found the intervention acceptable. The teachers were not directly involved in the intervention and this may have negatively impacted their completion of data forms and the rating scale. The present study provided comparisons of various outcome measures for high school students: CWS, TWW, CIWS, and CPM. The present study supports CWS and CIWS, questions the validity of TWW for high school students, and questions the validity of CPM only in brief writing sessions. The present study also provided comparisons of two effect size measures for single-subject research, PND and NAP, and suggests that single-subject researchers report both effect sizes in future research and continue to compare and contrast PND and NAP. The present study supports goal setting as an effective writing intervention for high school students, so long
as the goal setting intervention: (a) uses an outcome measure that is reliable and valid for high school students and (b) incorporates performance feedback, self-graphing, and reinforcement. While this multi-element intervention joins both individual studies and extant meta-analyses (e.g., Graham & Perin, 2007a, 2007b; Rogers & Graham, 2008) in supporting goal setting as an effective writing intervention, more research still needs to be conducted in order for goal setting to be considered an evidence-based writing intervention for high school students. For the time being, the present study supports the use of a goal setting writing intervention as an effective way of improving the writing skills of high school students on an individualized level of support.
APPENDIX A

DATA SHEET FOR CLASSROOM TEACHER
Appendix A

Data Sheet for Classroom Teacher

**Student Name:** ____________________________

**Teacher Name:** ___________________________

**Directions:** Please keep a weekly log of the student’s grades for written assignments. Each week, please write down the dates for that week, the type of assignment (e.g., short essay, long essay, journal, OGT practice), and the grade the student received. If multiple writing assignments are given in one week, please give them each a new line. If no writing assignments are given in one week, please record that week’s dates and write “None” under the title and grade columns. You may use the other side of this paper, if needed.

<table>
<thead>
<tr>
<th>Week Dates (e.g., 2/4/2012-2/8/2012)</th>
<th>Writing Assessment Type</th>
<th>Student’s Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week Dates (e.g., 2/4/2012-2/8/2012)</td>
<td>Writing Assessment Title/Name</td>
<td>Student’s Grade</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

COVER LETTER PROMPTS PROVIDED BY COUNSELOR
Appendix B

Cover Letter Prompts Provided By Counselor

The following pages contain the job postings for which the school counselor asked the participants to write a cover letter. The participants were simply asked to write a cover letter in response to the job posting. Names and phone numbers of local companies and people have been redacted in order to maintain the confidentiality of the location of the participants. National chain names were not redacted, but the locations of the chains were redacted.
Taco Bell Team Member job in [Redacted]

Company: Taco Bell
Job Title: Team Member
Job Type: Full-time, Part-time
Hours: Varies
Pay Type: Hourly
Wages: To be discussed
Location: [Redacted]

We're looking for Service and Food Champions who love serving customers, have experience in the restaurant industry and who want to be a part of the largest restaurant company in the world! If you want to build a great career while providing fast, fun and friendly service to our customers, Taco Bell is the perfect place to learn, grow and succeed!

We offer the following:

• A commitment to promote from within
• Excellent training programs
• Reward and recognition culture
• Family friendly environment

The ideal candidates must want to have fun serving great food to our customers!
Job Summary

Company: [Redacted]
Location: [Redacted]
Industries: Hotels and Lodging, Restaurant/Food Services
Job Type: Part Time, Temporary/Contract/Project
Years of Experience: 2+ to 5 Years
Education Level: Some High School Coursework
Career Level: Entry Level
Salary: 8.00 - 9.00 USD/hour
Contact Information: Crown Services Inc
Phone: [Redacted]

Banquet Servers
About the Job

We are looking for part-time banquet servers!
Must be able to work weekends!
You have to supply your own clothes for the opening!
Starting pay is $6.00 to $8.00 an hour!
We need people that are serious about working part-time!
Call [Redacted] to apply today!
Job Summary

Company
Restaurant

Location

Industries
Food and Beverage Production

Job Type
Full Time

Contact Information

Phone:

Server/Waitstaff

About the Job

Job Purpose:
Serves patrons by providing information to help food and beverage selections; presenting ordered choices; maintaining dining ambiance.

Duties:
* Prepares room for dining by clothing tables and setting decorations, condiments, candles, masking, service plates, and utensils.
* Photocopies establishment and patrons by adhering to sanitation, safety, and alcohol beverage control policies.
* Helps patrons select food and beverages by presenting menu, offering cocktails and appetizers, suggesting courses, explaining the chef’s specialties; identifying appropriate wines; answering food preparation questions.
* Transmits orders to bar and kitchen by recording patrons’ choices; identifying patrons’ special dietary needs and special requests.
* Keeps kitchen staff informed by noting timing of meal progression.
* Serves orders by picking up and delivering patrons’ choices from bar and kitchen; delivering accompaniments and condiments from service bars.
* Responds to additional patron requests by inquiring of needs; observing dining process.
* Maintains table setting by removing courses as completed, replenishing stemware, refilling water glasses, being alert to patron spills or other special needs.
* Concludes dining experience by acknowledging choice of restaurant; inviting patrons to return.
* Obtains revenues by totaling charges, issuing bill; accepting payment, delivering bill and payment to hostperson; returning change or credit card and signature slip to patrons.
* Contributes to team effort by accomplishing related results as needed.

*Please apply in person!
Skills/Qualifications:
Job Summary

Company
Location
Industries
Other/Not Classified
Job Type
Full Time
Employee
Job Reference Code

CASHIER

About the Job
Cashier
Looking for energetic, self motivated individuals for our busy c-stores. Apply at:
Interested applicants contact:
Restaurant Manager

If you have 2+ years of full service, casual dining restaurant management experience, we invite you to join us and surround yourself with people who share our values.

We celebrate and reward our Managers by offering a variety of benefits along with a "no ceiling" philosophy in bonus potential.

What more could you ask for?

- Restaurant expansion - We are growing!
- Career growth opportunities
- Amazing culture and support structure
- Base, Bonus and benefits package
- Quality of life career programs

So bring your restaurant management experience to Red Robin and make the ordinary extraordinary and the mundane fun!

Related Keywords: front of house restaurant manager food service management food dining eatery kitchen red robin ARH FOH
Red Robin Assistant Managers Assistant Restaurant Management Assistant Management Assistant General Managers
APPENDIX C

INTERVENTION NOMINATION FORM
APPENDIX C

Intervention Nomination Form

Student Name: ______________________

DOB: ______

English Teacher Name: _________________________________

Student Demographic Data (for informational purposes only):

  Age: ______  Grade: ______  Race/Ethnicity: ________________

  Credits remaining to graduation: ______  Sex: (circle one) Male      Female

Does the student have an educational disability? (circle one) Yes     No

Student Scoring Data (record most recent scores):

  AIMSweb Writing Fluency CWS Score: ______________

  Current GPA: _________  Current English grade(s): _____________

  OAA/OGT Writing: __________

  Other classroom-based assessments: ______________________________

  Other curriculum-based assessments: ____________________________

  Other state assessments: _________________________________

Student’s Free Periods Schedule: (circle) Mon   Tues   Wed   Thurs   Fri

  Free Period Times: ___________________

Will this student still be enrolled in high school eight weeks from now? Yes -or- No

Nomination Date: ____________________
APPENDIX D

INFORMED CONSENT TO PARTICIPATE IN A RESEARCH STUDY
Informed Consent to Participate in a Research Study

**Study Title:** The Effectiveness of a Goal Setting with Reinforcement Intervention for Improving the Writing Skills of High School Students

**Principal Investigator:** Katherine B. Lavik, M.Ed.

**Purpose of the Study:**
The purpose of this study is to investigate how a goal setting writing intervention will affect the writing ability of high school students. Compared to the number of studies on interventions for reading and math, the number of studies on writing interventions is very small. When narrowing down the field of writing studies to those that have been conducted with high school students, the number declines even further. Writing is an essential life skill, however, and more research needs to be conducted to find out the best ways to improve the writing skills of high school students. Writing is important for high school, college, and jobs. An employee’s writing ability is a significant factor in hiring and promotional decisions at work. In college, coursework includes paper-writing, essays, and extended responses as measures of student achievement. Unfortunately, only 37% of eighth grade students and only 31% of 12th grade students have achieved at least a proficient level of writing ability. As such, business often have to remediate the writing abilities of their employees—a process that wastes company time, money, and productivity. Fortunately, there has been enough research on writing interventions for high school students in order to recommend some components of effective interventions. Based on meta-analyses conducted by a number of prominent researchers in the field of writing instruction, twelve interventions have been demonstrated to improve writing at all grade levels. These interventions are strategy instruction, summarization, peer-assisted writing, goal setting, handwriting/spelling/typing, word processing, sentence combining, inquiry, prewriting, process writing, writing-to-lean, and model study. The purpose of this study is to investigate how a goal setting intervention will affect the writing performance of high school students.

**Procedures:**
The student will meet one-on-one with the researcher three times per week to receive the intervention during a free period. The researcher is a doctoral school psychology intern who is employed by the school district. The researcher has been trained in this intervention technique. The student will write a quick writing sample during each intervention period. The student and researcher will then score the writing sample together. The writing sample will be scored with Correct Writing Sequences (CWS), a way to easily measure correct grammar, spelling, and punctuation. By scoring the sample together, the student will have a better understanding of what mistakes were made, what was correct, and how to improve during the following session. The student
and researcher will then set a goal CWS score for the student’s writing for the next session. The student will receive one point for meeting or exceeding the goal. For every three points the student earns, he or she may receive a small reward (e.g., candy, eraser). For every six points the student earns, he or she may receive a large reward (e.g., one mp3 download, a $5.00 gift certificate to a fast food restaurant).

Additionally, the student will be encouraged to talk about how he or she is able to use what he or she has learned during the intervention in other settings. In order to see if the student is using what he or she has learned during the intervention in other settings, the student’s English teacher will be asked to report on weekly grades. These grades will only be used to measure the degree to which the intervention will affect classroom performance. If the student has produced writing samples in other classes, the student’s teachers for those classes will be asked to provide a writing sample from before and after the intervention to see if the intervention affected writing in other class work. The student’s counselor will also give another job-related writing sample before and after the intervention. It is the intent of this intervention to improve writing ability not only in the intervention setting, but in other settings, as well, such as the classroom and for job applications.

The intervention will last for six weeks. A baseline period will precede the intervention to assess how the student scores on the writing sample before setting goals and earning reinforcements. A maintenance period will follow the intervention to see if the student is able to continue using what he or she has learned even after he or she is no longer setting goals or earning reinforcements with the researcher. The baseline and maintenance periods will simply consist of the student writing a quick writing sample.

**Benefits:**

Writing is an important life skill that often receives less instruction and less intervention than reading and math. By participating in this intervention, the student will have an additional opportunity outside of his or her regular education classes to improve his or her writing skills. During the intervention, the student will be encouraged to use what he or she has learned during the intervention in other settings, such as classes, job applications, and college. The student’s grades and writing ability may increase as a result of this intervention.

**Risks/Discomforts:**

There are no anticipated risks beyond those encountered in everyday life. The writing prompts are designed to elicit creative narrative writing and are not designed to elicit any feelings of discomfort.

**Privacy and Confidentiality:**

Identifying information will not be made available to any publications or presentations of the research data. The writing samples, goal setting charts, and weekly teacher reports will be kept confidential within the limits of the law. All materials that
contain identifying information will be kept in a secure location and only the researcher will have access to the data. Research participant data will be reported using pseudonyms. According to the law, confidentiality will not be maintained if there is an indication that the student may harm himself or herself or others. Consent forms will be kept in a locked cabinet in the dissertation advisor’s office.

In order to ensure that the researcher is following protocol, another school psychology graduate student may observe a few intervention sessions and/or score a few writing samples from the student. The observer will be observing the researcher and completing a checklist to make sure that the researcher is following protocol. The observer will not be taking notes about the student. The school psychology graduate student may also grade a few of the student’s writing samples in order to confirm that the researcher has scored CWS correctly. The student’s name will be removed from any writing samples that the school psychology graduate student will be scoring. In this way, the confidentiality of any student participating in this study will be maintained.

**Compensation:**
Participation or non-participation will have no effect on the student’s grade in the classroom. Compensation for participation will only be provided in the form of reinforcement for earning points for meeting writing goals.

**Voluntary Participation:**
Participation in this research study is entirely voluntary. The student and/or the parent/guardian may choose to not participate or may choose to stop participating at any time without penalty.

**Contact Information:**
If you have any questions or concerns about this research, you may contact the researcher, Katherine Lavik, at 440-667-8534, or her dissertation advisor, Dr. Frank Sansosti, at 330-672-0059. This project has been approved by the Kent State University Institutional Review Board (IRB). If you have any questions or complaints, you may call the IRB at 330-672-2704.

**Consent Statement and Signature:**
I have read the consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to grant permission for myself/my child to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

____________________________________   _____________________________
Participant Signature     Parent Signature
APPENDIX E

TREATMENT INTEGRITY CHECKLIST
Appendix E

Treatment Integrity Checklist

Date: __________________________

<table>
<thead>
<tr>
<th>Step</th>
<th>Check</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part 1: Review Goal</strong></td>
<td></td>
</tr>
<tr>
<td>1. Looked at goal setting chart with student.</td>
<td></td>
</tr>
<tr>
<td>2. Talked about goal for the present session (e.g., “What is your goal for this session?”) and reviewed current points.</td>
<td></td>
</tr>
<tr>
<td><strong>Part 2: AIMSweb Administration</strong></td>
<td></td>
</tr>
<tr>
<td>3. Administer AIMSweb probe following AIMSweb protocol.</td>
<td></td>
</tr>
<tr>
<td><strong>Part 3: Scoring, Praise, Correction</strong></td>
<td></td>
</tr>
<tr>
<td>4. Score CWS with the student.</td>
<td></td>
</tr>
<tr>
<td>5. Verbally praised student for CWS in each sentence (e.g., “You wrote [#] correct writing sequences in this sentence. Good job!”).</td>
<td></td>
</tr>
<tr>
<td>6. Told student number of IWS in each sentence (e.g., “You wrote [#] sequences incorrectly.”).</td>
<td></td>
</tr>
<tr>
<td>7. Asked student if they knew why the sequences were incorrect (e.g., “Do you know why they are incorrect?”).</td>
<td></td>
</tr>
<tr>
<td>8. Praised student for any correct answers and offered explanations for any incorrect answers or any unknown answers.</td>
<td></td>
</tr>
<tr>
<td><strong>Part 4: Goal Setting</strong></td>
<td></td>
</tr>
<tr>
<td>9. Plotted present session’s CWS score on the goal setting chart.</td>
<td></td>
</tr>
<tr>
<td>10. Recorded goal for next session with a line on the chart.</td>
<td></td>
</tr>
<tr>
<td>11. If the student has met 2 of 3 consecutive goals, set goal for next session at least one CWS higher than most recent goal. Keep in mind progress to end goal. -OR- If the student has met 1 of 3 consecutive goals, maintain current goal. -OR- If the student has met 0 of 3 consecutive goals, set goal for next session as highest goal previously attained.</td>
<td></td>
</tr>
<tr>
<td><strong>Part 5: Reinforcement</strong></td>
<td></td>
</tr>
<tr>
<td>12. If the student met or exceeded goal, add one point to the “points earned” column.</td>
<td></td>
</tr>
<tr>
<td>13. Added up the total points earned and reminded student of how to earn points.</td>
<td></td>
</tr>
<tr>
<td>14. If the student has earned 0-2 points total (or since the most recent receipt of a reinforcement), the student may not receive a prize. -OR- If the student has earned 3 or 9 points total (which represents 3 points from the beginning and 3 points from the receipt of a large prize), the student may choose a small prize. -OR- If the student has earned 6 or 12 points total, the student may choose a large prize.</td>
<td></td>
</tr>
</tbody>
</table>

Reinforcement during the final session:
If, during the final two intervention sessions the student has earned 1 point on only 1 day, the student may choose a small prize.
-OR-
If, during the final two intervention sessions the student has earned 1 point on both days, the student may choose a large prize.
### Part 6: Generalization

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>15.</strong> Researcher says, “Remember, you can use what you have learned here in other places. Where else have you been able to use what we learn here?”</td>
<td></td>
</tr>
</tbody>
</table>
| **16.** Researcher verbally praises any answers involving writing.  
- OR-  
  *If* student does not provide an answer, researcher says, “I think you can tell me at least one place. Here’s a hint—we’re in this place right now.” The researcher then praises any answer involving writing. |   |
| **17.** The researcher provides *at least 1* other suggestion for places where the student can use what s/he has learned. |   |
APPENDIX F

SAMPLE NEGATIVE/CORRECTIVE PERFORMANCE FEEDBACK
Appendix F

Sample Negative/Corrective Performance Feedback

General Procedures:

1. Before correcting any error, ask the student if he or she knows why the incorrect writing sequence is incorrect.
2. Praise (and restate/rephrase/summarize) any correct answers. Explain any incorrect or unknown answers.

Sample Conversation 1:

Examiner: Do you know why this writing sequence was incorrect?
Student: Yes, I see it now.
Examiner: What was the error you made?
Student: I spelled that word wrong.
Examiner: No, you spelled that word correctly. You forgot to put a period here to end your sentence. [onto next sequence]

Sample Conversation 2:

Examiner: Why this writing sequence was incorrect?
Student: I spelled canteen wrong.
Examiner: Yes, you’re right. You spelled canteen wrong. When figuring out words I don’t know how to spell, I like to pretend that I’m in first grade again and learning phonics. Canteen is spelled c-a-n-t-e-e-n [pronounces phonetically and writes out word while spelling it]. [onto next sequence]

Sample Conversation 3:

Examiner: Why this writing sequence was incorrect?
Student: I don’t know.
Examiner: You used the wrong verb tense of “to walk.” You wrote, “We walks to the store.” You are writing in plural first person. We and I are first person, so this should be the same as if you were saying, “I walk to the store.” “Walks” is for third person, like “he walks, she walks.” So, the correction for this sentence should be, “We walk to the store.” [onto next sequence]
APPENDIX G

GOAL SETTING CHART
Appendix G

Goal Setting Chart

Student Name: ______________________________

**Note:** Record CWS score in black ink. Record goal line in colored ink.

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Points Earned</th>
<th>Total Points</th>
<th>Day 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 2</td>
<td></td>
<td></td>
<td>Day 11</td>
</tr>
<tr>
<td>Day 3</td>
<td></td>
<td></td>
<td>Day 12</td>
</tr>
<tr>
<td>Day 4</td>
<td></td>
<td></td>
<td>Day 13</td>
</tr>
<tr>
<td>Day 5</td>
<td></td>
<td></td>
<td>Day 14</td>
</tr>
<tr>
<td>Day 6</td>
<td></td>
<td></td>
<td>Day 15</td>
</tr>
<tr>
<td>Day 7</td>
<td></td>
<td></td>
<td>Day 16</td>
</tr>
<tr>
<td>Day 8</td>
<td></td>
<td></td>
<td>Day 17</td>
</tr>
<tr>
<td>Day 9</td>
<td></td>
<td></td>
<td>Day 18</td>
</tr>
</tbody>
</table>
APPENDIX H

NOTES ON GOAL SETTING INTERACTION
Appendix H

Notes on Goal Setting Interaction

Overall, the participants were engaged with the researcher when setting goals. The researcher would have each participant mark his current level of performance and look at his previous performance. If the participant had met his goal in two of the previous three trials (or two trials in a row), the researcher would suggest raising the goal and ask the participant what he thought would be a good goal to set for next time, reminding the participant that the end goal was to get the student to the average level of performance by the end of the year (i.e., 56 CWS). The participants usually set their goals at least two points higher than their previous goals. Sometimes, the participants would jump 10 or 15 points in their goal setting, which would have been unrealistic for any of the participants. The researcher would remind the participant that she wanted him to be successful and to be able to earn more points (since most of the students were slow to start earning points). The participant would agree. Sometimes the participant would offer a more realistic goal (e.g., about 5 points higher), and sometimes the researcher would offer a different goal if the participant did not offer one. One student, Tyler, had the lowest CWS winter benchmark score but hit the average range during the first intervention session, so his goal setting became a little different from the others. His end-of-intervention goal was to simply continue his upward momentum from average to high average. During the intervention sessions, the researcher would continually praise him for his average performance (because average is not a bad place to be, especially for someone who had the lowest CWS winter benchmark score) and would discuss with him how to improve to the high average range. When a participant seemed to stall (e.g., Anderson, Kyrie, CJ), the researcher would encourage them to continue to strive for the same goal if he had met it in at least one of the three previous sessions. While some of these participants may have stalled or slipped backwards in their performances at some points, they never appeared discouraged. They maintained positive attitudes and told the researcher that they wanted to continue to strive for the goals that they had set.
APPENDIX I

STUDENT-IDENTIFIED REINFORCERS
Appendix I

Student-Identified Reinforcers

Small reinforcers (not to exceed $1.00 per item):
- Mechanical pencils
- Starburst
- Anything chocolate
- Erasers
- Small toys (e.g., bouncy ball, fake lizard to scare friends/teachers)

Large reinforcers (not to exceed $5.00):
- Subway gift card
- Game Stop gift card
- Gift card to download music (e.g., Amazon, iTunes)
- McDonald’s gift card
- TGIF gift card
APPENDIX J

SOCIAL VALIDITY SCALE FOR TEACHERS
Appendix J

Social Validity Scale for Teachers

Intervention Rating Profile – 15

The purpose of this questionnaire is to obtain information about the writing intervention that was provided to your students (or student). Please circle the number that best describes your agreement or disagreement with each statement using the scale below.

1 = strongly disagree  2 = disagree  3 = slightly disagree  4 = slightly disagree  5 = agree  6 = strongly agree

1. This would be an acceptable intervention for students who have writing problems. 1 2 3 4 5 6
2. Most teachers would find this intervention appropriate for students who have writing problems. 1 2 3 4 5 6
3. This intervention proved effective in changing my students' writing problems. 1 2 3 4 5 6
4. I would suggest the use of this intervention to other teachers. 1 2 3 4 5 6
5. My students' writing problems were severe enough to warrant the use of this intervention. 1 2 3 4 5 6
6. Most teachers would find this intervention suitable for students who have writing problems. 1 2 3 4 5 6
7. I would be willing to use this intervention in the classroom setting. 1 2 3 4 5 6
8. This intervention did not result in negative side effects for my students. 1 2 3 4 5 6
9. This intervention would be appropriate for a variety of students. 1 2 3 4 5 6
10. This intervention is consistent with those I have used in classroom settings. 1 2 3 4 5 6
11. The intervention was a fair way to handle my students' writing problems. 1 2 3 4 5 6
12. This intervention is reasonable for improving students' writing problems. 1 2 3 4 5 6
13. I liked the procedures used in this intervention. 1 2 3 4 5 6
14. This intervention was a good way to handle my students' writing problems. 1 2 3 4 5 6
15. Overall, this intervention was beneficial for my students. 1 2 3 4 5 6

APPENDIX K

SOCIAL VALIDITY SCALE FOR STUDENTS
APPENDIX K

Social Validity Scale for Students

Intervention Rating Profile – 15

The purpose of this questionnaire is to obtain information about the writing intervention in which you participated. Please circle the number that best describes your agreement or disagreement with each statement using the scale below.

1 = strongly disagree  2 = disagree  3 = slightly disagree  4 = slightly disagree  5 = agree  6 = strongly agree

1. This would be an acceptable intervention for students who have writing problems. 1 2 3 4 5 6
2. Most students would think this intervention is appropriate for students who have writing problems. 1 2 3 4 5 6
3. This intervention was effective in improving my writing. 1 2 3 4 5 6
4. I would tell other students to use this intervention. 1 2 3 4 5 6
5. My writing problems were bad enough to necessitate the use of this intervention. 1 2 3 4 5 6
6. Most students would think this intervention is suitable for students who have writing problems. 1 2 3 4 5 6
7. I would be willing to use this intervention again. 1 2 3 4 5 6
8. This intervention did not result in negative side effects for my writing. 1 2 3 4 5 6
9. This intervention would be appropriate for a variety of students. 1 2 3 4 5 6
10. This intervention is consistent with what I am learning in class. 1 2 3 4 5 6
11. The intervention was a fair way to handle my writing problems. 1 2 3 4 5 6
12. This intervention is reasonable for improving students’ writing problems. 1 2 3 4 5 6
13. I liked the procedures used in this intervention. 1 2 3 4 5 6
14. This intervention was a good way to improve my writing. 1 2 3 4 5 6
15. Overall, this intervention was beneficial for me. 1 2 3 4 5 6

REFERENCES
References


Graham (Eds.), *Handbook of research on learning disabilities* (pp. 383-402). New York: Guilford Press.


