

5-2015

Curriculum Based Measurement-Written Expression: Effects of Choice and Interest on Third Graders' Writing

Morgan L. Schrader

Western Kentucky University, morgan.schrader@allen.kyschools.us

Follow this and additional works at: <http://digitalcommons.wku.edu/theses>



Part of the [School Psychology Commons](#)

Recommended Citation

Schrader, Morgan L., "Curriculum Based Measurement-Written Expression: Effects of Choice and Interest on Third Graders' Writing" (2015). *Masters Theses & Specialist Projects*. Paper 1490.
<http://digitalcommons.wku.edu/theses/1490>

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.

CURRICULUM BASED MEASUREMENT-WRITTEN EXPRESSION:
EFFECTS OF CHOICE AND INTEREST ON THIRD GRADERS' WRITING

A Specialist Project
Presented to
The Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Specialist in Education

By
Morgan L. Schrader

May 2015

CURRICULUM BASED MEASUREMENT-WRITTEN EXPRESSION:
EFFECTS OF CHOICE AND INTEREST ON THIRD GRADERS' WRITING

Date Recommended March 31, 2015

Carl Myers

Dr. Carl Myers, Director of Specialist Project

Elizabeth A. Jones

Dr. Elizabeth Jones

Sally Kuhlenschmidt

Dr. Sally Kuhlenschmidt

Calvin
Dean, The Graduate School

4-25-15
Date

ACKNOWLEDGMENTS

First and foremost, I must thank my thesis chair, Dr. Carl Myers. This project would not have been possible without his constant support and hard work throughout this process. I have been very fortunate to be your student and have your support throughout these three years of graduate school. In addition, I would like to thank my committee members, Dr. Elizabeth Jones and Dr. Sally Kuhlenschmidt, for their hard work and for serving on my committee.

I would also like to thank my parents, Robert and Leanna Schrader, brother, Zachary, and other family members for their constant encouragement. I am truly blessed for such an amazing family who gives me constant love and support and would not be where I am today without you all. Finally, I would like to thank my friends and cohort members for the friendships and memories we have made and for helping me through the good and bad times. Thank you all.

TABLE OF CONTENTS

	Page
List of Tables.....	v
Abstract.....	vi
Introduction.....	1
Literature Review.....	3
Method.....	25
Results.....	32
Discussion.....	40
References.....	45
Appendix A: Consent and Assent Forms.....	49
Appendix B: Story Starters.....	51

LIST OF TABLES

	Page
Table 1. Frequency of Gender-Specific and Gender-Neutral Story Starters Chosen by Participants.....	33
Table 2. Mean Scores for Boys and Girls with Gender-Specific and Gender-Neutral Story Starters.....	35
Table 3. Two-Way ANOVA Results for Third Grade Students' Choice of Story Starter (n = 153).....	36
Table 4. Mean Scores per Level of Interest in the Story Starter by Gender....	38
Table 5. Two-Way ANOVA Results for Third Grade Students' Interest Level of Story Starter (n = 159).....	39

CURRICULUM BASED MEASUREMENT-WRITTEN EXPRESSION:
EFFECTS OF CHOICE AND INTEREST ON THIRD GRADERS' WRITING

Morgan Schrader

May 2015

52 pages

Directed by: Dr. Carl Myers, Dr. Elizabeth Jones, and Dr. Sally Kuhlenschmidt

Department of Psychology

Western Kentucky University

Curriculum-Based Measurement (CBM) has many effective uses in the school system and its use is increasing. Most CBM research focuses on the area of reading and little on writing. Only one study was found that examined the effects of choice in story starters on students' writing performance, and a few studies examined the effects of interest level of writing topics, although none using CBM as a measure. Past research has indicated girls score significantly higher on CBM measures of writing than boys. This study investigated if choice and interest level in a story starter topic have an effect on students' writing performance.

This study investigated the differences between third grade boys and girls on two different CBM-Written Expression probes using production-dependent, production-independent, and accuracy scores. The results indicate that when no choice is given, girls do score higher than boys. However, when given a choice of story starter topics, boys' performance on the CBM measures is comparable to girls' performance. Furthermore, higher interest in story starter topics lead to higher scores as well.

Introduction

Curriculum Based Measurement (CBM) is a brief fluency measure intended to indicate students' performance in necessary academic skills (Shinn, 1989). These academic skills are in the areas of written expression, reading, spelling, and math. CBM fluency measures examine proficiency with skills by only allowing a certain amount of time to complete the tasks. CBM measures typically use one to three minute probes, thus allowing for a quick way to evaluate a student's performance in a particular area. CBM is a standardized tool that uses standard instructions, probes, scoring guidelines, and forms for recording scores.

When CBM was first developed, specifications for developing the CBM probes were provided to make sure the probes represented the area they intended to assess (e.g., math, reading) and to ensure the probes were comparable, increasing the effectiveness of the tool for making instructional decisions (Shinn, 1989). Originally, CBM probes had to be developed by the examiner based on the curriculum used in that school district. Now, commercially available CBM materials (e.g., AIMSweb, EasyCBM) are typically used, minimizing the need to develop one's own CBM probes. Each commercially available program, like AIMSweb (2008), has standardized guidelines for administering and scoring CBM probes, and includes normative data for comparison purposes.

Each CBM area (i.e., written expression, reading, spelling, and math) has separate guidelines for administration and scoring (Deno, 2003). In reading, students are required to read aloud from text and sometimes select words deleted from text. In writing, students are required to write short stories when given a story starter. In spelling, teachers dictate a list of words for the students to spell. In math, students solve sheets of

computation problems. Thus, the materials, directions, and tasks for CBM are very similar to what children do on a daily basis in school (Hosp, Hosp, & Howell, 2007).

Academic performance is measured by determining the CBM score. Correct and incorrect responses made in a fixed time period are calculated. These data are collected through direct observation procedures. After CBM data are collected and scored, an individual student's scores are compared to norms (e.g., state, district) based on the student's age and grade.

Most CBM research focuses on the area of reading and few studies are on writing. This specialist project examined CBM-Written Expression (CBM-WE). CBM-WE examines different skills involved in writing (e.g., amount produced, spelling) by providing students with a topic (called a "story starter") and then allowing three minutes to write (AIMSweb, 2008). Past research has noted gender differences, with girls scoring significantly higher than boys. In the area of CBM-WE, only one study was found that examined the effects of choice in story starters on students' writing performance. While some research has examined the effects of interest related to students' writing (Albin, Benton, & Khramtsova, 1996; Benton, Corkill, Sharp, Downey, & Khramtsova, 1995), no studies were found that specifically looked at students' interest with CBM-WE story starters and scoring measures. The present study intends to add to current CBM-WE research by specifically addressing the following research questions:

Research Question 1: Does choice of a story starter have an effect on third grade students' writing performance?

Research Question 2: Does interest level in a story starter topic have an effect on third grade students' writing performance?

Literature Review

The current literature review provides an overview of the uses of Curriculum Based Measurement (CBM), along with a description of how it is an efficient measure that has been shown to be reliable and valid. An emphasis is provided specifically in the CBM area of Written Expression (CBM-WE) along with previous research on CBM-WE. Of note for this investigation is the research on how choice and interest impacts students' writing because this Specialist Project focuses specifically on those factors related to CBM-WE.

Uses of CBM

Deno (2003) explained how CBM was originally developed to test the idea that repeated assessment data could be used to evaluate instruction and help teachers improve their effectiveness. CBM data are sensitive to small changes in students' skills, making the evaluation of instruction easier. Since its original development, additional uses have been identified in the school. Deno (2003) notes some of the uses of CBM data include: (a) screening to identify students academically at risk, (b) progress monitoring, (c) improving individual instructional programs, (d) predicting performance on important criteria (e.g., high-stakes assessments), (e) enhancing teacher instructional planning, (f) increasing ease of communication, (g) reducing bias in assessment, (h) offering alternative special education identification procedures, (i) recommending and evaluating inclusion, (j) measuring growth in secondary school programs and content areas, (k) assessing English Language Learners, and (l) predicting success in early childhood education. Some applications of CBM data are utilized more often than others. Two of the main uses of CBM data are described in the following paragraphs.

Screening to identify students academically at risk is a main use of CBM data in the school setting (Cusumano, 2007; Espin et al., 2008). Because CBM administration and scoring procedures are standardized, an individual's performance can be compared to that of a group. Commercially developed norms (e.g., AIMSweb) or local norms are commonly used for this purpose. Norms are an effective way to determine an individual's performance, or an entire class's performance, compared to a particular group (e.g., class, school, district, etc.).

A second primary use of CBM is to progress monitor student or class performance. CBM measures are sensitive to small changes in progress, have high reliability with frequent administrations, and are quick and easy to administer (Goo, Watt, Park, & Hosp, 2012). With or without norms, CBM data can be used to monitor a student's progress over a period of time by comparing progress to prior student performance. Gathering progress-monitoring data can be used to help determine if interventions are working, for a student or class, by seeing if changes have occurred in performance. Using CBM data to progress monitor can guide decisions to make changes to interventions or classroom curriculum.

Efficiency of CBM

In the past, mastery measurement was the dominant approach to monitoring students' progress (Fuchs, 2004). With mastery measurement, a hierarchy is created of instructional objectives that create a sequence ending with the annual curriculum goal. Mastery measurement can be time-consuming and its effectiveness can vary depending on the teacher's skill with such measures. The use of CBM is time efficient, taking only 1 to 3 minutes to administer, depending on the performance being examined and the

number of samples needed to maximize reliability (Fewster & Macmillan, 2002). Short administration time provides the ability to conduct multiple administrations across time, providing multiple performance samplings. Another advantage of CBM is the ease in which professionals (e.g., teachers) can be taught to use the procedures of CBM and obtain reliable data. As previously mentioned, CBM is very similar to what children do on a regular school day (Hosp et al., 2007). This makes it easy for teachers to be taught the procedures and scoring methods, providing a method of progress monitoring for the class (Deno, 2003). The ease of administration is another advantage of CBM. CBM data can be collected fairly effortlessly and efficiently on an individual student or in a whole group setting. With the exception of reading, the other areas of CBM can be administered in a group format (e.g., classroom).

Reliability and Validity of CBM

When determining whether or not to use a test or measure, reliability and validity are two important aspects that need consideration. Reliability and validity are technical qualities required of any measurement tool to be used for educational decision-making (McMaster & Espin, 2007). Reliability is the precision, accuracy, and consistency of a measurement procedure (e.g., across settings, people, time, etc.), according to Bruton, Conway, and Holgate (2000). The validity of a measure refers to how well it measures what it intends and claims to measure (Thorndike, 2005). Considerable research has been conducted on CBM and the overall general conclusion is that CBM has substantial research support for its reliability and validity (Cusumano, 2007; Deno, 2003; Fuchs, 2004). However, the reliability and validity conclusions regarding CBM have not always been consistent and depend on the area (e.g., writing vs. reading) being assessed or

scoring method used (Gansel, Noell, VanDerHeyden, Naquin, & Slider, 2002). For example, CBM-Reading yields higher reliability when three probes are administered and the average of the three probes are used (AIMSweb, 2008).

Summary of CBM

In general, research strongly supports CBM as an effective tool in the school system for many different purposes (Cusumano, 2007; Deno, 2003; Espin et al., 2008). CBM data are especially useful for screening to identify at risk students and progress monitoring students' growth in skills. Most research regarding CBM has focused on the area of reading and little on the other areas assessed by CBM, including Written Expression (Fuchs, 2004). The following section provides an overview of CBM-Written Expression (CBM-WE) and research in that area.

Curriculum Based Measurement-Written Expression Overview

Curriculum Based Measurement-Written Expression (CBM-WE) assesses a student's writing skills (Cusumano, 2007). CBM-WE is a brief fluency measure that results in a sample of writing, allowing the assessment of several different skills involved in writing (e.g., amount produced, spelling). To administer CBM-WE, a predetermined story starter is read aloud to the students and the students are given one minute to think about the story starter and what they are going to write, being prompted to continue to think about the story starter after 30 seconds. After one minute, students are asked to begin writing for three minutes. Ninety seconds after beginning, students are prompted to continue to write about the story starter. Once the three minutes has expired, students are asked to put their pencils down.

There are three categories of methods of scoring CBM-WE probes: (a) production-dependent indices, (b) production-independent indices, and (c) accurate production indices (Malecki & Jewell, 2003). Originally, only production-dependent indices were used for scoring CBM-WE (Shinn, 1989). The three traditional production-dependent measures are Total Words Written (TWW), Words Spelled Correctly (WSC), and Correct Word Sequences (CWS). These measures were called production-dependent because the more a student “produces,” the higher the score. TWW is defined by the total number of groupings of letters broken by a space. The total number of legible words spelled correctly defines WSC. A CWS is defined by two adjacent words containing correct grammar, spelling, punctuation, and context (Malecki & Jewell, 2003).

Production-independent indices measure writing accuracy rather than just the amount (Malecki & Jewell, 2003). With this type of measure, a student’s score is based on percentage correct and a higher score is a result of a higher level of writing accuracy. Thus, even though a student may not write as much as same-aged peers, a high level of accuracy would provide information that could decrease the concern about a student’s writing skills. Production-independent indices include Percentage of Words Spelled Correctly, Percentage of Words Correctly Sequenced (Percent CWS), and Percentage of Error. With the exception of Percentage of Error, all percentages are determined by dividing the number correct by the total number possible and multiplying by 100. Percentage of Error is computed by taking the number of word sequence errors, dividing them by the total number of possible correct sequences, and multiplying by 100.

Accurate-production indices measure both writing fluency (amount produced) and accuracy (Malecki & Jewell, 2003). Students who write the most with high levels of

accuracy obtain the highest scores. Students who do not write much and make several errors would have the lowest scores. Students who have only production or accuracy deficits would still have relatively low scores. Correct Minus Incorrect Word Sequence (CMIWS) is the most common accurate-production index. CMIWS is computed by subtracting the number of incorrect word sequences from the number of correct word sequences. CMIWS is sometimes referred to as Correct Word Sequence Minus Incorrect Word Sequence (CWS-IWS) in some studies.

CBM-WE has multiple effective uses within the school system. Progress monitoring, determining intervention effectiveness, screening, driving Individual Education Plans (IEPs), and driving instruction are some of the main uses of CBM-WE (Hessler & Konrad, 2008). Through progress monitoring and screening, CBM-WE data allow a comparison of a student's writing to a norm group (e.g., class, grade, age), helps guide decisions when implementing interventions, and provides academic information when considering referring a student for special education services. Frequent CBM-WE monitoring allows the teacher to examine the student's writing progress over a particular amount of time. By examining the different areas where students may be having difficulty, teachers can change/add to the current instruction. If a student is currently receiving special education services, these data can be effectively used to form IEP writing goals for a student.

CBM-WE is time efficient, taking only 3-5 minutes to administer, providing the ability to conduct multiple administrations across time to gain multiple performance samplings (Fewster & Macmillan, 2002). Professionals can easily be taught to use the administration and scoring procedures of CBM-WE to obtain reliable data. Because the

writing probes are very similar to what children do on a regular school day (Hosp et al., 2007), teachers can easily be taught the procedures and scoring methods, providing a method of progress monitoring for the class (Deno, 2003). The ease of administration is another significant feature of CBM-WE. CBM-WE data can be collected effortlessly and efficiently on a group (e.g., class) basis. One limitation is that scoring procedures can be time consuming, depending on the number of probes given to a student and total number of students given probes.

Research supports CBM-WE as an effective tool in the school system for many different purposes. CBM-WE is especially useful for progress monitoring, making sure students are progressing in writing proficiency at a rate similar to the student's same age- and grade-level peers. The following section provides an overview of CBM-WE research.

Curriculum Based Measurement-Written Expression Research

While many studies have been conducted on the CBM area of reading, relatively few studies have been conducted on CBM-WE. McMaster and Espin (2007) conducted a literature review on the technical features, development, and technical adequacy of CBM-WE. Their article reviewed 14 technical reports and 18 published articles on CBM-WE. The technical reports originated from the Institute for Research on Learning Disabilities at the University of Minnesota where CBM measures were first developed, during the 1980's. McMaster and Espin's review of those technical reports showed high test-retest, alternate-form, and internal consistency reliability for CBM-WE ranging between $r = .71$ to $.92$. High criterion validity coefficients with other measures of writing, $r = .69$ to $.88$, were also found for CBM-WE.

Overall, McMaster and Espin (2007) concluded that the research on CBM-WE

demonstrated it is sensitive to growth, allowing the ability to use CBM-WE to monitor students' progress. Their reviewed articles also showed that CBM-WE data can be successfully used to assess students at different skill levels, for screening purposes, and to measure beginning writers' written expression skills. McMaster and Espin's review of the CBM-WE literature indicated a large number of CBM-WE studies supported the many uses of these data in schools. The published articles reviewed by McMaster and Espin up to that point in time were obtained and those most pertinent to this Specialist Project were selected to be reviewed in this section of the literature review. In addition, more recent studies on CBM-WE will also be included in this section.

Alternate CBM-WE scoring methods. Gansle, Noell, VanDerHeyden, Naquin, and Slider (2002) conducted one of the first published studies to examine CBM-WE. Gansle et al. examined the predictor criterion relationship between several measures of writing competence and teachers' assessment of students' writing skills, along with standardized group tests of writing. The authors were trying to determine if novel-scoring methods, in addition to the traditional production-dependent measures, could be applied to CBM-WE and still have strong correlations with the criterion variables. In their study, 179 third and fourth graders received two 3-minute CBM-WE probes on two consecutive days. Predictor variables included: TWW, parts of speech, long words, WSC, total punctuation marks, correct punctuation marks, correct capitalization, complete sentences, words in complete sentences, CWS, sentence fragments, simple sentences, and computer-scored variables. Criterion variables included teacher rankings of students' writing skills and standardized writing test scores. Results showed that correct punctuation marks, CWS, and words in complete sentences correlated the highest

with other written expression measures. The authors concluded that correct punctuation marks and words in complete sentences showed the most promise to serve as additional indices of writing skills, along with the traditional production-dependent measures like CWS and TWW. The study noted that the main limitation was the small number of participants and in only two grades, making generalization difficult.

In a later study, Gansle et al. (2004) examined students' writing improvement following an intervention and the relationship between the nationally norm-referenced, individually administered Woodcock Johnson-Revised tests of achievement and CBM-WE. The 47 participants were randomly chosen third and fourth grade students. Students were given two 3-minute CBM-WE probes. One probe was given before a 22-25 minute intervention and the other following the intervention. Participants were also administered the Woodcock Johnson-Revised during the same week. During the intervention, participants brainstormed ideas for a story starter as a group, and produced a sentence on paper, which received writing quality feedback. For this study, six variables, as determined by Gansle et al. (2002) as having the highest correlation coefficients, were used to score the CBM-WE probes: TWW, CWS, total punctuation marks, correct punctuation marks, words in complete sentences, and simple sentences. Total punctuation marks, simple sentences, and words in complete sentences were the best predictors of the Woodcock Johnson Revised Writing Samples subtest scores, based on regression analyses. However, only students' TWW scores improved following the intervention.

Gender differences. Malecki and Jewell (2003) conducted another one of the earlier published studies examining CBM-WE and revealed gender differences in their

results. Malecki and Jewell administered CBM-WE probes to 946 students in first through eighth grade using production-dependent, production-independent, and accurate-production scoring indices, twice in a school year (i.e., fall and spring). The study investigated which scoring indices are appropriate with regard to grade and gender. Results were examined by groups of early elementary (grades 1 and 2), elementary (grades 3, 4, and 5) and middle school (grades 6, 7, and 8) grades. Results, as would be expected, showed that older students out performed younger students, and that students performed higher at the end of the school year than at the beginning.

Malecki and Jewell's results also indicated that girls out performed boys across each age grouping. Furthermore, results showed that at the early elementary and elementary grade levels, the production-dependent (fluency), production-independent (accuracy), and accurate-production (accurate fluency) writing measures were all significantly related, leading to the conclusion that all types of measures were useful at the elementary levels (i.e., grades 1-5). However, for middle school students, how much the students wrote was not significantly related to the accuracy measures. The authors concluded that production-dependent measures were less reliable at older grade levels. Thus, for older grade levels in particular, production-independent and accurate-production measures were considered the most suitable types of scores to use.

Jewell and Malecki (2005) conducted a follow-up study that examined the use of CBM-WE assessment measures in the three categories of production-dependent, production-independent, and accurate-production indices. These categories were used to compare written language scores across grade level and gender. There were a total of 203 students from second, fourth, and sixth grades who were included as participants. In

addition to CBM-WE, the following measures were used: Tindal and Hasbrouck Analytic Scoring System (THASS), Stanford Achievement Test (SAT), and the students' classroom Language Arts grades for fall semester.

Jewell and Malecki's results showed gender differences in production-dependent measures, where girls outperformed boys. However, there were no significant gender differences on the production-independent or accurate production indices. When examining grade levels, they found that students in higher grades had significantly higher scores on all CBM assessment indices except for production-independent scores between fourth and sixth grades. Similar to their past findings (i.e., Malecki & Jewell, 2003), production-dependent measures became less reliable when students reached the sixth grade, but were appropriate for the younger elementary levels (i.e., grades 2 and 4). Many of the CBM-WE scoring indices were significantly correlated with SAT subtest scores, language grades, and scores on the THASS for lower grades, but fewer significant correlations occurred as grade levels increased. For example, by grade 6, CWS was the only CBM-WE scoring index that was significantly related to the THASS scores.

Melloy (2012) examined a data set already established by Youngman (2010), consisting of 1,348 students in first through fifth grade. CBM-WE probes were administered three times over the course of a school year (i.e., fall, winter, spring) and were scored using the standard production-dependent procedures of TWW, WSC, and CWS. Results showed students obtained higher scores in higher-grade levels and also indicated there were gender differences with writing skills, where girls outperformed boys, in every grade. As the grade levels increased, so did the size of the differences between girls' scores and boys' scores.

Hogston and Schrader (2012) did a follow-up study on Melloy's (2012) data to determine if gender differences would still exist on production-independent (i.e., Percent CWS) and accurate-production (i.e., CMIWS) scoring indices. Results were examined by grade and gender. Students in higher-grade levels scored significantly higher on the production-independent measure (i.e., Percent CWS) and accuracy measure (i.e., CMIWS) than did students in lower grades. While girls' mean scores were higher than boys on both types of measures, a statistically significant difference between boys and girls was only found for the dependent variable, CMIWS. There was no significant difference comparing boys and girls at any grade level using the production-independent measure, Percent CWS.

Ferrington et al. (2014) found parallel results in a study that included 1,240 students in grades 3-8. Participants were from five different schools, three elementary and two middle schools, in a rural southeastern school district. CBM-WE probes from AIMSweb were administered to participants, class by class, three times over the school year, with approximately 12 weeks between each administration (i.e., fall, winter, spring). Each CBM-WE probe was scored using TWW and CWS. A two-way repeated measures analyses of variance was used to examine differences between gender and grade levels. Results were comparable to previous research and found that girls significantly outperformed boys on both scoring indices (i.e., TWW and CWS) at each grade level. Overall, higher-grade levels outperformed lower grade levels for both genders, but girls' scores increased significantly more than boys' scores from fall to spring benchmarks.

Beginning writers. Coker and Ritchey (2010) examined CBM-WE with kindergarten and first grade students. This study examined the adequacy of CBM-WE

with younger students in the early primary grades in an effort to help develop and validate measures for beginning writers. Participants included 233 kindergarten and first grade students from a school district with students from diverse ethnic backgrounds. Pearson correlation coefficients between alternate CBM writing probes were computed for alternate-form reliability on the production-dependent scores. The following results were found for alternate-form reliability: .74 and .77 for TWW, .80 and .75 for CWS, and .81 and .77 for WSC. Other qualitative measures of the students' writing were determined and compared with the CBM-WE production-dependent measures to assess criterion-related validity: .20 - .46 with basic writing and contextual writing for kindergarten, .43 - .46 with the total qualitative score for kindergarten, .25 - .57 with spelling and writing samples subtests for first grade, and .53 - .59 with spelling, writing samples, and broad writing scores for first grade. Overall, results showed that both CBM production-dependent measures and qualitative scores of writing performance had favorable alternate-form reliability, low to moderate criterion-related validity, and were sensitive to growth for primary grade school students.

McMaster, Xiaoqing, Parker, and Pinto (2011) described how few research studies have been conducted to examine the reliability and validity for early primary students' writing, defined as kindergarten through second grade. The authors reviewed the literature on CBM-WE related to beginning writers. Their conclusions indicated that CBM-WE yields adequate reliability and validity to assess beginners' (i.e., kindergarten through second grade) writing skills. CBM-WE was also noted to be sensitive to growth at that young age level, making the method useful for progress monitoring. Finally, the authors noted that the multiple ways in which CBM-WE can be measured provides a

good view of young students' overall writing ability.

Predictive validity. Espin et al. (2008) examined if CBM-WE could predict success on state standardized tests for high school students. For the study, 183 high school students from diverse backgrounds were assessed using the measures WSC, TWW, CWS and Correct Minus Incorrect Word Sequence (CMIWS). Tenth grade students completed two, 10-minute CBM-WE probes in the fall. The predictive variables were scores of the students' writing, taken as samples after 3, 5, 7 and 10 minutes. In January, students also received the Minnesota Basic Standards Test (MBST), a high stakes test required for graduation. Students' performance on the MBST was the criterion variable.

Alternate-form reliability for the two CBM-WE probes ranged from .64 to .85 (Espin et al.). Alternate-form reliability was higher the longer amount of time writing occurred, with scores for the 7 and 10 minute probes yielding the highest reliability coefficients. Predictive validity correlations did not show significant differences between time frames of writing but showed significant differences between scoring indices (e.g., TWW vs. WSC). Predictive validity correlations with the MBST were as follows: .23 to .31 for TWW and WSC, .43 to .48 for CWS, and .56 to .60 for CMIWS. This study demonstrated a significant relationship between CBM-WE and success on state standardized tests for high school students, especially when using CMIWS.

Amato and Watkins (2011) conducted a study that examined the predictive validity of CBM-WE on a sample of 447 eighth-grade students. Previously, little research had been conducted to demonstrate CBM-WE's predictive validity for older students. CBM-WE probes and the Test of Written Language-3rd edition (TOWL-3) were

given and results of the two were compared to determine the relationship between them. One CBM-WE probe was used and the predictor variables were 10 writing indices: TWW, WSC, %WSC, CWS, %CWS, CMIWS, number of sentences, number of correct capitalizations, number of punctuation marks, and number of correct punctuation marks.

Amato and Watkins' analyses indicated three of the predictor variables in the multiple regressions were highly correlated. To reduce estimation problems, three predictor variables (i.e., WSC, CWS and total punctuation marks) were deleted from this analysis. These three predictor variables were eliminated because of their high correlations with other predictor variables and because of past research showing accuracy-based measures as more reliable for secondary students than fluency-based measures. All seven of the remaining predictor variables collectively accounted for 44% of the variance in students' overall writing quotient scores on the TOWL-3. Regression analyses revealed that more complex fluency measures (i.e., number of correct punctuation marks) and accuracy measures (i.e., CWS) were the best predictors of written expression for eighth-grade students, at least as measured by the TOWL-3. Amato and Watkins suggested that simple fluency measures were not adequate for assessing secondary students' writing. Of all the measures, Percent CWS showed the strongest bivariate correlation with the TOWL-3 scores ($r = .61$). The authors concluded that CBM-WE does have significant predictive validity at the secondary level, but the usable CBM-WE indices are limited.

Summary of CBM-WE research. Relatively few research studies have examined CBM-WE. However, for the majority of studies examining CBM-WE, results are fairly consistent. While CBM-WE studies indicate lower reliability and validity

coefficients than other areas (e.g., reading), the results generally indicate adequate reliability and validity with several of the scoring indices. Multiple purposes and uses of CBM-WE have been demonstrated in the school setting, such as screening to identify at-risk students and progress monitoring.

CBM-WE data can effectively be used even when examining beginning writers' performance. CBM-WE data can also be used to examine secondary level students' writing, but certain scoring indices become less reliable as grade level increases. Studies have shown that across grade levels, girls outperform boys on CBM-WE measures and that older students outperform younger students.

Overview of Interest and Choice in Writing

While the literature reviewed on CBM-WE has been favorable, the consistent finding that girls score higher on CBM-WE measures than boys is of particular interest to this Specialist Project. Can the difference between boys and girls be minimized if a choice of writing topic is given, or students' interest in the topic is enhanced? The current section of this literature review is an overview of research examining the effects of interest factors and choice on writing, along with gender differences in writing performance.

In a recent dissertation, Bleck (2013) examined the effects of choice versus no choice of story starter by evaluating its impact on writing production (i.e., TWW and CWS) and writing accuracy (i.e., %CWS). This study also evaluated the score differences between males and females. Participants included 83 ninth graders in a Midwest high school. CBM-WE probes were administered over four consecutive weeks, using AIMSweb story starters and administration guidelines. The independent variable

was choice or no choice. Students wrote two stories each week; one where they were allowed to choose among eight story starters and one where they were not given a choice of story starter.

Bleck's results were mixed but mostly indicated that choice in story starters did not significantly increase students' writing production (i.e., TWW and CWS) or accuracy (i.e., %CWS) over writing samples when the students were not given a choice, when boys' and girls' scores were combined. In fact, during the first week when the students were not given a choice, their TWW and CWS scores were significantly higher than when they were given a choice. There were no significant differences for TWW or CWS during any other week. There were no significant differences for %CWS for any week. Bleck did find girls' scores were significantly higher than boys' scores, regardless of choice/no choice condition, on all indices (e.g., TWW, CWS, and %CWS), except during week 3. During the third week, an interaction effect between gender and choice was significant in that boys had significantly higher TWW and CWS scores when given a choice and the boys' scores were very close to the girls' scores. While such a finding only occurred in one of her four CBM-WE administrations, Bleck concluded that "it appears that offering male students a choice of writing topic could improve their writing performance and make it comparable to the performance of female students" (p. 76).

Troia, Harbaugh, Shankland, Wolbers, and Lawrence (2013) examined the relationships between writing motivation, writing activity, and writing performance, but did not use CBM-WE. In the study, 618 students participated from nine Midwest schools and six Pacific Northwest schools. Participants came from diverse backgrounds and were in grades 4 through 10, excluding eighth grade. Measures used for the study were the

Writing Activity and Motivation Scales (WAMS) instrument, teacher judgment of writing ability, and narrative writing quality. The WAMS was group administered, about 3 months after starting school and consists of 30 items related to writing motivation. These items give separate scores for self-efficacy beliefs, interest, perceived task value, attitudes, goal orientations, and attributions for success and failure. Teachers' judgment of students' writing abilities included a 9-point decile scale, where students were ranked compared to peers. For the narrative writing, two fictional prompts were given to provide participants a choice of topics. The scoring rubric included a 6-point scale (poor to outstanding) for each of five traits: conventions, sentence fluency, word choice, organization, and ideas.

Results from Troia et al. showed similar findings to previous studies (Farrington et al., 2014; Hogston & Schrader, 2012; Jewell & Malecki, 2005; Malecki & Jewell, 2003), where female participants and older students wrote more than male and younger participants. Students rated with higher levels of writing ability, based on teacher judgment, also wrote better quality stories. Findings showed that gender, teacher judgment, and writing activity (i.e., WAMS) directly correlated with features of writing motivation. Furthermore, writing motivation (e.g., self-efficacy beliefs, interest, perceived task value, attitudes, goal orientations, and attributions for success and failure) were found to have a direct effect on the student's writing performance. When students had higher scores on the scales related to writing motivation on the WAMS (e.g., interest in a topic), students' writing performance (e.g., amount written, quality) increased.

Benton, Corkill, Sharp, Downey, and Khramtsova (1995) conducted a study which examined the relationship between knowledge, interest, and narrative writing.

CBM-WE was not used as a measure. Participants included 106 ninth graders and 203 undergraduate college students. Participants wrote a story for 25 minutes about baseball. Afterwards, participants completed a 39-item, multiple-choice test about specific baseball topic knowledge and answered six interest questions about baseball. A hierarchical regression analysis was conducted to examine the results. Topic knowledge was a better predictor of interest than was the participants' ratings of their individual interest. Female participants, those with higher interest in baseball, and undergraduates showed significantly higher quality narrative writings. Overall, findings indicate that topic knowledge and writing motivation (e.g., interest) highly correlate with writing performance.

Albin, Benton, and Khrantsova (1996) conducted a follow-up study to the Benton et al. (1995) study and found a similar finding - interest facilitates writing performance. This study examined if differences in interest between two topics were related to narrative writing for undergraduate college students. Albin et al. reviewed and cited past research that concluded interest in a topic guides attention and enhances learning. Participants included 224 undergraduate students. Participants wrote about two topics, a baseball game and a soccer game, for 20 minutes each. Afterwards, participants completed two interest inventories, six items each, on baseball and soccer. Topic knowledge tests on baseball and soccer knowledge were also completed. A hierarchical regression analysis was conducted to examine the results. Results showed that baseball was an overall higher interest topic than soccer, and that participants wrote more topic-relevant information on the higher interest topic (i.e., baseball). When interest was

higher, participants also wrote more about game actions (e.g., hitting a home run) as compared to irrelevant-nongame actions (e.g., people observing the game).

Purpose of the Present Study

Fuchs (2004) discussed past and present research about CBM and how there are many more changes and studies needed in the future to make CBM even better.

Compared to the hundreds of studies examining CBM in the area of reading, relatively few research studies have examined CBM-WE. The few research studies that exist have examined CBM-WE validity and reliability characteristics, as well as differences between genders and grade levels. However, evaluations of aspects of the story starter used in CBM-WE are rare. Sources such as AIMSweb (2008) and Hosp et al. (2007) provide lists of story starters, classified from primary to upper level grade levels, but they do not describe their criteria for including or excluding story starters, nor how it was determined at what grade level to place a story starter.

More importantly, very few research studies have evaluated the effect students' choice and interest in the story starters have on their writing performances. One study that did examine the effects of choice versus no choice in story starters with a sample of ninth grade students generally did not find statistically significant increases on CBM-WE measures when students were given a choice (Bleck, 2013). However, Bleck did find a gender and choice interaction effect for one of her four samples where boys did write at a comparable level to girls.

The impact of different story starters on students' writing performance is a topic in need of more research. While Bleck's (2013) research evaluated ninth grade students, it is unknown what effect choice may have on students' writing at the elementary level.

Furthermore, while Bleck gave a choice of story starters, she did not assess students' interest level in the story starter topics. It is possible that even though the students had a choice of story starters, some or many of the students had little interest in the topics. A limited interest in the topics could account for the inconsistent results Bleck reported. It is important for school personnel to know if differences in story starters can impact students' writing performance in order to know if schools should use choice and/or topic specific story starters for assessing writing.

The present study examined the effects of choice and students' interest in a story starter on elementary students' writing performance. Specifically, this study evaluated the differences in boys' and girls' production-dependent scores (i.e., TWW, WSC, and CWS), production-independent scores (i.e., %CWS), and accurate-production scores (i.e., CMIWS) for a sample of third grade students. The following research questions were addressed:

Research Question 1: Does choice of a story starter have an effect on third grade students' writing performance? The current study provided the students a choice between six gender-specific and gender-neutral story starters. It was anticipated, that when given a choice of story starters, students would pick one of high interest. Given previous research on gender differences (e.g., Jewell & Malecki, 2005; Malecki & Jewell, 2003), it was hypothesized there would be a gender effect with girls performing significantly better than boys, regardless of choice of story starter.

Research Question 2: Does interest level in a story starter topic have an effect on third grade students' writing performance? Students were given a story

starter and asked to rate their interest level on the topic. Based on previously reviewed research (Albin et al., 1996; Benton et al., 1995), it was predicted that a higher level of interest in the story starter would result in significantly higher CBM-WE scores for both boys and girls, with girls scoring higher than boys.

Method

Participants

The participants in this study were obtained from the author's school psychology practicum site. Participants included third grade students from an elementary school in a south central Kentucky district. All 214 third grade students from the school district were invited to participate in the study. Two weeks before the CBM-WE probes were administered, the students' parents were sent an opt-out consent form. Thus, parents had three days to sign the form and return it to the school if they did not want their children to participate in the study, but were not collected until two weeks had passed to give extra time. Prior to the CBM-WE administration, students were asked to give signed assent to participate in the study. The consent and assent forms, which indicate that Western Kentucky University's Institutional Review Board approved this research, can be found in Appendix A. Only eight of the nine classes of students were assessed. Besides the one class of students not assessed, which comprised 24 students, additional students were not included in the study due to a lack of parent permission or student assent, as well as absences from school on the day of the assessments. A final sample of 159 third grade students (77 boys, 82 girls) participated in the assessment, which is 74.3% of the third grade students in that district. Specific demographic data on the participants were not collected. However, the school district is a rural one, with 93.9% of its students White, 3.1% Hispanic, 1.0% Black, and the rest other or multiple ethnicities. Many of the district's students (62.1%) are on free or reduced lunch.

Materials

For the part of the study addressing the first research question on the effects of choice, six CBM-WE story starters appropriate for a third grade level were selected from AIMSweb (2008). The six story starters were selected based on likely interest for the students. Based on the author's and specialist project advisor's clinical judgment, two story starters were chosen as likely to appeal to boys, two for girls, and two were judged to be gender-neutral topics. Gender-neutral story starters were determined by finding topics not specifically related to gender stereotypes and that both genders would likely enjoy.

For the second research question on the effects of interest, one CBM-WE story starter was created. A story starter was created using guidelines for writing CBM-WE story starters from AIMSweb (2008). One was created, rather than chosen from AIMSweb, with the intent to have a story starter that would be less appealing to students of that age. The story starter topic was on cleaning one's room. That topic was chosen after two third grade teachers discussed various topics with the students in their classrooms that would be uninteresting for a story. Of the options generated by the students, the students ranked the topics of cleaning one's room or doing chores as low. One teacher was from the district where the participants were chosen and the other teacher was from another school district.

A sheet of lined paper with six story starters printed on it was used for the first research question related to choice. The back of that sheet also consisted of lined paper with a single story starter on it, and was used for the research question addressing interest level. Those sheets containing the CBM-WE probes used in this study can be found in

Appendix B. For both probes, a Likert scale with a 4-point rating scale was on the bottom for the students to rate how interesting they found the story starter after they wrote the story. Specifically, this scale included the ratings: 1 - not interesting at all, 2 - not very interesting, 3 - a little interesting, and 4 - very interesting. At the top of each sheet of paper, the students were to indicate their gender. Names were not requested to ensure anonymity of the data.

Procedure

Opt-out consent forms were given to homeroom class teachers of all third grade students to send home with each student. Parents were given two weeks after the consent form to return the signed form, indicating they did not give permission for their child to participate in this study. Teachers collected the opt-out consent forms as they were returned. Once it was determined who could be assessed, that group of students in each homeroom class was given assent forms on the day CBM-WE probes were administered.

The author of this Specialist Project administered all of the CBM-WE probes to eight of the nine third grade homeroom classrooms. Only eight of the nine classes were used to keep the number of students who initially received each type of probe even. That is, half of the classes were administered the CBM-WE choice probe first, followed by the probe with the single story starter, and the other half of the classes received the probe with the single story starter first. The two probes were administered consecutively to each class of students. Probes were administered over two consecutive days at the end of the school year.

CBM-WE probes were handed out with the probe to be completed first facing up. Students were first instructed to circle, at the top, whether they were a boy or girl. Non-

participants, who did not have consent or did not assent, sat quietly doing something else (e.g., reading). Then, the examiner read the standardized directions to the participants. The examiner used the following standardized administration directions with each class based on Hosp et al. (2007). The directions for the probe with one story starter were exactly as Hosp et al. listed. Wording that was added or changed has been bolded for the probe with story starter choices.

Research question one (choices):

- Say: “Today I want you to write a story. **I am going to read you six sentences first and then I want you to choose the one you find MOST interesting and circle the number that goes with it. After you have selected the sentence you find the most interesting,** I want you to compose a short story about what happens. You will have **one minute to think about what sentence you find most interesting,** one minute to think about what you will write and three minutes to write your story. Remember to do your best work. If you do not know how to spell a word, you should guess. Are there any questions?” (Pause for questions.) “Put your pencils down and listen. For the next minute, think about... **[insert the 6 story starters] and which one you find MOST interesting.**”

- **After reading the six story starters, begin your stopwatch and allow one minute for the students to think about which one they find most interesting. Monitor students so that they do not begin writing. After 30 seconds say, “You should be thinking about which sentence you find MOST interesting.”**

- **At the end of one minute, restart your stopwatch and** allow one minute for the students to think. Say, “For the next minute, **think about which sentence you chose and** what you will write.” Monitor students so that they do not begin writing. After 30 seconds say, “You should be thinking about **which sentence you chose to write about.**” At the end of one minute, restart your stopwatch for three minutes and say, “Now begin writing.”

- Monitor students’ attention to the task. Encourage the students to work if they are not writing.

-After 90 seconds say, “You should be writing about **the sentence you chose.**”

-At the end of three minutes say, “Thank you. Put your pencils down.” (Hosp et al., 2007, p. 88)

Research question two (interest level):

- Say: “Today I want you to write a story. I am going to read you a sentence first and then I want you to compose a short story about what happens. You will have one minute to think about what you will write and three minutes to write your story. Remember to do your best work. If you do not know how to spell a word, you should guess. Are there any questions?” (Pause for questions.) “Put your pencils down and listen. For the next minute, think about... [insert story starter].”

- After reading the story starter, begin your stopwatch and allow one minute for the students to think. Monitor students so that they do not begin writing. After 30 seconds say, “You should be thinking about...(insert story starter).” At the

end of one minute, restart your stopwatch for three minutes and say, “Now begin writing.”

- Monitor students’ attention to the task. Encourage the students to work if they are not writing.

-After 90 seconds say, “You should be writing about...(insert story starter).”

-At the end of three minutes say, “Thank you. Put your pencils down.” (Hosp et al., 2007, p. 88)

After being told to stop writing, students were instructed to indicate their interest level in the story starter topic. The second administration was then conducted, following the same administration procedures. When both probes were completed, the examiner immediately collected them. CBM-WE probes were kept in class groups to know which probe was administered first.

The writing probes were scored at a later time by the examiner. The author of this specialist project was previously trained on the standardized administration and scoring procedures for CBM-WE and has been involved in previous research projects on CBM-WE. A second graduate student in the school psychology program, who was also trained on the standardized scoring procedures for CBM-WE, provided inter-rater agreement ratings.

The writing probes were scored using the standard production-dependent indices of Total Words Written (TWW), Words Spelled Correctly (WSC), and Correct Word Sequence); the production-independent index of Percent Correct Word Sequence (%CWS); and the accurate production index of Correct Minus Incorrect Word Sequences (CMIWS).

After all of the writing probes were scored by the examiner, 20% of the probes from each classroom were randomly selected and re-scored by the second graduate student for inter-rater agreement purposes. McMaster and Espin (2007) noted that a minimum of 80% inter-rater agreement is recommended. Overall, inter-rater agreement for TWW was 94.5%, ranging from 92% to 100% per classroom. For WSC, it was 94.8%, ranging from 90% to 100%; and for CWS the average was 93.8, ranging from 88% to 100%. None of the individual classrooms' inter-rater agreements fell below the recommended 80%, so the inter-rater agreements were considered at an acceptable level and suggested accurate scoring of the probes. Where differences in scores did occur, the probes were discussed between the two raters and a determination was made for the correct score to be recorded.

Results

Half of the classes received the choice story starter probe first and the other half received the non-choice story starter first. To evaluate if the order of the two probes had an impact on students' writing, independent sample *t*-tests were conducted. The *t*-tests indicated that it did not matter if the non-choice story starter was administered first or second. However, when choice in story starter is given after a non-choice story starter, the students scored significantly higher on all measures except %CWS. Specific results for the significant findings were: TWW $t(157) = -3.01, p = .003$; WSC $t(157) = -2.86, p = .005$; CWS $t(157) = -2.70, p = .008$; CMIWS $t(157) = -2.05, p = .042$.

The first research question addressed whether or not choice of a story starter would have an effect on third grade students' writing performance. It was hypothesized that girls would score higher than boys, regardless of choice of story starter. There were two story starters that were judged to be masculine, two judged to be feminine, and two gender-neutral. The number of boys and girls picking each of the story starters is presented in Table 1. The results showed that when given a choice, more than half (53.7%) of the boys picked a masculine story starter and only 8.5% picked a feminine story starter. Almost twice as many girls picked a feminine story starter over a masculine one; however, girls most frequently picked a gender-neutral story starter (46.8%). The gender-neutral story starters were popular with the boys as well (37.8%).

To verify the students picked a story starter that was of high interest, students were asked to rate their interest in the story starters they picked. All but six students rated the story starter they picked as a "3" (a little interesting) or "4" (very interesting) on the 4-point Likert scale. Given this analysis presumes the students are interested in the

Table 1

Frequency of Gender-Specific and Gender-Neutral Story Starters Chosen by Participants

<u>Story Starter</u> <u>Type/Number</u>	<u>Boys (n = 82)</u>		<u>Girls (n = 77)</u>		<u>Total (n = 159)</u>	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Masculine						
2	36	43.9	10	13.0	46	28.9
5	8	9.8	4	5.2	12	7.5
Total	44	53.7	14	18.2	58	36.5
Feminine						
3	4	4.9	12	15.6	16	10.1
6	3	3.6	15	19.5	18	11.3
Total	7	8.5	27	35.1	34	21.4
Gender-Neutral						
1	11	13.4	8	10.4	19	11.9
4	20	24.4	28	36.4	48	30.2
Total	31	37.8	36	46.8	67	42.1

Note. Story starters for each number can be found in Appendix B.

topics they picked, the six students (five boys, one girl) that rated their interest as a “1” (not interesting at all) or a “2” (not very interesting) were eliminated from further data analyses. For descriptive purposes, Table 2 provides the mean scores for all the CBM-WE scoring measures for boys and girls based on whether the story starter was gender-specific or gender-neutral.

As can be seen in Table 2, for every type of gender-specific and gender-neutral story starter and across every CBM-WE scoring method, girls’ mean scores were higher than the boys’ mean scores. To determine if those differences were statistically significant, as well as to assess differences in the type of story starter and interaction effects, a two-way ANOVA was conducted. The use of the standard p value of $\leq .05$ was the criterion for statistical significance. The F values and p values from that two-way ANOVA are reported in Table 3. The main effect examining the interaction between gender and type of story starter was not significant for any type of CBM-WE scoring method. Contrary to the hypothesis that expected girls to score significantly higher than boys, there were no significant differences on any of the scoring methods between boys and girls when the students were given a choice of story starters.

A significant difference based on the type of story starter was found only for the scoring method of WSC. A Tukey HSD post hoc analysis indicated the significant difference was between the masculine and gender-neutral probes, with both boys and girls having higher WSC scores on the masculine probes than the gender-neutral probes. However, the p value on the Tukey HSD was .058, a non-significant level. Thus, further post-hoc analyses were conducted separately for boys and girls. An independent samples t -test indicated the difference between the masculine and gender-neutral probes for girls

Table 2

Mean Scores for Boys and Girls with Gender-Specific and Gender-Neutral Story Starters

	Type of Story Starter		
	Masculine	Feminine	Neutral
Boys (<i>n</i> = 77)			
TWW	36.4	35.6	29.5
WSC	32.4	31.1	25.6
CWS	29.6	28.6	22.8
%CWS	75.8	72.3	72.3
CMIWS	22.6	19.1	15.1
Girls (<i>n</i> = 76)			
TWW	37.9	36.8	34.5
WSC	34.8	32.8	29.9
CWS	31.5	29.1	28.9
%CWS	75.9	74.3	80.7
CMIWS	23.6	20.2	22.3

Note. TWW = Total Words Written, WSC = Words Spelled Correctly, CWS = Correct Words Sequence, %CWS = Percent Correct Word Sequence, and CMIWS = Correct Minus Incorrect Word Sequence.

Table 3

Two-Way ANOVA Results for Third Grade Students' Choice of Story Starter (n = 153)

	<i>F</i> values	<i>p</i> values
TWW		
Gender	1.179	.279
Story Starter Type	2.491	.086
Interaction	0.348	.706
WSC		
Gender	1.466	.228
Story Starter Type	3.289	.040*
Interaction	0.145	.865
CWS		
Gender	1.578	.211
Story Starter Type	2.130	.123
Interaction	0.650	.524
%CWS		
Gender	0.532	.467
Story Starter Type	0.603	.549
Interaction	0.290	.749
CMIWS		
Gender	1.258	.264
Story Starter Type	1.233	.294
Interaction	0.756	.471

* $p < .05$.

Note. Story Starter Types: Masculine, Feminine, and Neutral. TWW = Total Words Written, WSC = Words Spelled Correctly, CWS = Correct Words Sequence, %CWS = Percent Correct Word Sequence, and CMIWS = Correct Minus Incorrect Word Sequence.

was not significant, $t(47) = 1.408, p = .166$. However, the difference for those two types of probes for boys was significant, $t(68) = 2.373, p = .020$.

The second research question evaluated the effects of students' interest in a story starter topic on writing performance. One story starter was provided to the students and they also indicated their level of interest in the story starter. It was hypothesized that a higher level of interest in the story starter would result in significantly higher CBM-WE scores and that girls would score higher than boys. For descriptive purposes, Table 4 presents how many boys and girls rated the story starter at each level of interest (i.e., 1 through 4) and their mean scores on each of the CBM-WE scoring method. In general, higher interest in the story starter topic resulted in higher scores and girls usually scored higher than boys at each level.

To test for significance, a two-way ANOVA was used to evaluate the main interaction effect (gender x interest level) and the simple effects of gender and interest level. The F values and p values from the two-way ANOVA are presented in Table 5. None of the main interaction effects were significant for any of the CBM-WE scoring methods. There were no simple effects for gender or interest level for the production-independent scoring method of %CWS. However, for all other scoring methods, there was both a significant effect for gender (i.e., girls' scores were higher than boys) and for level of interest in the story (i.e., higher interest resulted in higher scores).

Table 4

Mean Scores per Level of Interest in the Story Starter by Gender

Interest Level	Means				
	TWW	WSC	CWS	%CWS	CMIWS
Boys					
1 (<i>n</i> = 18)	21.33	18.11	15.50	71.44	8.89
2 (<i>n</i> = 10)	29.50	26.00	24.70	75.70	20.10
3 (<i>n</i> = 29)	34.48	32.10	30.41	81.34	24.90
4 (<i>n</i> = 25)	39.16	34.44	30.80	76.64	20.84
Girls					
1 (<i>n</i> = 5)	26.40	25.40	24.00	71.40	20.60
2 (<i>n</i> = 14)	38.93	37.50	36.71	90.57	33.29
3 (<i>n</i> = 33)	38.00	35.27	33.12	84.70	26.85
4 (<i>n</i> = 25)	44.88	41.12	38.64	78.20	30.36

Note. Numbers 1 - 4 are the interest level ratings: 1 = not interesting at all, 2 = not very interesting, 3 = a little interesting, and 4 = very interesting. TWW = Total Words Written, WSC = Words Spelled Correctly, CWS = Correct Words Sequence, %CWS = Percent Correct Word Sequence, and CMIWS = Correct Minus Incorrect Word Sequence.

Table 5

Two-Way ANOVA Results for Third Grade Students' Interest Level of Story Starter (n = 159)

	<i>F</i> values	<i>p</i> values
TWW		
Gender	5.615	.019*
Interest Level	7.880	.000***
Interaction	0.287	.835
WSC		
Gender	8.421	.004**
Interest Level	6.269	.000***
Interaction	0.587	.624
CWS		
Gender	10.165	.002**
Interest Level	5.410	.001**
Interaction	0.878	.454
%CWS		
Gender	1.989	.160
Interest Level	2.216	.089
Interaction	0.811	.490
CMIWS		
Gender	11.150	.001**
Interest Level	2.842	.040*
Interaction	1.255	.292

* $p < .05$. ** $p < .01$. *** $p < .001$.

Note. TWW = Total Words Written, WSC = Words Spelled Correctly, CWS = Correct Words Sequence, %CWS = Percent Correct Word Sequence, and CMIWS = Correct Minus Incorrect Word Sequence.

Discussion

Curriculum-Based Measurement (CBM) consists of brief fluency measures that are considered valid indicators of students' performance in basic academic skills (i.e., reading, mathematics, spelling, and written expression). CBM has been demonstrated to be useful for numerous activities, such as screening to identify at risk students and progress monitoring students' growth. The current study adds to the literature on CBM-WE by evaluating aspects of the story starter, a research topic that has rarely been addressed. Specifically, this study evaluated how providing a choice of story starters and the interest level in the story starter topic impacts boys' and girls' writing performance. This study investigated the differences in 159 third grade students' writing performance using the standard production-dependent indices (i.e., TWW, WSC, and CWS), a production-independent index (i.e., %CWS), and an accurate production index (i.e., CMIWS).

In reference to the first research question about whether or not choice would have an impact on student's writing performance, it was correctly anticipated that when given a choice, both boys and girls picked a story starter topic of high interest. Only six of the 159 students picked a story starter that they perceived to be of low interest. It was hypothesized that girls would outperform the boys, regardless of choice in story starter. Results showed that there were no significant differences between boys and girls when given choice of story starters and there were no significant interaction effects (story starter type x gender). When given a choice, boys scored at a comparable level to girls. The finding that there was not a significant difference between boys' and girls' writing performance is not consistent with previous research findings (Fearington et al., 2014;

Hogston & Schrader, 2012; Jewell & Malecki, 2005; Malecki & Jewell, 2003; Melloy, 2012). Indeed, most research on writing performance in general shows that girls outperform boys (Benton et al., 1995; Troia et al., 2013). Bleck (2013), in one of her four administrations of choice/no choice CBM-WE probes with ninth graders, also found girls' and boys' scores were comparable when given a choice. The current finding that choice results in comparable scores between boys and girls is remarkable and is an important addition to the research literature on CBM-WE.

Generally, there were no significant differences in the type of story starter chosen by the students (i.e., gender-specific or gender-neutral). The only statistically significant difference was for the CBM-WE scoring method of WSC, where boys who picked a masculine story starter scored higher than when they picked a gender-neutral story starter. It is difficult to know how important this finding is, given that none of the other CBM-WE scoring methods resulted in significant differences. However, the implication that the type of story starter chosen by boy's impacts their writing performance is noteworthy, given it is consistent with the finding that boys' and girls' writing performance is comparable when given a choice of story starters.

When examining the second research question, it was hypothesized that both boys and girls who perceive higher interest in the story starter will write significantly more than students who perceive lower interest in the story starter. It was also predicted that the girls would outperform the boys, regardless of perceived interest in the story starter. Results indicate no simple effects for gender or interest level for the production-independent scoring method of %CWS. However, for all other scoring methods (i.e., TWW, WSC, CWS, CMIWS), there was both a significant effect for gender (i.e., girls'

scores were higher than boys) and for level of interest in the story (i.e., higher interest resulted in higher scores). There were no interaction effects (gender x interest level) for any of the CBM-WE scoring methods. The finding that girls outperformed boys on multiple scoring indices is consistent with previous findings (Farrington et al., 2014; Hogston & Schrader, 2012; Jewell & Malecki, 2005; Malecki & Jewell, 2003; Melloy, 2012). The finding that higher interest resulted in higher scores is consistent with Albin et al. (1996) and Benton et al. (1995).

The current study is an important addition to the current literature on CBM-WE regarding the effects of choice, interest, and gender on third grade students' writing performance. Results from this Specialist Project indicate that when no choice is given, girls outperform boys in writing performance, which is consistent with previous findings (Jewell & Malecki, 2005; Malecki & Jewell, 2003). But, when given a choice of story starter, boys' scores are comparable to girls' and no gender differences are found for any of the CBM-WE scoring methods. These results are imperative for school personnel to know when using CBM-WE. Previous research findings on the CBM-WE differences between boys and girls resulted in calls for the development of gender-specific CBM-WE norms (Jewell & Malecki, 2005; Malecki & Jewell, 2003). Perhaps by giving choices in story starters, separate norms are not necessary.

Strengths and Limitations

A strength of the current study is that it examines gender differences in writing performance with choice of writing topics. Most previous research examining CBM-WE gender differences did not include choice (Jewell & Malecki, 2005; Malecki & Jewell, 2003). Bleck (2013) is the only study found that examined writing performance

differences for CBM-WE when students were given a choice in story starters. The current study adds to the research on the effects of choice on writing performance. A second area of strength for this study is the methodology. Prior to data collection, two third grade teachers assessed the interest level of story starters to find one that would result in varied interest levels. Students' interest in the story starter was assessed for both research questions, which is a unique addition to CBM-WE research. This allowed a quantifiable measure of interest from the students for each story starter. Finally, inter-rater agreement was measured and was above the acceptable minimal level (i.e., 80%), ensuring accuracy in the obtained scores.

A limitation of the present study is generalizability, which is a common issue for most studies. Participants only included third grade students from one rural school district, and the student population had very limited diversity. Results may not generalize to other school districts or other grade levels. A second limitation is the small sample size ($n = 159$), also affecting the generalizability. While the overall sample size was acceptable, when subgroups were created (i.e., interest rating by gender), the group sizes ranged from five to 33. The subgroups for choice of story starter by gender ranged from seven to 44. A third limitation is that the selection of the story starters could have been more systematic. That is, the story starters were judged to be gender-specific or gender-neutral, but were not empirically validated as such. Furthermore, the story starters varied in terms of length. It is unknown if story starters of different length have an effect on students' writing.

Future Research

This study only examines choice and perceived interest of story starters for a sample of third grade students. Additional research with larger sample sizes (to obtain large subgroups) and other grade levels would add to the generalizability of these important results.

The finding that students produced significantly more when given a choice after receiving a non-choice probe is a curious result that needs replication. If these results are replicated, future research should investigate reasons for such a finding.

Another possible factor to examine would be students' familiarity with the story starter topic. Interest versus familiarity in story starters was not examined for this study. It is unclear whether familiarity of story starter topic plays a role in writing performance. Teasing out the differences between interest and familiarity, if any, would help determine if the current results were truly based on interest or if familiarity in the story starter topic accounts for the findings.

Future research could also examine the length, word frequency, or word difficulty of the story starter to determine if any of these factors would have an effect on students' writing performance. Examining such differences would help determine what types of story starters to use with different grade levels of students.

REFERENCES

- AIMSweb. (2008). *AIMSweb progress monitoring and RTI system*. Retrieved November 6, 2012, from <http://www.aimsweb.com/>
- Albin, M. L., Benton, S. L., & Khrantsova, I. (1996). Individual differences in interest and narrative writing. *Contemporary Educational Psychology, 21*, 305–324.
- Amato, J. W., & Watkins, M. W. (2011). The predictive validity of CBM writing indices for eighth-grade students. *Journal of Special Education, 44*, 195-204.
- Benton, S. L., Corkill, A. J., Sharp, J., Downey, R., & Khrantsova, I. (1995). Knowledge, interest, and narrative writing. *Journal of Educational Psychology, 87*, 66–79.
- Bleck, A. A. (2013). Examining the impact of student choice of writing topic on writing production and accuracy. *Dissertation Abstracts International Section A, 74*, 1-114.
- Bruton, A., Conway, J., & Holgate, S. (2000). Reliability: What is it, and how is it measured? *Physiotherapy, 86*, 94-99.
- Coker, D., & Ritchey, K. D. (2010). Curriculum-based measurement of writing in kindergarten and first grade: An investigation of production and qualitative scores. *Exceptional Children, 76*, 175-193.
- Cusumano, D. (2007). Is it working?: An overview of curriculum-based measurement and its uses for assessing instructional, intervention, or program effectiveness. *Behavior Analyst Today, 8*, 24-34.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *Journal of Special Education, 37*, 184-192.

- Espin, C., Wallace, T., Campbell, H., Lembke, E. S., Long, J. D., & Ticha, R. (2008). Curriculum-based measurement in writing: Predicting the success of high-school students on state standards tests. *Exceptional Children, 74*, 174-193.
- Ferrington, J. A., Parker, P. D., Kidder-Ashley, P., Gagnon, S. G., Mcane-Bowling, S., & Sorrell, C. A. (2014). Gender difference in written expression curriculum-based measurement in third- through eighth-grade students. *Psychology in the Schools, 51*, 85-96.
- Fewster, S., & Macmillan, P. D. (2002). School-based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial & Special Education, 23*, 149-156.
- Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*, 188-192.
- 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*, 477-497.
- Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Slider, N. J., Hoffpauir, L. D., & Whitmarsh, E. L. (2004). An examination of the criterion validity and sensitivity to brief intervention of alternate curriculum-based measures of writing skill. *Psychology in the Schools, 41*, 291-300.
- Goo, M., Watt, S., Park, Y., & Hosp, J. (2012). A guide to choosing web-based curriculum-based measurements for the classroom. *Teaching Exceptional Children, 45*, 34-40.

- Hessler, T., & Konrad, M. (2008). Using curriculum-based measurement to drive IEPs and instruction in written expression. *Teaching Exceptional Children, 41*, 28-37.
- Hogston, E., & Schrader, M. (2012, September). *Various scoring methods and gender differences on CBM-Written Expression*. Poster presented at the annual meeting of the Kentucky Association for Psychology in the Schools, Louisville, KY.
- Hosp, M., Hosp, J., & Howell, K. (2007). *The ABC's of CBM. A practical guide to curriculum-based measurement*. New York, NY: Guilford Press.
- Jewell, J., & Malecki, C. K. (2005). The utility of CBM written language indices: An investigation of production-dependent, production-independent, and accurate production scores. *School Psychology Review, 34*, 27-44.
- Malecki, C. K., & Jewell, J. (2003). Development, gender, and practical considerations in scoring curriculum-based measurement writing probes. *Psychology in the Schools, 40*, 379-390.
- McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing: A literature review. *Journal of Special Education, 41*, 68-84.
- McMaster, K. L., Xiaoqing, D., Parker, D. C., & Pinto, V. (2011). Using curriculum-based measurement for struggling beginning writers. *Teaching Exceptional Children, 44*, 26-34.
- Melloy, A. (2012). *Gender differences in written expression at the elementary level*. (Unpublished specialist project). Western Kentucky University, Bowling Green, KY.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York, NY: Guilford Press.

Thorndike, R. M. (2005), *Measurement and evaluation in psychology and education* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Troia, G. A., Harbaugh, A. G., Shankland, R. K., Wolbers, K. A., & Lawrence, A. M. (2013). Relationships between writing motivation, writing activity, and writing performance: Effects of grade, sex, and ability. *Reading and Writing: An Interdisciplinary Journal*, 26, 17-44.

Youngman, E. A. (2010). *The development of curriculum-based measurement local norms in the area of written expression*. (Unpublished specialist project). Western Kentucky University, Bowling Green, KY.

APPENDIX A: Consent and Assent Forms

Allen County Schools

570 Oliver Street, Scottsville, KY 42164
Phone (270) 618-3181 Fax (270) 618-3185

PARENT/GUARDIAN PERMISSION FORM

Your child is invited to participate in a research study conducted by Morgan Schrader, Practicum Student in School Psychology in the Allen County School District, and Dr. Carl Myers, from Western Kentucky University. The collection of data will contribute to completion of Morgan's research thesis. Morgan is examining brief assessments of students writing skills across two grade levels. Many participants from the grades 3rd and 5th are needed for this study. Given that your child is in the required grade, your child has been selected as a possible participant in this study.

If your child participates, he or she will be asked to write for 3 minutes, two different times, in one day. Students will be asked to write about an interesting topic and an uninteresting topic.

No names will be collected for this study. No name or code will be needed to keep the writings together, since the interesting topic will be on the front page and the uninteresting topic will be on the back. There will be no identifying information linking your child to their writing; all that is asked is the child's gender. All writings collected in this study will be kept strictly confidential and is accessible only to the project staff. Data will be kept within a locked file cabinet in Dr. Myer's office. Only an overall summary of the results will be written up for Morgan's thesis, not individual results.

Your child's participation is voluntary. Your decision whether or not to let your child participate will not affect your relationship with school personnel. If you do decide to allow your child to participate, you are free to withdraw your consent and discontinue your child's participation before the administration begins. After the collection of the writing samples, there is no way to know which writing sample is your child's. If you or your child decides not to participate, it will have no negative outcome for you or your child in any way. Given the task is a commonly expected one in a school setting and is brief (3 minutes), we anticipate no discomfort, or risks as a result of your child's participation in this study.

Complete and return this form to your child's teacher or school office within 3 days of receiving this letter only if you **DO NOT** want your child to participate in the collection of data. A request to have your student opt out of participation will become effective on the day it is received by the school.

I request that my child _____, in grade _____, **NOT**
Participate in the collection of data.
Signature of parent or guardian: _____
Date: _____
Student's name: _____
(please print)
Student's school: _____
(please print)

If you have any questions, please feel free to contact Morgan Schrader, 270-945-8989, or Dr. Carl Myers, 270-745-4410, School Psychology Program Director at Western Kentucky University, Thesis Advisor.



THE DATED APPROVAL ON THIS CONSENT FORM INDICATES THAT
THIS PROJECT HAS BEEN REVIEWED AND APPROVED BY
THE WESTERN KENTUCKY UNIVERSITY INSTITUTIONAL REVIEW BOARD
Paul Mooney, Human Protections Administrator
TELEPHONE: (270) 745-2129

WKU IRB# 14-312
Approval - 5/5/2014
End Date - 10/1/2014
Full Board
Original - 3/31/2014

Allen County Schools

570 Oliver Street
Scottsville, KY 42164
Phone (270) 618-3181
Fax (270) 618-3185

I, _____, understand that my parents/guardians have said its okay for me to take part in a project about writing under the direction of Morgan Schrader.

I am taking part because I want to. I have been told that I can stop at any time I want to. If I do not want to or if I stop at any time, I know that nothing will happen to me and it will not affect my grade.

Signature _____ Date _____

THE DATED APPROVAL ON THIS CONSENT FORM INDICATES THAT
THIS PROJECT HAS BEEN REVIEWED AND APPROVED BY
THE WESTERN KENTUCKY UNIVERSITY INSTITUTIONAL REVIEW BOARD
Paul Mooney, Human Protections Administrator
TELEPHONE: (270) 745-2129



WKU IRB# 14-312
Approval - 5/5/2014
End Date - 10/1/2014
Full Board
Original - 3/31/2014

