

The Effects of an Interdependent Group Contingency on Student Writing

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Madeline Birch Auge

August 2021

Copyright © 2021 by Madeline Birch Auge
All rights reserved.

ACKNOWLEDGEMENTS

First, I want to thank my chair, Dr. Merilee McCurdy, for her mentorship, expertise, and endless support throughout my graduate career. Next, I want to thank my committee members, Dr. Chris Skinner, Dr. David Cihak, and Dr. Tara Moore, for providing their time, support, and feedback through the dissertation process. I also want to thank the research team who helped me collect, score, and input data: Brooke Killion, Jada White, Jessica Kotik, Lynnette Neu, and Valerie Hogan-Sandi. Another person who I want to acknowledge is my colleague and friend, Janet Schwartz-Micheaux. Her endless encouragement and support was instrumental in helping me complete my dissertation during the pandemic.

I want to also thank my parents, Craig and Elizabeth, and sister, Meredith, for their unconditional love and support through all of my endeavors. They, along with all the other members of my family, always remind me to celebrate each success along the way. They have also been great at challenging me and giving me advice. One of the most salient pieces of advice I have used through my life, including for my dissertation, is from my dad: “You haven’t written until you’ve rewritten.”

ABSTRACT

National data indicate that many students in the United States are not proficient in writing at grade-level expectations (Persky et al., 2003). However, there is not enough research, resources, or support for school personnel to improve student writing (Graham & Harris, 2003). Previous writing intervention studies involving performance feedback methods have shown positive impacts on student writing fluency, but it may be too time consuming for teachers to use in the classroom (Truckenmiller et al., 2014). Teachers need feasible, evidence-based writing interventions that are easy to implement in their classrooms and that motivate students to improve their writing skills.

The present dissertation study examines the effects of a classwide interdependent group contingency writing intervention using randomized criteria on student writing production. Participants included 39 students from three first and second grade classrooms at a rural elementary school. A single-case A-B-A design was used to analyze the impact of the intervention on student total words written, the maintenance of student writing production after the intervention was removed, and social validity among students and teachers.

Results indicated that students wrote more with the group contingency intervention in place, but experimental control was not established within classrooms due to time constraints. Percent nonoverlapping data and Hedges' *g* effect sizes were calculated and yielded minimal to significant effects of the group contingency on student writing production. Students and teachers responded positively to the intervention, with

teachers reporting the ease of implementation. Limitations and applications of these findings are discussed.

TABLE OF CONTENTS

Chapter I Introduction and Literature Review	1
Student Writing	2
Writing Theories	3
Instructional Hierarchy	7
Writing Intervention Research in the Classroom	9
Group Contingency	11
Limitations and Gaps in Literature	17
Present Study	18
Research Questions	18
Chapter II Method	19
Participants and Setting	20
Materials	20
Dependent Variable	21
Design	23
Procedures	24
Procedural Integrity	27
Social Validity	27
Chapter III Results	29
Class A	30
Class B	32
Class C	33
Social Validity	35
Interscorer Reliability and Procedural Integrity	36
Chapter IV Discussion	37
Limitations and Future Research	45
Other Future Research	47
Applied Implications	47
Conclusion	49
List of References	50
Appendices	59
Appendix A. Tables and Figures	60
Appendix B. Teacher Informed Consent Form	68
Appendix C. Parent Consent/Child Assent Form	71
Appendix D. Essay Example Page	74
Appendix E. Story Starter Prompts	75
Appendix F. Scripts for Baseline and Intervention Phases	76
Appendix G. Procedural Integrity Checklist Baseline	77
Appendix H. Procedural Integrity Checklist Intervention	78
Appendix I. Teacher Social Validity Measure	79
Appendix J. Student Social Validity Measure	81
Vita	83

LIST OF TABLES

Table 1. <i>Descriptive Statistics of Total Words Written for Baselines and Intervention Phases</i>	60
Table 2. <i>Percentage Non-Overlapping Data (PND) for each adjacent phase</i>	61
Table 3. <i>Hedges' g Calculations</i>	62
Table 4. <i>Student Responses to Social Validity Survey</i>	63
Table 5. <i>Teacher Responses to Social Validity Survey</i>	64

LIST OF FIGURES

<i>Figure 1. Average total words written by Class A during baseline and intervention phases.</i>	65
<i>Figure 2. Average total words written by Class B during baseline and intervention phases.</i>	66
<i>Figure 3. Total words written by Class C during baseline and intervention phases.</i>	67

CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

Student Writing

Writing is an essential skill that is used in various aspects of school, employment, and everyday communication. Successful writers can access better jobs, pursue higher education, and advance in their professions (Graham & Perin, 2007). Individuals who do not have the skills to write well are at a disadvantage compared to those who have acquired these complex skills. In fact, students who have deficits in writing perform significantly lower across subject areas compared to peers who have the skills to write well (Bangert-Drowns et al., 2004). National data indicate that many students in the United States (U.S.) are not proficient in writing at grade-level expectations (Persky et al., 2003). Only 27% of eighth grade students and 28% of fourth grade students in the U.S. perform at or above a proficient level of writing (National Center for Education Statistics, 2011; National Center for Education Statistics, 2002).

Writing development begins with a “learning-to-write” stage, where students learn how to carry out basic writing skills. Students then progress into the “writing-to-learn” stage, where they rely on higher-level cognitive functioning skills to compose written products (Bangert-Drowns et al., 2004). If students are unable to perform basic writing tasks, such as handwriting, grammar, and spelling, then they are likely to experience difficulty as they progress through school and are required to write more advanced essays (National Center for Education Statistics, 2011; Graham et al., 1991).

Given the complexity in learning to write, it is not surprising that students with writing concerns are commonly referred to school psychologists (Bramlett et al., 2002). Students who are not proficient in writing typically struggle with various aspects of the

writing process, such as planning, transcribing, and editing (Troia & Graham, 2003). Students with a specific learning disorder (SLD) in writing have difficulty translating their thoughts onto paper in an organized manner and spend more time and effort trying to carry out basic writing skills, such as spelling and grammar. Subsequently, they have more difficulty performing higher-order levels of writing, which includes the planning and organization processes, compared to typical peers (Graham et al., 1991). For these reasons, students with deficits in writing are at a disadvantage compared to typical peers when writing for high-stakes positions, such as college admissions or job applications.

Despite the fact that a majority of students do not write at grade-level expectations, there are not enough research-based writing interventions available to remediate students' writing deficits (Eckert et al., 2009). Additionally, there is not enough research, resources, or support for school personnel to improve student writing (Graham & Harris, 2003). National organizations, such as the National Writing Project (2019), have provided professional development opportunities for educators and practitioners to emphasize the need for evidence-based practices in writing instruction. To develop evidence-based practices in writing, there needs to be an understanding that writing is a complex skill that develops over time (Harris & Graham, 2016). Several theories have explained the complexity of learning to write and outline the process of writing development.

Writing Theories

Several theories have attempted to explain the intricate writing processes and cognitive functions involved in the act of writing. Writing is a complex developmental

process that may be difficult to explain through theories (MacArthur et al., 2015). The Hayes and Flower model (1980) is a broad framework that was the first to describe the cognitive processes involved in writing. While there are modern theories based on current research, the Hayes and Flower model serves as the backbone of the theories that followed (Alamargot & Chanquoy, 2001). From this model, the Berninger and Swanson model of writing (1994) emerged as a prominent model that is better used to describe the development of writing in early learners.

Hayes and Flower Model

The Hayes and Flower (1980) model of writing is the first published cognitive framework explaining the process of writing. This model explains the cognitive and environmental factors that contribute to the process of writing, as well as, individual differences in writing construction (Alamargot & Chanquoy, 2001; Hayes & Flower, 1980). There are three main components of the Hayes and Flower model: task environment, long-term memory, and monitoring. Task environment refers to the factors outside of the writer's control, such as the topic of the writing assignment (Alamargot & Chanquoy, 2001). If the topic of the writing assignment is interesting to the writer, then he is more likely to put forth more sustained attention and effort toward the task. Whereas, if a writer finds the assignment boring, he may have decreased motivation to complete the assignment in an effortful manner. Students' motivation to write can be directly targeted by modifying environmental variables. For example, a teacher can allow students to choose a writing topic instead of the teacher assigning a topic (Britton et al., 1975).

The second component of the Hayes and Flower (1980) model is the writer's long-term memory. Long-term memory incorporates three different types of knowledge to apply to the writing task: domain, pragmatic, and linguistic knowledge. Domain, or general, knowledge refers to what the writer knows about the topic of the assignment. Pragmatic, or communication, knowledge refers to what the writer knows about the audience, such as a teacher reading a student's response. Finally, linguistic, or grammatical, knowledge refers to what the writer knows about the formatting of the writing task. According to the Hayes and Flower theory, all three types of knowledge are necessary for an individual to possess in order to progress to the next phase of physical writing.

The monitoring process, or the physical process of writing, involves using long-term memory to incorporate background knowledge into the writing task. This process transfers domain (topical) knowledge into linguistic (grammatical) knowledge (Alamargot & Chanquoy, 2001). While this process occurs, the task environment factors continually influence the three main monitoring components: planning, translating, and reviewing. If the writer does not have prior experience, or linguistic knowledge, of the writing task, he can still plan a successful written product through the retrieval of various pieces of domain knowledge. Then, the writer organizes the information in a way that achieves the overall goal of the writing task (i.e., answering the writing prompt accurately). Next, in the monitoring process, the writer translates the plans onto the paper through use of pragmatic and linguistic knowledge. Finally, the writer follows the reviewing process, including reviewing and editing his work (Hayes & Flower, 1980).

Berninger and Swanson Model

The Berninger and Swanson model (1994) was developed to address the specific cognitive processes that students undergo when learning how to become proficient writers. This model is considered one of the most comprehensive writing models that explains writing development (Alamargot & Chanquoy, 2001). Berninger and Swanson (1994) used the Hayes and Flower (1980) model to lay the groundwork, but then extended this model further by attempting to explain how each process develops from a cognitive perspective. Berninger and Swanson's model is more appropriate to explain student writing development because it takes into account the varying developmental stages of young writers. For example, elementary-aged students do not have as strong of a working memory capability as middle school-aged students. Hence, younger students would not be able to store as much information into long term memory as older students. Due to these developmental considerations, Berninger and Swanson (1994) explain that it is important to know a student's level of current writing ability to understand what stage of writing skill development they are in and what they still need to learn to become successful writers.

Although the Hayes and Flower (1980) model and the Berninger and Swanson (1994) model explain the cognitive theory behind writing development, another theory is needed to explain how writing development can be applied to teaching and learning. The Instructional Hierarchy model describes the process in which students effectively learn in the classroom. This model explains the development of academic skills, including writing, in an ecologically valid setting, such as in the schools (Haring et al., 1978).

Instructional Hierarchy

The Instructional Hierarchy (IH) pairs instructional stages and instructional activities to promote student learning. The IH describes how students learn based on the four common stages of learning: acquisition, fluency, generalization, and adaptation (Haring et al., 1978). While all IH stages are important in student learning, the current research project focuses on student acquisition and fluent use of writing skills. Therefore, this summary will focus on the first two stages of the Instructional Hierarchy. The acquisition stage is the beginning stage focused on skill accuracy that requires teacher modeling of the skill with feedback to the student. At this stage, the student cannot yet perform the task on his own, so he relies on a model to assist him. When the student attempts the task, he requires significant support from the model. Once the student can perform the task accurately on his own or with little support, he moves to the second stage of the IH. In the fluency stage, the student performs the task independently, accurately, and practices the task to mastery. The teacher continues to provide feedback to the student until he masters the task by performing it correctly several times (Haring et al., 1978).

The IH provides useful guidelines that can help teachers and practitioners identify a student's learning stage and the instructional strategies that will benefit that student. With this stage-based information, researchers have developed specific interventions (i.e., repeated readings for reading fluency) that can apply directly to the student's learning stage and help the student in a particular stage (Martens & Eckert, 2006; McCurdy et al., 2006). For the acquisition and fluency stages, instruction should rely on modeling/direct

instruction, practice, performance feedback, and rewards for effort, accuracy, or speed, depending on the learning stage.

Direct instruction uses explicit and deliberate instructional strategies to deliver new material to students (Kinder & Carnine, 1990). Direct instruction is a process that involves breaking material down into segments that can be easily taught to students (acquisition) and allowing students to practice using the newly learned material (fluency) in structured contexts developed by the teacher. Students are given opportunities to attempt the new skill by practicing among peers in a structured manner so that they can become fluent in the skill (Kinder & Carnine, 1990).

Opportunities to respond, or practice, involves students repeating a newly acquired skill until they can fluently (i.e., accurately and quickly) perform that skill. Practice in the classroom has shown numerous benefits on student learning, with the most notable benefit being an increase in student engagement and motivation (Menzies et al., 2017). From a cognitive perspective, repetition allows information to be encoded from short-term into long-term memory.

In addition to instruction and practice, performance feedback methods are used by teachers in two ways. In the acquisition stage, performance feedback can be used as an error correction technique to prompt students to accurately complete a skill (Visscher & Coe, 2002). During this stage, the student relies on the teacher as a model to help him correct mistakes and complete the assigned task. During the fluency stage, the student works to independently complete a task with accuracy and speed, and performance feedback is delivered as a motivator. Therefore, performance feedback may be based on

accuracy of the skill or on the amount of time required to accurately complete the skill. Performance feedback can serve as reinforcement for students and motivate them to complete similar tasks in the future (Menzies et al., 2017). For example, performance feedback in the form of a chart or graph can motivate students to improve their performance. Additionally, performance feedback in the form of graphic representation can be encouraging to students.

The cognitive models of writing development and the IH framework can provide guidance for educators. A survey of writing educators indicated that teachers reported implementing various writing strategies reflected in the IH. However, teachers report being unsure of how to balance different types of writing instruction to benefit their students (Cutler & Graham, 2008). Additionally, teachers vary significantly on how much time they spend teaching writing to students. Only 65% of 8th grade students reported spending at least 15 minutes writing per day in English class (National Center for Education Statistics, 2011). For students to improve their writing skills, there needs to be more support for teachers, including providing teachers with supplemental instructional strategies easily incorporated into the classroom setting (National Commission on Writing, 2003).

Writing Intervention Research in the Classroom

Despite the importance and the need to develop writing skills, there has been minimal research examining writing interventions compared to other academic areas, such as reading and math (Eckert et al., 2009). Teachers need classroom writing strategies that reflect the instructional hierarchy and include modeling, practice,

feedback, and rewards to encourage skill acquisition and fluency. Teachers need effective, yet feasible, intervention techniques to address writing deficits within their classrooms. One particular writing intervention method that has shown to be effective in the classroom setting is performance feedback (Truckenmiller et al., 2014).

Performance Feedback

Performance feedback is one of the most researched writing interventions in recent years. Studies have found that performance feedback interventions can significantly improve student writing production (Truckenmiller et al., 2014). For example, third grade students created bar graphs based on the total number of words they wrote in order to visually represent their progress. Next, researchers added a goal-setting condition to determine whether having the students set their own goals on their graphs would enhance writing performance more than performance feedback alone (Koenig et al., 2016). Results found that student writing in the performance feedback and the performance feedback plus goal setting condition were significantly higher than the control. However, there were no significant differences between the performance feedback and the performance feedback plus goal setting conditions. This result indicates that the performance feedback on its own has an effect on writing production. These findings, as well as findings from similar studies, support the use of performance feedback to increase writing production among elementary-aged students (Truckenmiller et al., 2014). However, most performance feedback studies were conducted by researchers, so the studies have little ecological validity in the classroom setting (Hier & Eckert, 2014).

Although performance feedback has shown positive impacts on student writing fluency, it may be difficult for teachers to use in the classroom. First, this method can be time-consuming for the teacher when providing detailed, individualized feedback to each student on writing production and writing quality (Truckenmiller et al., 2014). Consider a teacher who spends 15 minutes providing feedback on one student's narrative story. For a class of 20 students, at this rate, it would require that teacher to spend 5 hours providing feedback on a single classwide writing assignment. In addition, feedback is more effective if it is delivered as soon as possible. Based on basic behavioral principles, the shorter the interval of time between performing an action and receiving a reward, the more likely the behavior will occur in the future (Skinner, 1963). Unfortunately, prompt feedback on a written product may be more difficult for a classroom teacher due to time and resource constraints. Classwide interventions, such as group contingencies, can be implemented by the teacher using methods that allow the teacher to deliver the most feedback in the shortest amount of time.

Group Contingency

Group contingency (GC) is an intervention that has been used to improve emotional, behavioral, and academic skills of students. GC aims to address a general problem among a group of students by increasing a desired task, such as homework completion (Little et al., 2015). Essentially, students work as a group to achieve a target goal in order to receive a reward. Several GC studies have found behavioral improvements among students, such as increased prosocial behaviors and decreased negative statements (Hansen & Lignugaris/Kraft, 2005). GC has been used in various

settings, such as special education classrooms and residential facilities, as well as with students who have various presenting concerns, such as Emotional Disturbances and Specific Learning Disorders (Little et al., 2015; Popkin & Skinner, 2003).

Types of GC

There are three types of GC methods: dependent, independent, and interdependent (Litow & Pumroy, 1975). Each method has benefits and consequences that should be considered before selecting which method is appropriate for a given scenario. Dependent GC requires that one person or subgroup must meet the goal for other members of the group to receive the reward. Dependent GC can be detrimental if the group relies on one individual to reach a particular goal for the group. If the goal is for an entire group to improve on a task, then the dependent GC method is not ideal. On the other hand, dependent GC can allow for individuals to access a reward that they may not obtain on their own. In contrast, an independent GC allows only the individuals who achieve a goal to have access to a reward. Consequently, those who do not meet the goal are unable to earn a reward. One negative result of independent GC is that individuals are likely to give up trying to meet a goal if they are consistently underperforming. Finally, interdependent GC is an all-or-none approach, where either everyone in a group receives a reward, or no one does. This method is generally perceived as the preferred approach, especially when using it with children due to the “fairness” quality of each student receiving the reward rather than some students being left out (Litow & Pumroy, 1975; Romeo, 1998).

Interdependent GC

With the interdependent GC method, there are three types of group performance levels that can be used to measure overall performance: the group as a whole, one random performance, and averaged performances. When using the group as a whole, every individual in the group needs to meet a criterion, or performance goal, to earn a reward. This method can be beneficial because all students are held accountable. However, this method requires that the teacher evaluate every student's performance which is time consuming. In contrast, the second method evaluates one individual's work which serves as the class' performance level. This method is less time-consuming, so the group can receive feedback as quickly as possible. However, a single random performance is not an accurate representation of the entire group and may lead to negative social repercussions if the group does not receive the reward (Romeo, 1998). For example, a lower-performing student's math homework might be selected to serve as the class performance level. If the chosen student does not meet the goal, peers may isolate that student or say negative comments to them (Romeo, 1998). Finally, the averaged performances method allows for an accurate representation of the group, as well as provides more accountability among students. The averaging method adds up all the performances and divides it by the number of individuals. If the average performance is greater than the criterion, then the entire group earns the reward (Litow & Pumroy, 1975).

Interdependent GC Research. Several studies in the academic areas of reading, spelling, and math have looked at the effectiveness of an interdependent GC in the classroom setting (Scott et al., 2017; Shapiro & Goldberg, 1990; Sharp & Skinner, 2004;

Turco & Elliott, 1990). Recently, Scott et al. (2017) evaluated the effects of an interdependent GC on math problem accuracy in a first-grade classroom. They found that the class' average math accuracy score was significantly greater than scores at baseline. Additionally, some studies have shown that students who perform lower at baseline benefit the most from an interdependent GC method. For example, Pappas, Skinner, and Skinner (2010) administered an interdependent GC in an Accelerated Reader program with fourth-grade students. Results showed that lower-performing students' reading accuracy increased significantly whereas the higher-performing students' reading accuracy remained the same. Although there is support that interdependent GC can be used to increase academic performance, only one study has evaluated the effects of an interdependent GC on writing production.

McCurdy, Skinner, Watson, and Shriver (2008) targeted 17 ninth-grade students with disabilities in writing to improve their use of adjectives, full sentences, and conjunctions. Using a multiple baseline design across classrooms, the three writing skills were introduced sequentially. A new skill was presented following mastery of the previous skill. For each writing skill, students were asked to try to meet a classroom criterion goal. All student essays were scored by a researcher, who wrote their performances privately on each essay. The class averages for each essay were displayed on a graph in the classroom. Therefore, all students were able to compare their own performances with the class' overall performance. Results indicated that the interdependent GC intervention improved student writing across all skill areas, but only one writing quality maintained when the intervention was withdrawn. McCurdy et al.

(2008) discussed that future studies should focus on the feasibility of implementing a GC intervention in the classroom. For example, in the study, a researcher directed all aspects of the intervention; whereas, a teacher would not have the time or resources available to attend to each students' essay. Additionally, McCurdy et al. (2008) recommend that future researchers use randomized criterion rather than targeting one skill at a time and explicitly telling students what they will measure on a given day. Randomized criterion can provide uncertainty among students, so they are more likely to try to write using as many skills as possible.

Randomized criterion. The criterion for an interdependent GC needs to be clearly identified so that all members of the group are motivated to attempt the task. If a criterion is set too high, then lower-performing students might not put in as much effort as higher-performing students. In contrast, if the criterion is set too low, then all students might put in as little effort as possible to reach the goal. Criterion-setting can be difficult, given the variability of assignment difficulties and abilities among students. If the criterion is known prior to the task, then the students might only perform to the criterion level and then stop once they have reached that goal. Essentially, the students might perform the bare minimum requirements, which would not be a valid representation of what they are capable to do.

Research using randomized criterion with students has primarily targeted student behaviors such as increasing on-task behavior while decreasing off-task behaviors. For example, Denune et al. (2015) analyzed the effects of an interdependent group contingency on on-task, off-task, and disruptive behaviors among sixth grade students

with emotional and behavioral disorders. They established rules and awarded points for students who followed the rules. As a group, students accumulated points for following these rules, such as sitting in their seats and not displaying disruptive behaviors. The researcher randomized the criterion that served as the group's goal total points and pulled that number out of a bag. If the group's total points were higher than the goal criterion, then they were awarded with a randomized reinforcing item. Results from the Denune et al. study found that students' on-task behaviors improved with the interdependent GC while their off-task behaviors reduced.

In the area of academics, few studies have examined the effects of a GC using randomized components. First, Popkin and Skinner (2003) administered an interdependent GC using randomized criteria that were selected after the task was completed. The study included five middle school students who had emotional and behavioral disorders. The criteria possibilities included accuracy in math, English, and one randomly selected subject. Therefore, the students did not know how many tasks they had to complete correctly or which subject would be scored, so they attempted to complete as many problems as possible across subjects. Results yielded significant improvements across all subject areas, particularly spelling (Popkin & Skinner, 2003).

Few studies have examined the use of a GC with randomization components on academics with typically developing students. One of these studies evaluated the use of a GC with randomization components on homework accuracy among six fourth graders (Reinhardt et al., 2009). Three homework assignments in different academic areas, reading comprehension, mathematics, and spelling, were assigned to each student every

day. The teacher calculated the accuracies of each assignment and randomly selected which academic subject would be the targeted goal criterion as well as the percentage of accuracy on the specified assignment. Results indicated that students improved in the area of reading comprehension more than spelling and math. Reinhardt et al. (2009) also explain that student reading comprehension was the lowest scoring subject at baseline. Despite the effectiveness of the interdependent GC and randomized criterion design, no studies have examined using this method on writing in a classroom setting with typically developing students.

Limitations and Gaps in Literature

Several studies have examined the effectiveness of writing interventions using performance feedback methods that often are not feasible in a typical classroom setting (Truckenmiller et al., 2014). In many of these studies, results were not ecologically valid due to the fact that researchers implemented the interventions rather than teachers (Hier & Eckert, 2014). From a classwide perspective, group contingencies have shown positive effects in academic areas, such as reading and math, but have not been applied to writing skills (Pappas et al., 2010; Scott et al., 2017). One study has attempted to apply a group contingency intervention to student writing skills with the researcher as the primary administrator (McCurdy et al., 2008) and skill growth did not maintain. Additionally, it has been recommended that future studies use randomized criterion to motivate students to write longer and better essays (McCurdy et al., 2008). Teachers need feasible writing interventions that are easy to implement in their classrooms and that motivate students to improve their writing skills.

Present Study

Based on theories of writing (Hayes & Flower, 1980; Berninger & Swanson, 1994), the IH (Haring et al., 1978), writing intervention research (Koenig, et al., 2016; McCurdy et al., 2008; Truckenmiller et al., 2014), and GC research (Pappas et al., 2010; Scott et al., 2017), the present study examines an interdependent group contingency writing intervention using a randomized criterion method. The randomization of the criteria aims to influence students to write the most words during each essay writing session. Additionally, students counted their own words to reduce the amount of work by the teacher and to provide immediate feedback.

Research Questions

1. Does participation in a classwide interdependent group contingency increase student writing production, as measured by Total Words Written (TWW)?
2. Following the use of a classwide interdependent group contingency, will students maintain writing production (i.e., TWW)?
3. Does the interdependent group contingency have social validity for the teachers and students?

CHAPTER II

METHOD

Participants and Setting

This study took place at a rural primary school in the Southeast region of the United States. All phases of the study took place in the students' typical classrooms during a pre-designated writing time, so as not to take away from academic instruction. One 1st grade and two 2nd grade teachers agreed to participate in this study.

All 54 students participated in this study; however, only 39 students with appropriate consent and assent were included in the data collection. Of those participants, 18 students were male and 19 were female. Additionally, 18 students were in 1st grade and 19 students were in 2nd grade. There were no exclusionary criteria.

Materials

First, teachers signed a University of Tennessee Knoxville IRB-approved teacher consent form (Appendix B). Next, parent consent and youth assent materials were sent home (Appendix C). Additional study materials included essay papers, pencils, index cards for the intervention criterion and rewards, a stopwatch, and a calculator for researcher and teacher use. Each student was provided an individual essay paper with lines on the front and back to guide their handwritten responses to story starters. All participants were given the same story starter during each session with the prompt written across the first line (Appendix D).

The story starters were obtained from AIMSweb (Powell-Smith & Shinn, 2004). Prompts selected by researchers for inclusion were randomly selected across sessions. All prompts were selected for inclusion based on their grade-level appropriateness. The story starters were narrative based so students could write about themselves (i.e., "The best part

about summer is...” and “When I grow up, I want to be...” and not need to rely on past, learned information. In addition, past research results support the use of narrative prompts as a valid measure of student writing abilities across grades (McMaster & Campbell, 2008). The list of prompts is included in Appendix E.

To implement the intervention with integrity, other study materials for the researchers and teachers included: researcher-developed scripts for baseline and intervention phases (Appendix F) and procedural integrity checklists (Appendices G and H). At the end of the study, social validity questionnaires were dispersed to teachers and students (Appendices I and J). These questionnaires were adapted from a Likert-type scale developed by Kear et al. (2000).

Dependent Variable

In the classroom, teachers typically use writing rubrics to evaluate student writing quality, which can be time-consuming. Teachers need simple strategies to measure student writing production that can be used within a classroom setting. Additionally, researchers have acknowledged the difficulty in developing a standardized writing measure because of the various complex skills that are involved in the writing process (McMaster & Campbell, 2008). However, researchers have found that production measures are reliable measures of writing quality in young students. Some researchers have found that the use of quality-related measures, such as Correct Writing Sequences, has been shown as the primary valid indicator of student writing skills (Gansle et al., 2002). However, this method is too time-consuming and confusing for teachers to implement in the classroom. In writing research, writing is commonly scored using

production-based curriculum-based measurement (CBM) scoring techniques, such as total words written (TWW; Wright, 2013). Student TWW was used to measure writing quantity and will serve as the primary dependent variable.

Total Words Written (TWW)

TWW is a CBM used to measure student writing quantity or production. TWW is measured by counting the number of “words” written, regardless of whether they are spelled correctly. A “word” is defined as a group of letters that have a clear space between another group of letters (Wright, 2013). Since words do not have to be spelled correctly to count TWW, this is an easy way for students to assess their own performance. In this study, students counted their own TWW after completing each essay to reduce teacher workload. The researchers then rescored TWW for the student papers and reported the data for each session as class averages of students who had consent to participate.

Inter-Scorer Agreement

Prior to the start of the study, supporting researchers were trained by the lead researcher to score TWW. Supporting researchers were trained to score student essays with reliability above 90% compared with the trained primary researcher. Then, during baseline and intervention phases, inter-scorer agreement was calculated across 17% of sessions to evaluate TWW. Interscorer agreement was calculated by dividing the higher scorer’s TWW by the other scorer’s TWW and multiplying by 100. The minimum acceptable agreement was 90%. If agreement was below 90%, the researchers scored the

essays again and discussed their scores with each other until they reached an agreement. However, in the current study, all inter-scorer agreement was above 90%.

Design

An A-B-A single-subject withdrawal design was used, and all data were entered into a secure database. With this design, experimental control is established when the baseline data are relatively stable (i.e., low variability). Then, the subsequent intervention phase is implemented and continues until the trend is stable again. During the second baseline phase, the intervention is removed to evaluate the impact of withdrawing the intervention (Gast & Baekey, 2014). These trends are determined by visual analysis, examining level, trend, and variability in the data. In the current study, students in all classrooms wrote five essays at baseline. The intervention phase length differed among the classrooms based on the trends of their average TWW. In addition, effect sizes, using percent non-overlapping data (PND), and Hedges' *g*, were calculated.

PND is a common method used to visually analyze the effectiveness of a single-subject design. It is widely recognized as a valid measure of treatment effect among researchers (Olive & Franco, 2008; Scruggs et al., 1987). When conducting PND, a horizontal line is drawn at the level of the highest data point during a phase (i.e., baseline or treatment phase). In the following phase, the number of data points that are above the horizontal line are counted and divided by the total number of points in the phase. This proportion is then multiplied by 100% in order to attain a percentage. A treatment is considered very effective if PND is greater than 90%; effective if PND is between 70% and 90%; questionable if PND is between 50% and 70%; and ineffective if PND is under

50% (Olive & Franco, 2008; Scruggs & Mastropieri, 1998). The greater the PND, the stronger the treatment effect (Scruggs & Mastropieri, 1998).

Additionally, Hedges' g was calculated to obtain further effect size measures (Hedges, 1981). Hedges' g is commonly used in single-subject research, as opposed to other effect size calculations (i.e., Cohen's d), when there are fewer than 20 participants per group. A treatment is considered to have a large effect if the value is 0.08; a moderate effect size if the value is 0.05; and a small effect if the value is 0.02 (Sawilowsky, 2009).

Procedures

Approval for the study was obtained through the IRB at the University of Tennessee Knoxville along with the vice principal of the primary school and the Director of Schools. The study was developed to minimize interruptions in the typical classroom setting. For example, the intervention was conducted during the students' dedicated writing time scheduled by their teachers. All first and second-grade teachers were approached regarding the study by the primary researcher and three teachers expressed interest in participating. The teachers each signed an informed consent form (Appendix B). Parent consent and youth assent forms were sent home with all students in the participating classrooms (Appendix C). Students and parents were informed that all students were to participate in the intervention, but the consent/assent form was to allow for inclusion in data collection and analysis.

Researcher and Teacher Training

The primary researcher trained graduate student researchers in scoring procedures and in the administration of the intervention steps. Graduate student researchers

participating in data collection were trained to score student essays using TWW scoring guidelines from AIMSweb (Wright, 2013). Each graduate researcher scored three sample student essays and were allowed to aid in data collection if their scores were at least 90% reliable. Reliability was calculated by dividing one rater's score by another rater's score of TWW, then multiplying by 100 for the final percentage of agreement.

Additionally, the primary researcher trained the teachers to implement the intervention using provided intervention scripts (Appendix F). During baseline, the primary researcher led the procedures in the classroom with the teachers observing. Then, during the first two intervention phase essays, the researchers modeled the intervention for the teachers. Once each teacher felt comfortable administering the intervention independently, the researcher observed the teacher implementation and collected procedural integrity data (Appendix H). The teachers then took over as the primary individual administering the writing intervention in their classrooms. The primary researcher observed the teachers and provided feedback during the first sessions to ensure procedural integrity. The teachers were provided with any necessary support they needed from the primary researcher. At least one graduate researcher was present with the teacher to collect implementation integrity measures (Appendices H and I). With the teachers' permission, some procedural integrity measures were collected via audio recording to allow the teachers flexibility in administering the intervention.

Baseline

The baseline phase served as the control condition. In the classroom, students were prompted to plan their essay for one minute then handwrite their story for seven

minutes. Each story starter was written on the essay page and read aloud to the students to ensure that each student understood the story prompt. Students then turned in their papers to the teacher or researcher. No feedback was provided to the students.

Intervention

Interventions took place two to three times per week in each classroom over the course of eight weeks. Once baseline data were stable, the interdependent group contingency intervention was implemented. The intervention writing instructions were the same as the baseline writing, but the students also were informed of the group contingency component (Appendix F). During the intervention, the students worked as a group to write as many words as they could in their essays. After seven minutes, the students counted the number of words they wrote, wrote that number at the top of their papers, and turned it in to the teacher. The teacher then randomly selected five student papers and averaged the TWW. The student names chosen were not disclosed to the class due to possible undesired social effects. This average TWW was considered as the class score and was announced to the students.

The teacher then selected a random card with a number on it, which served as the criterion, or class goal. The numbers on the cards were determined by the researchers and were calculated dependent on the average TWW each class wrote during baseline. These numbers ranged from 30% below to 30% above the average TWW the class wrote at baseline. This range was used to ensure a mixture of challenging and easy-to-reach goals (i.e., a mixture of difficulty levels). Since students did not know the criterion prior to writing, they were encouraged to do their best work in case the higher, more challenging

criterion was drawn. If the average TWW from the five student papers (i.e., the class score) was greater than the criterion (i.e., the class goal), then the class received a candy reward that was pre-approved by the teachers. If the class score was less than the goal, then the class did not receive a candy reward, but were encouraged to write more during the next essay. The first day of intervention was purposely altered by the researchers in order to ensure that the students accessed the reward. All following intervention sessions were not altered by researchers.

Procedural Integrity

To ensure procedural integrity, the researcher introduced the study to the students and modeled the procedure for the teacher. A second graduate researcher was present to complete a procedural integrity checklist derived from the researcher-generated script (Appendix H). Procedural integrity was calculated by taking the total number of completed checklist items and dividing it by the total possible checklist items, then multiplying by 100. The teachers then conducted the intervention on their own with the lead researcher observing and providing feedback on their performances. Once the teacher implemented the intervention with 100% integrity, and felt confident enough to do so, the graduate researcher attended and/or recorded 33.3% of the remaining sessions to collect procedural integrity using the checklist.

Social Validity

A brief survey developed by the researchers was administered to all students and the teacher during the last day of data collection (Appendices I and J). The items were

read aloud to them by the researcher to ensure that all students comprehended each item. The survey was derived from The Writing Attitude Survey (Kear et al., 2000) and consisted of age-appropriate statements, such as, “I enjoyed writing as a team to get a reward.” Students circled one of the four Likert-scale options that best describes their views of the statement: 1 “Not at all like me,” 2 “Sort of not like me,” 3 “Sort of like me,” and 4 “Definitely like me.” Each Likert-scale option was depicted by a picture of a cartoon rather than words in order to allow for students to comprehend each option and respond accordingly. Researchers scanned the room to ensure all students circle one option for each statement, but they did not look at student answers, to ensure confidentiality. Students also had an opportunity to write an open-ended response regarding what they did or did not like about the intervention.

The teachers also completed a social validity survey by rating statements such as, “I think this writing project improved my students’ writing skills,” on a five-point Likert scale (see Appendix J). The teachers circled one option that best described their views of each statement: 1 “Strongly Disagree,” 2 “Somewhat Disagree,” 3 “Neutral,” 4 “Somewhat Agree,” and 5 “Strongly Agree.” The teachers also had an opportunity to write an open-ended response regarding what they did or did not like about the study, as well as how the study could have been improved. Results from student and teacher surveys provided insight on the social acceptability and feasibility of the writing study. Social validity results can also provide future directions for designing effective writing interventions.

CHAPTER III

RESULTS

Each classroom's average performance on Total Words Written (TWW) was averaged across all students who had parental consent. Although students counted and recorded their TWW for each essay, the TWW in Figures 1-3 was recorded by the researchers. The researchers' decision making of when to introduce the intervention phase was based on a visual analysis of the data in Figures 1-3 in Appendix A. Results, described below, include visual analyses, descriptive statistics, and effect size calculations for each classroom.

Class A

Class A was a second-grade classroom with 11 students (5 females, 6 males) whose parents consented to have their data included in the study. Students wrote 14 stories in total. The class had five opportunities to earn a reward for their writing and earned a total of five rewards.

Visual Analysis

Visual analysis of TWW (Figure 1) across baseline and intervention phases found a high degree of variability in the initial baseline phase. An immediate but small level change occurred upon the introduction of the intervention phase. The first two intervention data points (6, 7) were approximately at the same level as the highest baseline data point (2). Once exposed to the intervention, Class A's writing showed moderate variability, but displayed an overall high level and increasing trend. A withdrawal phase was then implemented by removing the GC intervention and returning to baseline procedures. During the second baseline phase, Class A's average TWW remained relatively stable and at the same level as the GC intervention phase.

Descriptive Statistics and Effect Sizes

Information on the descriptive statistics for Class A's writing is included in Table

1. Class A wrote 70.9 total words at baseline and 85.9 words under GC. When returning to the second baseline phase, TWW increased to an average of 88.35 words.

The primary effect size calculation used in this study is the percentage of non-overlapping data (PND) and are shown in Table 2. Across the initial baseline and GC intervention phases, the variability in average TWW is reflected in a PND of 40%.

Across the GC intervention phase and second baseline, Class A had a minimally effective PND calculation with 0%.

Further effect size calculations were computed using Hedges' g (Hedges, 1981) and are shown in Table 3. Class A had a large positive effect on TWW from the initial baseline to intervention phases ($g = 2.24$), which indicates that the GC intervention outperformed Class A's initial TWW baseline. When examining the effect size between the intervention and the second baseline phase, a moderate positive effect size was present ($g = 0.50$).

Summary

Overall, Class A's writing production (i.e., TWW) during intervention did improve compared to the initial baseline. However, variability in the initial baseline negatively impacted experimental control. Therefore, the impact of the GC intervention is difficult to interpret.

Class B

Class B was a first-grade classroom with 18 students (9 females, 9 males) whose parents consented to have their data included in the study. Students wrote 14 stories in total. The class had six opportunities to earn a reward for their writing and earned a total of five rewards.

Visual Analysis

Visual analysis of TWW (Figure 2) across baseline and intervention phases found a moderate degree of variability in the baseline phase, as well as a slightly increasing trend. When the intervention phase was introduced, the students' TWW remained similar to the baseline level, with a slight increasing trend as the intervention progressed. The first four intervention data points (6, 7, 8, and 9) were approximately at the same level as the highest baseline data points (2, 4, and 5). On the tenth story, the average TWW increased sharply and remained above the baseline levels. A withdrawal phase was then implemented by removing the GC intervention and returning to baseline procedures. During the second baseline phase, Class B's average TWW remained relatively stable and at the same level as the GC intervention phase during the first two data points. Their average TWW during the last story then sharply decreased to the same level as the initial baseline phase.

Descriptive Statistics and Effect Sizes

Information on the descriptive statistics for Class B's writing is included in Table 1. Class B averaged 24.8 words at baseline and 32.6 words written under GC. When

returning to the second baseline phase from the intervention phase, the class' average TWW increased to an average of 36.7 words.

One effect size calculation is the percentage of non-overlapping data (PND) which are included in Table 2. Across the initial baseline and GC phases, the variability in average TWW is reflected in a PND of 50%. Across the GC intervention phase and second baseline, Class B had a PND of 33%. Experimental control could not be established due to a continuous increasing trend across the baseline and GC phases.

Further effect size calculations were calculated using Hedges' g (Hedges, 1981) and are shown in Table 3. The intervention had a large positive effect on Class B's TWW from the initial baseline to intervention phases ($g = 1.52$), which indicates that the GC intervention impacted Class B's TWW. When examining the effect size between the intervention and the second baseline phase, a large positive effect size was also present ($g = 0.71$).

Summary

Overall, Class B's writing production (i.e., TWW) during intervention did improve compared to the initial baseline. However, the GC condition did not result in clear level changes from baseline to intervention. Due to poor experimental control, the GC intervention cannot explain why students in Class B wrote more compared to their initial baseline.

Class C

Class C was a second-grade classroom with 10 students (7 females, 3 males) whose parents consented to have their data included in the study. Students wrote 14

stories in total. The class had seven opportunities to earn a reward for their writing and earned a total of six rewards.

Visual Analysis

Visual analysis of TWW (Figure 3) across baseline and intervention phases found an increasing trend in the baseline phase. An immediate, but small, level change occurred upon the introduction of the intervention phase. During the intervention phase, Class C's writing showed some variability, and a slight increasing trend. A withdrawal phase was then implemented by removing the GC intervention and returning to baseline. During the second baseline phase, Class C's average TWW initially increased above the GC intervention level. During the final story, the average TWW sharply decreased, but remained at a level greater than the initial baseline phase.

Descriptive Statistics and Effect Sizes

Information on the descriptive statistics for Class C's writing is included in Table 1. Class C wrote 39.7 words at baseline which increased to 52.4 words under GC. When returning to the second baseline from the intervention phase, the class' average TWW increased to an average of 58.1 words.

One effect size calculation is the percentage of non-overlapping data (PND) which are shown in Table 2. Across the initial baseline and GC intervention condition, the variability in average TWW is reflected in a strong PND of 100%. Across the GC intervention phase and second baseline, the PND for Class C was 50%. Experimental control could not be established due a to the increasing trend in the initial baseline.

Further effect size calculations were calculated using Hedges' g (Hedges, 1981) and are in Table 3. The GC had a large positive effect on TWW from the initial baseline to intervention phases ($g = 2.93$), which indicates that the GC intervention impacted Class C's TWW. When examining the effect size between the intervention and the second baseline phase, a moderate positive effect size was present ($g = 0.66$).

Summary

Overall, Class C's writing production (i.e., TWW) during intervention did improve compared to the initial baseline. However, the GC condition did not result in clear level changes from intervention to baseline due to an increasing trend in baseline. Therefore, the GC intervention cannot explain why students in Class C wrote more compared to their initial baseline.

Social Validity

Social validity data were taken from a survey completed by student participants after termination of the second baseline phase. Overall, students reported that they enjoyed the writing intervention, thought that they were able to write more than they did before the intervention, and enjoyed working as a group to obtain a reward (see Table 4).

Teacher social validity data were taken at the same time during the student social validity data collection. Teachers reported that they found the intervention easy to implement, thought their students enjoyed it, and thought that their students wrote more than they did before the intervention (see Table 5). All teachers wrote responses to open-ended questions. Two of the most common written responses were that they wanted to

use their own prompts for the essays and that they want an intervention that would help their students write better quality essays rather than writing more.

Interscorer Reliability and Procedural Integrity

Interscorer reliability was calculated for 17% of sessions across each phase. Overall agreement was taken from the compilation of all stories written by participants. Across all classrooms during the intervention and second baseline phases, the interrater agreement for TWW was 97.8%. During the intervention, interrater was 98%, with values ranging from 91.6% to 100%. During the second baseline phase, interrater was 98% with values ranging from 84% to 100%. These interrater reliability percentages well exceed the minimum standard of 80% agreement.

Procedural integrity was recorded for 33.3% of sessions conducted across all baseline and intervention phases. Across classrooms and phases, procedural integrity was calculated at 100%. This strong procedural integrity indicates that all components of the baseline and intervention procedures were implemented correctly by teachers and researchers.

CHAPTER IV
DISCUSSION

Using a single-subject A-B-A withdrawal design, this study examined the effects of a group contingency writing intervention on elementary students' writing production. Specifically, this study addressed the use of an interdependent group contingency with randomized criterion method on first and second graders' TWW. This intervention was developed with a goal for teachers to implement it in the classroom setting. The first research question focused on whether students' TWW increased with the writing intervention in place. The second question addressed whether students' TWW maintained after the removal of the writing intervention. The third research question aimed to analyze the social validity of the intervention via student and teacher responses.

Research Question 1: Impact of the GC on TWW

Overall, the GC intervention did appear to impact student writing production; however, experimental control was not achieved for any classroom making interpretation difficult. Across all three classrooms, students' average TWW increased compared to the initial baseline.

Based on visual analysis, Class A was the only classroom that displayed any significant effect on TWW from the intervention. The three final baseline data points of Class A are decreasing, which provides the closest approximation to experimental control. Unfortunately, Classrooms B and C displayed an increasing trend during baseline. For experimental control to be established, the baseline data should be relatively stable before introducing the intervention (Gast & Baerkey, 2014). Class A's baseline data are variable, but they are the only classroom that displays a slight decreasing trend prior to the implementation of the intervention. This trend helps to demonstrate some

experimental control because it is moving in the opposite direction of where the intervention aimed to target. Class B and C's baselines were on an increasing trend, thus limiting the support of the intervention impacting the students' TWW. For example, it is plausible that if Class B and C were to continue on baseline, they may have increased their TWW without the intervention in place. Therefore, according to visual analysis, Class A displays the closest proximity to experimental control – even though experimental control was not established.

Since Classes A and C were the only two second-grade classrooms, we can compare their baselines with each other and see that Class C's average TWW during baseline (TWW = 39.7 words) was lower than that of Class A (TWW = 70.9 words). This difference in average TWW during baseline is possibly due to the difference in experience of the teachers. The teacher for Class A was well known in the school for being an effective teacher who implemented beneficial classroom management strategies. Additionally, the teacher for Class A exhibited more enthusiasm with their students compared to the teacher's reactions for Class C. This difference in intervention effectiveness between classrooms is similar to the findings from Pappas, Skinner, and Skinner (2010). They found that students who performed lower at baseline were the most impacted by the reading group contingency intervention. The current study supports this previous finding since Class C was the lowest performing group at baseline who benefited the most from the writing intervention.

Based on visual analyses, Classes A and B did not appear to benefit as much from the writing intervention as Class C. This finding is not consistent with the previous GC

writing intervention study by McCurdy et al. (2008). They found that the GC writing intervention improved all students' writing skills. However, McCurdy et al. (2008) used participants who were ninth grade students with learning disabilities; whereas the participants in the present study were much younger and were not identified as having learning disabilities. Other writing intervention studies that involved young elementary-aged students, such as the performance feedback interventions with third graders by Truckenmiller et al. (2014), also yielded effective results across students. Students in the performance feedback studies were still slightly older than the participants in the current study. It is possible that the age of students in the current study may have been a factor in the effectiveness of the intervention on TWW.

The intervention effectiveness may have been impacted due to a novelty effect by the presence of the graduate researchers. When the graduate researchers entered the classrooms for the first time, students were eager to participate. Teachers even reported that the students wanted to impress the researchers by “showing off” their writing abilities. This behavior may have impacted the baseline data because students may have written more than they typically would on an in-class assignment. Thus, the intervention effectiveness may have appeared minimal compared to baseline because the students may have reached their ceiling level threshold of writing. For example, Class A started at a higher baseline level compared to Classes B and C, yet the intervention effect was minimal compared to that of Class C. Class A may have started out close to their TWW threshold and were not able to write any more due to their young ages. According to national TWW norms, second grade students, on average, write approximately 32 words

during a 3-minute writing probe (AIMSweb, 2017). Class C's TWW is considered below average compared to these norms. However, Class A's TWW at baseline aligns with other second grade students. The national norms indicate first graders' TWW is 21 words written based on a 3-minute writing sample. Class B's average TWW at baseline is considered below average compared to other first graders (AIMSweb, 2017).

At times, during the intervention phase, students looked up during their essays and stated that they did not know what else to write about. The researcher and/or the teacher encouraged these students to continue writing for the entire seven-minute period. It is possible that older students may have more to write about since writing theories explain that they have more advanced cognitive capabilities (i.e., planning and transcription skills) compared to younger students (Berninger & Swanson, 1994).

When looking at the effect size calculations, PND and Hedges' g , they were not consistent with the visual analyses across classrooms. The descriptions of PND effectiveness were derived from Scruggs and Mastopieri (1998). When analyzing the PND across the initial baseline and intervention phases, the intervention was shown to be very effective for Class C only. The PND for Classes A and B did not support an effect on students' TWW during the intervention. However, the results of Hedges' g indicated moderate to large effect sizes across all classrooms and phases. Since Hedges' g uses the average TWW in the calculation, it is not surprising that this effect size calculation yielded significant results. These differences in effect size results are likely due to the targeted components of the calculations. For example, PND is derived from the variability of the data paths whereas Hedges' g is more influenced by the average.

Despite the differences in effect size calculations, researchers evaluating single-case design suggest that visual analysis yields the most accurate representation of the data since the number of participants are small (Kratochwill et al., 2010; Maggin et al., 2019). Therefore, the current study results are best explained by use of a visual analysis method. The visual analysis also allows for accounting nuances within the data set, such as the variability, trend and level changes.

Research Question 2: Maintenance of Results

Overall, student writing production varied and was not maintained. When the intervention was removed. Average TWW initially remained high across all three classrooms. Although students were informed that they would not be working as a group or be rewarded, they continued writing at intervention levels. Classes B and C displayed a sharp decrease in average TWW on the 2nd day of maintenance. Class A had a slight decreasing trend in average TWW during the 2nd baseline phase. Similar to the first research question, visual analyses represent these results more accurately than the average TWW. The average TWWs show that all classes increased from intervention to the last baseline phase: Class A increased by 2.5 words; Class B increased by 4.1 words; and Class C increased by 5.7 words. However, the PND calculation from intervention to 2nd baseline across classrooms indicated an ineffective intervention (Scruggs & Mastopieri, 1998). Due to the high variability, decreasing trends during the second baseline phase, and lack of data it cannot be concluded that the intervention yielded maintenance in writing production. Ideally, this study would have run more essay trials due to the decreasing trends in TWW; however, the timing of the school year did not

allow for this. Additionally, five data points per phase is considered a minimum standard for single-case design research (Kratochwill et al., 2010). Therefore, a maintenance explanation cannot be concluded with these data. This finding is comparable to the results of McCurdy et al. (2008), who also did not find writing skills maintained.

When the researchers returned to baseline, they explained to the students that they would not have to count their words anymore. Thus, it is likely that the students thought they were still able to work for a reward during the first essay of the second baseline phase. It was not until the second or third essays during the second baseline that the students realized they were not working for a reward anymore. Thus, the intervention may have been removed too abruptly to increase the chances of students maintaining TWW, it may have been beneficial to fade out the intervention in stages. According to Skinner (1963), reducing the density of reinforcement over time can allow for the maintenance of a target behavior. The intervention in this study was on a continuous reinforcement schedule and may have been more likely to maintain student TWW if the intervention was placed on a variable reinforcement schedule. An intermittent schedule of reinforcement would add an additional element of randomization and, subsequently, encourage students to continue to write more since they would not know when they would be rewarded for their writing production. Future studies in writing intervention research should examine the impact of thinning reinforcement schedules to aid in the maintenance of targeted skills.

Research Question 3: Social Validity

A primary goal of this study was to develop an intervention that would be easy to implement by teachers. Although previous writing studies have yielded effective results, these interventions are not useful if they are not adaptable to the classroom setting (Hier & Eckert, 2014; Truckenmiller et al., 2014). The present study aimed to have the teachers implement the intervention – and they did with 100% procedural integrity. Social validity measures indicated that the teachers found the intervention easy to implement in their classrooms. Additionally, all teachers reported that their students wrote more with the intervention. The first-grade teacher even stated that their students were writing more in other academic areas as well, such as science. Overall, the teachers found the intervention easy to implement, helpful with increasing the amount of writing among students, and enjoyable to their students.

On the other hand, some teachers suggested that they want their students to write better quality essays rather than writing more. This was not a surprising response since writing quality is a common concern among teachers due to the lack of quality writing instructional material and evidence-based interventions (Graham & Harris, 2003). The present study actually served as a pilot experiment that provided direction to another study addressing writing quality using adjectives and conjunctions as the targeted skills. However, due to the COVID-19 pandemic, this study could not be completed.

Regarding social validity, most students reported that they enjoyed the intervention and working as a group to earn a reward. They also reported that they thought their writing improved with the intervention. This finding is important because as

students progress through school, they can lose motivation to perform well and may require strategies to engage them in the classroom (Menzies, Lane, Oakes, & Ennis, 2017). The group contingency intervention was enjoyable, interactive, and required the participation of all students in order to earn a group reward. Teachers and students reported positive feedback regarding ease of implementation and enjoyable experience.

Limitations and Future Research

Serval limitations to the study have been identified throughout the discussion of the research questions. The primary limitation is that experimental control was not established within classrooms. Second, the young age of the students may have impacted the effectiveness of the study. Additionally, the time of the year was not ideal and impacted the ability to continue gathering data for the second baseline phase. Finally, the removal of the intervention was abrupt and may have impacted the students' ability to maintain their writing production. These limitations can be addressed by collecting more data in each phase, especially in the maintenance phase; examining the impact of GC on writing with older students; and fading out the intervention using an intermittent reinforcement schedule.

In addition to these limitations, another limitation to this study was with the IRB-approved consent requirements. While Class A did have 100% parent consent, Class A had 61% and Class C had 56% of students with consent/assent. If students did not have parent consent, their data could not be included in the analysis of this study. However, students without parent consent were allowed to participate in the classwide intervention. random selection process. Since the rewards were based on average class performance,

the current data do not reliably display how the classes performed as a whole (except for Class B, who had 100% consent/assent). It is possible that some students who did not return their consent/assent forms did not want to share their writing with the graduate students. When the graduate students first introduced the study, some students appeared nervous and asked whether their stories will be given to their parents or teachers. Additionally, some students may have assumed that they would not have to write the essays if their parent did not provide consent. However, teachers repeatedly described to the class that all students were taking part in the writing intervention regardless of parent consent. Future studies should keep classwide consent in mind when developing interventions and interpreting results based on classroom performance.

While evaluating classwide performance is more time-efficient, it may also be worthwhile for future studies to evaluate the effects of a GC intervention on an individual participant level. It is possible that the lowest performing students benefited more from the intervention compared to students who wrote more during baseline. This information can provide insight into how GCs affect various students with differing academic skill levels. For example, Scott et al. (2017) evaluated the effects of an interdependent GC on 16 first-grade students' math performance. The dependent GC was added to the interdependent GC intervention and yielded significant treatment effects for students who were initially lower performers. Analysis at an individual level can provide educators with understanding of which students should be targeted with more concentrated intervention methods, such as a dependent GC.

Other Future Research

Prior to the pandemic, this research group had planned and started a new study with a focus on writing quality over writing production. The procedures were similar to the group contingency designed for the present study, except adjectives and conjunctions were added as targeted writing skills. This study planned to provide direct instruction in the form of “mini-lessons” on conjunctions and adjectives and would allow students to practice these skills via repetition (Kinder & Carnine, 1990). After students wrote their essays, the teacher would randomly select a card that says “TWW, “Adjectives”, or “Conjunctions” from a bag. The frequency criterion differed per writing skill and ranged in value between average scores at baseline. Future studies should evaluate the use of group contingency writing interventions on various writing skills in order to improve the quality of writing among students.

Applied Implications

This study addresses a gap in writing intervention studies by replicating the use of previous group contingency methods in a classroom setting (Pappas et al., 2010; Scott et al., 2017; Truckenmiller et al., 2014). Further, this study addresses previous research recommendations to examine the use of a randomized criterion to motivate students to write longer essays (McCurdy et al, 2008). It was hypothesized that students would write more with the intervention in place, which is supported by the data from the current study. However, due to lack of experimental control, there is no way of knowing whether the group contingency or other unknown classroom variable impacted the students’ writing production.

One primary goal of this study was to develop a writing intervention that is easy to use in the classroom setting. Previous research on writing interventions, specifically performance feedback, has described the time-consuming process of providing individualized feedback to all students in a classroom (Truckenmiller et al., 2014). For teachers to implement research-based interventions in the classroom, more research is needed to examine the feasibility of teachers implementing the procedures in their classrooms (McCurdy et al., 2008). The teachers in the current study reported the ease of implementation and some reported the possibility of continuing the intervention on their own in the future. Practically, this group contingency intervention can be adapted to different academic subjects, ages of students, targeted skills, and types of rewards. For example, older students may obtain tokens when they meet their class goal and receive a reward (i.e., a pizza party) after obtaining five tokens. Each classroom may be more motivated to work for specific rewards rather than a piece of candy, as in the current study.

Although this study presents an intervention method that is feasible in the classroom, there needs to be more research on effective and easy interventions that can be completed by teachers. Additionally, teachers need intervention methods that are the least time-consuming. Teachers may be able to incorporate their own writing prompts with this intervention to increase the likelihood of implementation. Incorporating teachers in the intervention development also may increase their sense of ownership and increase the likelihood of them continuing to use the intervention over time.

Conclusion

Despite the various limitations to the present study, students wrote more when the GC intervention was introduced. Both students and teachers found the intervention to be impactful, easy to follow, and enjoyable. This research is the first that examines the use of a group contingency on classwide writing with the teacher implementing most of the procedures. While much more research is needed, this research provides initial support for using a group contingency to increase writing production in young students.

LIST OF REFERENCES

- Alamargot, D. & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht, Netherlands: Springer Science and Business Media.
- AIMSweb (2017). *AIMSweb national norms table: Written expression – total words written*. NCS Pearson, Inc.
- Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74(1), 29-58.
- Bramlett, R. K., Murphy, J. J., Johnson, J., Wallingsford, L., & Hall, J. D. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools*, 39(3), 327-335. doi: 10.1002/pits.10022
- Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). *The Development of Writing Abilities*. London: Macmillan.
- Cutler, L. & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, 100(4), 907-919. doi: 10.1037/a0012656
- Denune, H., Hawkins, R., Donovan, L., McCoy, D., Hall, L., & Moeder, A. (2015). Combining self-monitoring and an interdependent group contingency to improve the behavior of sixth graders with EBD. *Psychology in the Schools*, 52(6), 562-577. doi: 10.1002/pits.21846
- Eckert, T., Coddling, R. M., Truckenmiller, A. J., Rheinheimer, J. L. (2009). Improving children's fluency in reading, mathematics, spelling, and writing: A review of evidence-based academic interventions. In A. Akin-Little, S. G. Little, M. A.

- Bray, & T. J. Kehle (Eds.), *Behavioral Interventions in Schools: Evidence-Based Positive Strategies* (111-124). Washington, DC: American Psychological Association.
- Gast, D. L. & Baekey, D. H. (2014). Withdrawal and Reversal Designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp. 219-221). London: Routledge.
- Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. M., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternative measures for curriculum-based measurement in writing. *School Psychology Review*, 31(4), 477-497.
- Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 383–402). New York, NY: Guilford Press.
- Graham, S., Harris, K. R., MacArthur, C. A., & Schwartz, S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disability Quarterly*, 14, 89-114.
- Graham, S. & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445-476. doi: 10.1037/0022-0663.99.3.445

- Hansen, S. D. & Lignugaris/Kraft, B. (2005). Effects of a dependent group contingency on the verbal interactions of middle school students with emotional disturbance. *Behavioral Disorders*, 30(2), 170-184.
- Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). *The fourth R: Research in the classroom*. Columbus, OH: Charles E. Merrill Publishing Co.
- Harris, K. R. & Graham, S. (2016). Self-regulated strategy development in writing: Policy implications of an evidence-based practice. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 77-84. doi: 10.1177/2372732215624216
- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. In L. E. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (3-30). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Hedges, G. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, 6, 107-128.
- Hier, B. O. & Eckert, T.L. (2014). Programming generality into a performance feedback writing intervention: A randomized controlled trial. *Journal of School Psychology*, 56, 111-131.
- Kinder, D. & Carnine, D. (1990). Direct instruction: What it is and what it is becoming. *Journal of Behavioral Education*, 1(2), 193-213. doi: 1053-0819/91/0600-0193
- Koenig, E. A., Eckert, T. L., & Hier, B. O. (2016). Using performance feedback and goal setting to improve elementary students' writing fluency: A randomized controlled trial. *School Psychology Review*, 45(3), 275-295.

- Kratochwill, T. R., Hitchcock, T. J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.
- Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavioral Analysis*, 8, 341-347.
- Little, S. G., Akin-Little, A., & O'Neill, K. (2015). Group contingency interventions with children — 1980-2010: A meta-analysis. *Behavior Modification*, 39(2), 322-341. doi: 10.1177/0145445514554393
- Maggin, D. M., Cook, B. G., & Cook, L. (2019). Making sense of single-case design effect sizes. *Learning Disabilities Research and Practice*, 34(3), 124-132. doi: 10.1111/ldrp.12204
- MacArthur, C. A. & Graham, S., & Fitzgerald, J. (2016). *Handbook of writing research* (2nd ed.). New York, NY: The Guilford Press.
- Martens, B. K. & Eckert, T. L. (2006). The instructional hierarchy as a model of stimulus control over student and teacher behavior: We're close but are we close enough? *Journal of Behavioral Education*, 16, 83-91. doi: 10.1007/s10864-006-9026-6
- McCurdy, M., Daly, E., Gortmaker, V., Bonfiglio, C., & Persampieri, M. (2006). Use of brief instructional trials to identify small group reading strategies: A two experiment study. *Journal of Behavioral Education*, 16, 7-26. doi: 10.1007/s10864-006-9021-y

- McCurdy, M., Skinner, C., Watson, S., & Shriver, M. (2008). Examining the effects of a comprehensive writing program on the writing performance of middle school students with learning disabilities in written expression. *School Psychology Quarterly*, 23(4), 571-586. doi: 10.1037/1045-3830.23.4.571
- McMaster, K. L., & Campbell, H. (2008). New and existing curriculum-based writing measures: Technical features within and across grades. *School Psychology Review*, 37(4), 550-566.
- Menzies, H. M., Lane, K. L., Oakes, W. P., & Ennis, R. P. (2017). Increasing students' opportunities to respond: A strategy for supporting engagement. *Intervention in School and Clinic*, 52(4), 204-209. doi: 10.1177/1053451216659467
- National Center for Education Statistics (2002). NAEP 2002: National assessment of educational progress report card. *The nation's report card*. Retrieved from: <https://nces.ed.gov>
- National Center for Education Statistics (2011). Writing 2011: National assessment of educational progress at grades 8 and 12. *The nation's report card*. Retrieved from: <https://nces.ed.gov>
- National Commission on Writing (2003). *The neglected "r": The need for a writing revolution*. Retrieved from: <https://www.nwp.org/cs/public/download/>
- National Writing Project (2019). *What we do*. Retrieved from: nwp.org/what-we-do
- Olive, M. L. & Franco, J. H. (2008). (Effect) size matters: And so does the calculation. *The Behavior Analyst Today*, 9(1), 5-10.

- Pappas, D. N., Skinner, C. H., & Skinner, A. L. (2010). Supplementing accelerated reading with classwide interdependent group-oriented contingencies. *Psychology in the Schools*, 47(9). doi: 10.1002/pits.20512
- Persky, H. R., Daane, M. C., & Jin, Y. (2003). *The Nation's Report Card: Writing 2002*. Retrieved from: <https://nces.ed.gov/nationsreportcard/pubs/>
- Popkin, J. & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with randomly selected components. *School Psychology Review*, 32(2), 282-295.
- Powell-Smith, K. A. & Shinn, M. R. (2004). *Administration and Scoring of Written Expression Curriculum-Based Measurement (WE-CBM) for Use in General Outcome Measurement*. Retrieved from: <https://www.cnyric.org>
- Romeo, F. F. (1998). The negative effects of using a group contingency system of classroom management. *Journal of Instructional Psychology*, 25(2), 130-133. Retrieved from: <http://web.a.ebscohost.com.proxy.lib.utk.edu>
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597-599. doi: 10.22237/jmasm/1257035100
- Scott, K. C., Skinner, C. H., Moore, T. C., McCurdy, M., Ciancio, D., & Cihak, D. F. (2017). Evaluating and comparing the effects of group contingencies on mathematics accuracy in a first-grade classroom: Class average criteria versus unknown small-group average criteria. *School Psychology Review*, 46(3), 262-271.

- Scruggs, T. E., & Mastropieri, M. A. (1998). Synthesizing single subject studies: Issues and applications. *Behavior Modification*, 22, 221–242.
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single- subject research: Methodology and Validation. *Remedial and Special Education*, 8(2), 24-33.
- Shapiro, E. S. & Goldberg, R. (1990). In vivo rating of treatment acceptability by children: Group size effects in group contingencies to improve spelling performance. *Journal of School Psychology*, 28, 233-250. doi: 0022-4405/90
- Sharp, S. R., & Skinner, C. H. (2004). Using interdependent group contingencies with randomly selected criteria and paired reading to enhance class-wide reading performance. *Journal of Applied School Psychology*, 20(2), 29-45. doi: 10.1300/J370v20n203
- Skinner, B. F. (1963). Operant behavior. *American Psychologist*, 18(8), 503-515.
- Troia, G. A. & Graham, S. (2003). The consultant’s corner: “Effective writing instruction across the grades: What every educational consultant should know”. *Journal of Educational and Psychological Consultation*, 14(1), 75-89. doi: 10.1207/S1532768XJEPC1401_04
- Truckenmiller, A. J., Eckert, T. L., Coddling, R. S., & Petscher, Y. (2014). Evaluating the impact of feedback on elementary aged students’ fluency growth in written expression: A randomized controlled trial. *Journal of School Psychology*, 52, 531-548. doi: 10.1016/j.jsp.2014.09.001

- Turco, T. L. & Elliott, S. N. (1990). Acceptability and effectiveness of group contingencies for improving spelling achievement. *Journal of School Psychology, 28*, 27-37.
- Visscher, A. J. & Coe, R. (2002). *School Improvement Through Performance Feedback*. New York, NY: Routledge.
- Wright, J. (2013). Curriculum-based measurement-written expression: Guidelines for use. *'How the common core works' series*. Retrieved from: <https://interventioncentral.org>

APPENDICES

Appendix A. Tables and Figures

Table 1. *Descriptive Statistics of Total Words Written for Baselines and Intervention Phases*

	<i>TWW</i>		
	<u>Initial Baseline</u> M (SD) <i>Range</i>	<u>GC</u> M (SD) <i>Range</i>	<u>Second Baseline</u> M (SD) <i>Range</i>
<i>Class A</i>	70.9 (7.35) 66.4-83.0	85.9 (5.96) 80.1-94.6	88.4 (3.75) 85.6-93.9
<i>Class B</i>	24.8 (4.88) 18.4-30.7	32.6 (5.34) 26.7-41.1	36.7 (6.14) 29.9-41.8
<i>Class C</i>	39.7 (4.66) 32.4-45.4	52.4 (3.97) 47.4-58.7	58.1 (11.60) 49.9-66.3

Note. GC indicates Group Contingency condition. TWW indicates Total Words Written.

Table 2. *Percentage Non-Overlapping Data (PND) for each adjacent phase*

Participants	Baseline- Intervention PND [Size]	Intervention- Withdrawal PND [Size]
Class A	40% [Ineffective]	0% [Ineffective]
Class B	50% [Questionable]	33% [Ineffective]
Class C	100% [Very Effective]	50% [Questionable]

Note. *PND = percentage non-overlapping data points.

Table 3. *Hedges' g Calculations*

	<u>Initial Baseline to GC</u> g	<u>GC to second baseline</u> g
<i>Class A</i>	2.24***	0.50**
<i>Class B</i>	1.52***	0.71***
<i>Class C</i>	2.93***	0.66**

Note. *** Indicates a large effect size; ** Indicates a moderate effect size

Table 4. *Student Responses to Social Validity Survey*

	Class A Average	Class B Average	Class C Average	Average Across Classes
I liked writing stories every week	1.2	1.6	2.1	1.6
I think I wrote more when we did the project than in my regular class	1.2	1.6	1.6	1.5
I think I wrote better stories during the project than in my regular class	1.7	1.8	2.1	1.8
I liked working as a group to get a reward	1.3	1.3	1.4	1.3
I think students in my class liked working as a group to get a reward	1.3	1.7	1.6	1.6
I think it was easy to count my own words for each story	1.3	1.8	1	1.5
I think my teacher should do this group project again	1.9	1.6	2.1	1.8
I think the group project was fair for everyone	1.3	1.8	2	1.7
I would recommend this project to a friend	1.9	1.7	2.1	1.8

Note. Participants responded across a four-point Likert scale modified to use character faces.

1=Extremely agree, 2=Agree, 3=Disagree, 4=Extremely disagree

Table 5. *Teacher Responses to Social Validity Survey*

	Class A Teacher	Class B Teacher	Class C Teacher	Average Across Teachers
I think my students can write better essays because of the intervention.	3	5	3	3.7
I think my students can write longer essays because of the intervention.	4	5	4	4.3
It was easy to implement the intervention on my own.	5	5	4	4.7
I think my students enjoyed the intervention.	5	5	4	4.7
I enjoyed the intervention.	4	5	3	4.0
I would like to continue this intervention for future class writing activities.	4	5	2	3.7
I would like to continue this intervention for future subjects other than writing.	4	4	2	3.3
I think the group reward helped motivate students to write more.	5	5	3	4.3
I would recommend this intervention to other teachers/colleagues.	4	5	3	4.0
There was too much class time used up from the writing intervention.***	4	5	2	3.7
The researchers helped me conduct the intervention on my own.	5	5	2	4.0

Note. *** indicates reverse-scored items. Participants responded across a five-point Likert scale.

1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree

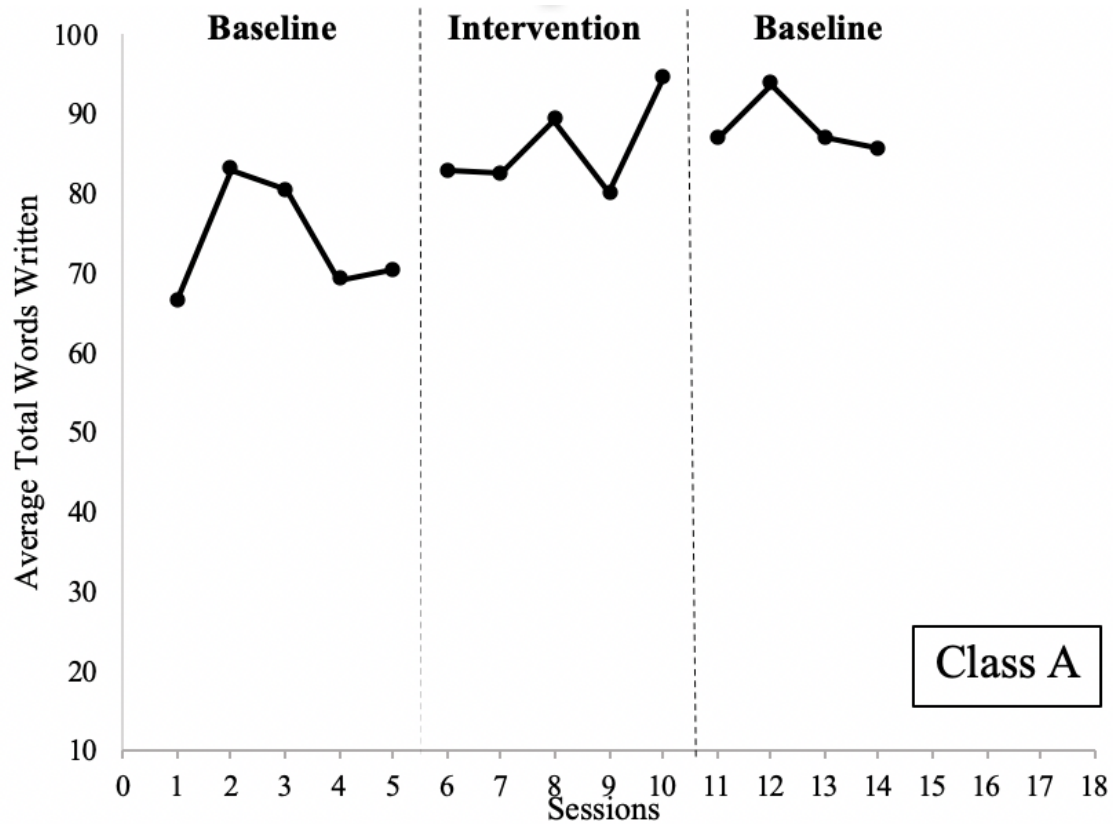


Figure 1. Average total words written by Class A during baseline and intervention phases.

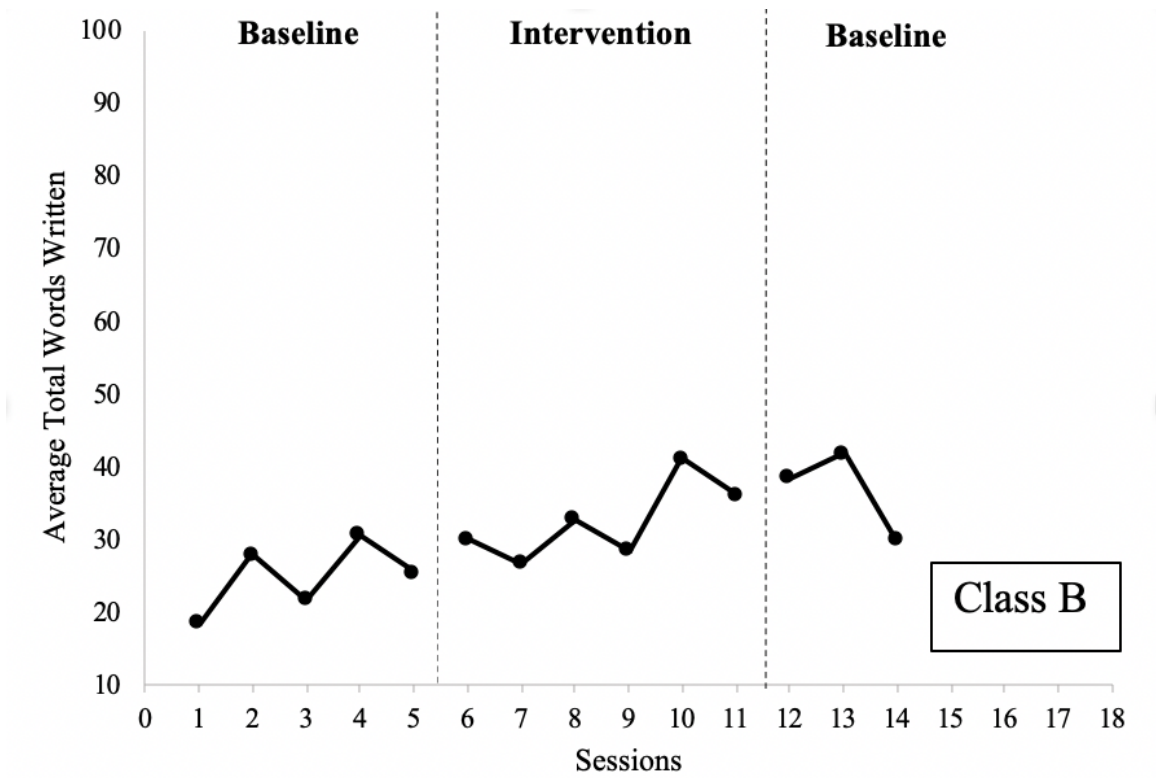


Figure 2. Average total words written by Class B during baseline and intervention phases.

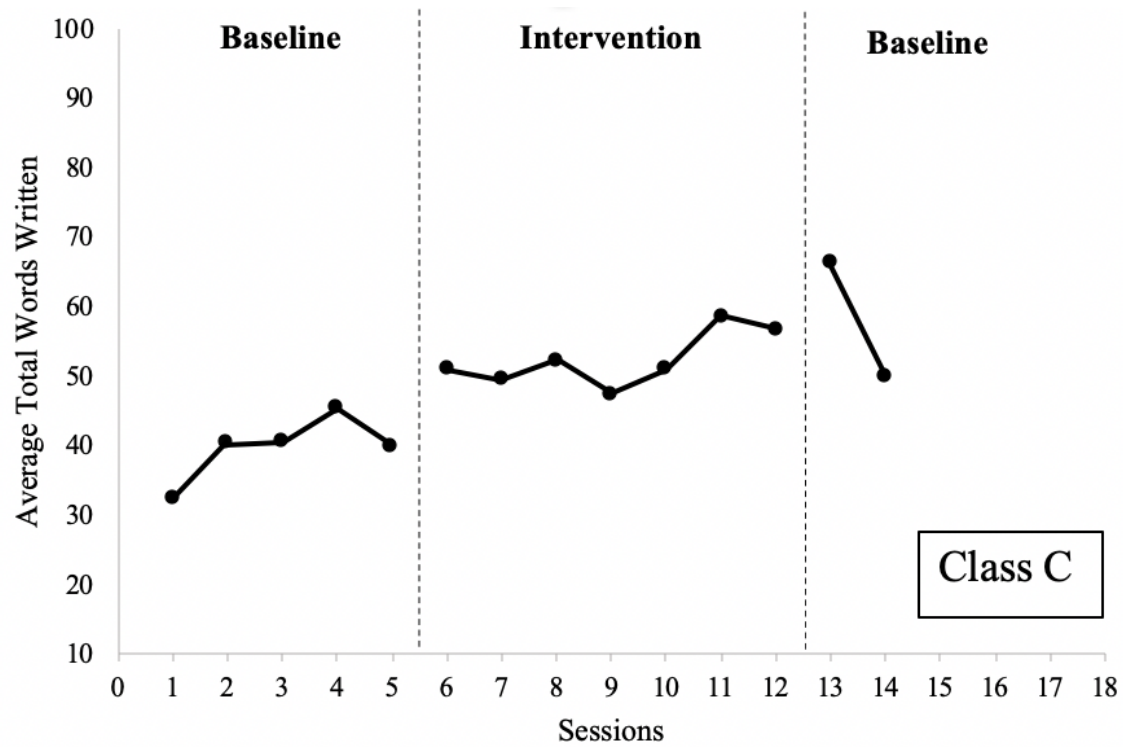


Figure 3. Total words written by Class C during baseline and intervention phases.

Appendix B. Teacher Informed Consent Form

Writing Group Contingency Program Teacher Informed Consent Form

You are invited to be part of a research study being conducted by Madeline Auge, B.S. and Merilee McCurdy, Ph.D. at the University of Tennessee, Knoxville. You are being invited because your classroom will participate in a writing instruction program this year at [REDACTED]. Being in this research study is voluntary, and you should only agree if you completely understand the study and want to volunteer your material and your students' material to be used. This form contains information that will help you decide if you want to be part of this research study or not. Please take the time to read it carefully, and if there is anything you do not understand, please ask questions.

Purpose:

The purpose of this study is to investigate the effectiveness of a classroom group contingency designed to improve students' writing production. Group contingency is an intervention technique that involves a group aiming to achieve a goal in order to access a reward. We plan to publish articles and make presentations at conferences to share the results of this research.

Procedure:

If you choose to participate, we will analyze the materials you and your students create during your time involved in the group contingency writing instruction program. Students will respond to essay prompts for 7 minutes and will count the number of words they wrote and turn it in to you. You will randomly select 5 student papers and record the average word count of those papers. Then, you will randomly pull a number out of a bag. That number will serve as the class goal. If the average word count is higher than the goal, then the entire class will earn a token toward a class reward. This reward will be determined by consulting with you prior to the intervention. Other than the reward, there is no compensation provided to the students. The entire study will take place 3 times a week for up to 3 months. These procedures will require 15-20 minutes each day (approximately 9 hours total). After the intervention, you will be asked to fill out a short survey on your perceptions about the intervention.

A researcher will conduct the program with your students during the first few days. Then, you will take the lead in conducting the procedure with the researcher present to provide any feedback. Once you are comfortable to do this on your own, the researcher will only be present every few days to answer any questions or comments. Ideally, you will be able to implement the program on your own and maintain the stories in your desk until the researchers come to gather the stories. The researchers will also provide you with any extra support and/or resources you may need.

Because these are all things that are part of your regular program evaluation activity for the writing instruction program, being included in the research will not require any additional time.

Teacher initials _____

Risks:

Participation in the program poses no known risks to students. We will monitor students for frustration levels while writing and provide frequent breaks, if needed. There is a slight risk of breach of confidentiality. To minimize the risk of a potential breach of confidentiality, all student data will be de-identified. Prior to any student data leaving the building, all identifying information will be replaced with a researcher generated code. Student names will be removed from all writing samples and any other study data.

Benefits:

Through your participation, you will be helping us to learn more about the role of group contingencies on improving student writing. Additionally, students will learn strategies to help them write longer and better essays.

Confidentiality:

Any information gathered during this program, which may identify your students, will be kept strictly confidential. The only individuals who have access to these records are the participating teachers and researchers. As the teacher, you will be responsible for collecting assent and consent forms from the students and keeping them in your desk until the researchers come to pick them up. The information obtained in this research may be published in scientific journals or presented at professional meetings, but data reported will not identify any individual participant or instructor.

Contact Information:

If you have questions at any time about the study or the procedures, you may contact the researcher, Dr. Merilee McCurdy 520 Bailey Education Complex or 865-974-8144. If you have questions about your rights as a participant, contact the Office of Research Compliance Officer at (865) 974-7697.

Freedom to Withdraw:

Your participation in this research study is voluntary. You are free to decide not to participate in this study or to withdraw your participation at any time without adversely affecting your relationship with the investigators or the University of Tennessee - Knoxville or those at [REDACTED]. Your decision will not result in any loss of benefits to which you are otherwise entitled.

Participation:

Your participation in this research study is voluntary; you may decline to participate without penalty.

Teacher initials _____

CONSENT

I have read the above information. I have received a copy of this form. I understand that my participation in this research study includes allowing Madeline Auge and Merilee McCurdy to use my materials for research purposes. I agree to be included in this study.

Signature of teacher

Teacher name – printed

Signature of researcher

Date

Appendix C. Parent Consent/Child Assent Form

Group Contingency Writing Instruction Parent Permission Form

Your child is invited to be part of a research study being conducted by Madeline Auge, B.S. and Merilee McCurdy, Ph.D. at the University of Tennessee, Knoxville. Your child is being invited because he/she will participate in a group contingency writing instruction program this year at [REDACTED]. Being in this research study is voluntary, and you should only agree if you completely understand the study and want to volunteer to allow your child's material to be used. This form contains information that will help you decide if you want your child to be part of this research study or not. Please take the time to read it carefully, and if there is anything you do not understand, please ask questions.

Purpose:

The purpose of this study is to investigate the effectiveness of a classroom group contingency designed to improve students' writing production. A group contingency is a type of intervention that involves a group working to achieve a goal to access a reward.

Participation:

If you choose to allow your child to participate and your child also agrees, we will analyze the materials your child will create during their time involved in the writing instruction program. Your child will write stories. Then, they will count the number of words they wrote and turn it in to the teacher. The teacher will randomly select 5 student papers and record the average total words written of those papers. Then, the teacher will randomly pull a number out of a bag. That number will serve as the class goal. If the average total words written is higher than the goal, then the entire class will earn a token toward a class reward. If the class meets their goals, your child will receive a reward in class. Rewards will be determined after consultation with your child's teacher (i.e., extra time at recess or free time). The entire study will take place 3 times a week for up to 3 months. These procedures will require 15-20 minutes each day (approximately 9 hours total). Because these are all things that will be part of their regular classroom activities, participation in the research will not require any additional time.

Benefits:

Your child will not receive any direct benefit from allowing their materials to be used in the research project, but we hope to learn things that will benefit teachers and researchers in the future.

Risks:

This research is considered to be no more than minimal risk, which means there is no more expected risk to you than what your child might experience during a typical day. There is the risk of possible loss of confidentiality, as someone could find out your child

was in the study or see their study information, but we believe that risk is unlikely because of the procedures we will use to protect their information.

Confidentiality:

If you and your child agree to participate in the research, we will assign your child a code number and use that instead of their name on all of the materials before we begin analyzing them for the research study. These materials will be stored in a secure location on the UT campus and all data files will be password protected and securely stored. No information which could identify your child will be shared in publications and presentations about this study.

Future Research

Your child's materials may be used for future research studies or shared with other researchers for use in future studies without obtaining additional informed consent from you. If this happens, all of your child's identifiable information will be removed before any future use or sharing with other researchers.

Contact Information:

If you have any questions about this research, please contact me, Madeline Auge, at mauge@vols.utk.edu or my advisor, Merilee McCurdy, at mccurdy2@utk.edu or 865-974-8144. If you have any questions about your rights as a research participant, please contact the Institutional Review Board (IRB) of the University of Tennessee, Knoxville, at utkirb@utk.edu or 865-974-7697. You may also contact the IRB with any problems, complaints or concerns you have about a research study.

Voluntary Participation:

It is completely up to you and your child to decide to be in this research study. Even if you decide to be part of the study now, you may change your mind at any time and stop participating by informing your child's teacher that you no longer wish to participate. Your child will not lose any services, benefits, or rights they would normally have if you choose not to give permission, or if you or your child change your minds and stop participating later.

If you agree that your child may participate, please print and sign the **Parent Permission** section below, and have your child sign the **Assent** section, on both copies of this form. Return one copy to your child's classroom teacher and keep one copy for your records. If you do not wish for your child to participate in the research, it is not necessary to do anything, as we cannot use their materials without your permission and your child's assent.

Parent Permission

I have read the above information. I have received a copy of this form. I understand that my child's participation in this research study includes allowing Madeline Auge and

Merilee McCurdy to use my child's materials for research purposes. I agree that my child may participate in this study.

Child's Name (printed) _____

Parent's Name (printed) _____

Parent's Signature _____ Date _____

Child/Youth Assent

I have talked about this research with my parent(s) and I agree that Madeline Auge and Merilee McCurdy may use my materials for research purposes. If I change my mind, and decide not to participate later, I only need to let my classroom teacher know.

Youth Name (printed) _____

Youth Signature _____ Date _____

IRB NUMBER: UTK IRB-18-04539-XP
IRB APPROVAL DATE: 09/24/2018
IRB EXPIRATION DATE: 09/23/2019

Appendix D. Essay Example Page

The best part about summer is...

[illegible]

Appendix E. Story Starter Prompts

1. The best part about school is...
2. My favorite game to play is...
3. The best part about summer is...
4. My favorite animal is...
5. If I could have any superpower, I would...
6. My favorite memory is...
7. My favorite TV show is...
8. The best part about winter is...
9. When I grow up, I want to be...
10. After school, I like to...
11. If I had a magic pencil, I would...
12. My favorite sport is...
13. One time a spaceship landed in my backyard...
14. My favorite thing to do inside is...

Appendix F. Scripts for Baseline and Intervention Phases

Group Contingency Writing Study Scripts

Baseline

“Today you are going to write a story. The beginning of the story says (*say prompt of the story*). You will have 1 minute to think about the story you want to write. Then, when I say GO, I’d like for you to write as much as you can. After you have written for 7 minutes, I will say stop and ask you to put your pencils down. Please do your best work. Do you have any questions? (*pause*) Begin thinking about your story.”

(Time for 1 minute - if a student writes during this time, remind them to only think about his/her story)

“Go ahead and begin writing your story.”

(Time for 7 minutes)

“Stop. Please put your pencil down and turn your paper over.”

(Make sure all have stopped writing and collect papers)

Intervention

“Today you are going to write a story. The beginning of the story says (*say prompt of the story*). You will have 1 minute to think about the story you want to write. Then, when I say GO, I’d like for you to write as much as you can. After you have written for 7 minutes, I will say stop and ask you to put your pencils down. Then you will count the number of words you wrote and write the number at the top of your paper. Please do your best work. Do you have any questions? (*pause*) Begin thinking about your story.”

(Time for 1 minute - if a student writes during this time, remind them to only think about his/her story)

“Go ahead and begin writing your story.”

(Time for 7 minutes)

“Stop. Please put your pencils down and turn your papers over.” (*pause and check that all have stopped writing*) “Now I would like for you to count the number of words you wrote. When you are finished counting, please write the number of words at the top of your paper. Let me know if you have any questions.”

(Collect papers)

If class met goal: “Great job everyone! We needed _____ total words and we wrote _____ words!”

If class didn’t meet goal: “We needed _____ total words and we wrote ____ words. Let’s try to write more next time!”

Appendix G. Procedural Integrity Checklist Baseline

Group Contingency Writing Study Checklist

Baseline

- _____ Today you are going to write a story. The beginning of the story says (*say prompt of the story*).
- _____ You will have 1 minute to think about the story you want to write. Then, when I say GO, I'd like for you to write as much as you can. After you have written for 7 minutes, I will say stop and ask you to put your pencils down. Please do your best work.
- _____ Do you have any questions? (*pause*) Begin thinking about your story."
- _____ *Time for 1 minute - if a student writes during this time, remind them to only think about his/her story*
- _____ Go ahead and begin writing your story.
- _____ *Time for 7 minutes*
- _____ Stop. Please put your pencil down and turn your paper over.
- _____ *Make sure all have stopped writing and collect papers*
- _____ Thank students for participating.

Appendix H. Procedural Integrity Checklist Intervention

Intervention

_____ Today you are going to write a story. The beginning of the story says (*say prompt of the story*).

_____ You will have 1 minute to think about the story you want to write. Then, when I say GO, I'd like for you to write as much as you can. After you have written for 7 minutes, I will say stop and ask you to put your pencils down.

_____ Then you will count the number of words you wrote and write the number at the top of your paper. Please do your best work.

_____ Do you have any questions? (*pause*) Begin thinking about your story.

_____ *Time for 1 minute - if a student writes during this time, remind them to only think about his/her story*

_____ Go ahead and begin writing your story.

_____ *Time for 7 minutes*

_____ Stop. Please put your pencils down and turn your papers over.” (*pause and check that all have stopped writing*)

_____ Now I would like for you to count the number of words you wrote. When you are finished counting, please write the number of words at the top of your paper. Let me know if you have any questions.

_____ *Collect papers*

_____ Randomly choose 5 papers and average the total words written from those 5.

_____ Pick a number (goal) out of the bag and compare the average words written with the goal number.

If class met goal:

_____ Great job everyone! We needed _____ total words and we wrote _____ words!

If class didn't meet goal:

_____ We needed _____ total words and we wrote ____ words. Let's try to write more next time!

Appendix I. Teacher Social Validity Measure

Social Validity Survey - Teacher

Please circle the number that best represents how you feel about each statement.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think my students can write better essays because of the intervention.	1	2	3	4	5
2. I think my students can write longer essays because of the intervention.	1	2	3	4	5
3. It was easy to implement the intervention on my own.	1	2	3	4	5
4. I think my students enjoyed the intervention.	1	2	3	4	5
5. I enjoyed the intervention.	1	2	3	4	5
6. I would like to continue this intervention for future class writing activities.	1	2	3	4	5
7. I would like to continue this intervention for future subjects other than writing (i.e., math, reading, spelling).	1	2	3	4	5
8. I think the group reward helped motivate students to write more.	1	2	3	4	5
9. I would recommend this intervention to other teachers/colleagues.	1	2	3	4	5

10. There was too much class time used up from the writing intervention.	1	2	3	4	5
11. The researchers helped me conduct the intervention on my own.	1	2	3	4	5

12. What did you like about the intervention?

13. What did you not like about the intervention?

14. Is there anything you wish we would have done more of?

15. Is there anything you wish we would have done less of?

16. Would you recommend the intervention to a teacher or colleague?

17. Are there any additional comments you would like to share?

Appendix J. Student Social Validity Measure

Social Validity Survey – Students Adapted from Kear, Coffman, McKenna, & Ambrosio (2000)

Please circle how you feel about each statement.

1. I liked writing 3 stories every week.				
2. I think I wrote more when we did the project than in my regular class.				
3. I think I wrote better stories during the project than in my regular class.				
4. I liked working as a group to get a reward.				
5. I think students in my class liked working as a group to get a reward.				

6. I think it was easy to count my own words for each story.				
7. I think my teacher should do this group project again.				
8. I think the group project was fair for everyone.				
9. I would recommend this project to a friend.				

10. What was your favorite part of the project? (Circle one)

- a. Working as a team
- b. Learning to write more
- c. Getting the reward as a group
- d. Other: _____

11. Is there anything else you would like to share?

VITA

Madeline Birch Auge was born in Garfield Heights, Ohio to Craig and Elizabeth Auge. She graduated from the University of Dayton in 2015 with a Bachelor of Science degree in Psychology. In August 2016, Madeline accepted a position in the School Psychology doctoral program at the University of Tennessee, Knoxville. She earned a Master of Science degree in Applied Educational Psychology in 2019. Madeline completed a pre-doctoral psychology internship at the Marcus Autism Center through the Department of Pediatrics in Emory University's School of Medicine. She completed her major clinical rotations in the Severe Behavior Intensive Outpatient Clinic and the Language and Learning Clinic, with a minor rotation in Clinical Assessment and Diagnostics. In July 2021, Madeline will complete her doctorate degree.