# Examining the Potential Use of Instructionally-Relevant Assessment Data in Early Writing

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

David Charles Parker

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Matthew K. Burns, Advisor

June, 2012

Acknowledgements

I had a brief conversation with Jim Ysseldyke when I was first considering the University of Minnesota for graduate school. I met him in his office, he asked me to explain my interests, and moments later he was walking me down to Matt Burns's office so Matt and I could talk. I remain incredibly thankful for that conversation, and for the opportunity to observe and learn from Jim's commitment to improving education.

Shortly into my first year of studies, I sought out a way to meaningfully apply the skills and knowledge I was learning. I met Jennifer McComas within a few weeks, and four years later I was still applying those skills in high-needs schools Jennifer worked tirelessly to support. I learned many valuable lessons each of the years I worked with Jennifer, including how generous and kind she is.

One summer when I was working on a project with Kristen McMaster, I rushed into her office feeling distraught over what I thought was a vital research question I had failed to address with the data. She patiently explained that my perceived oversight was actually a common experience when conducting research. She helped me realize that research is somewhat like exploring a cave: you set out to bring something new to light, but it inevitably leads to the realization that more remains to be discovered.

Matt and I talked about two things in our first follow-up conversation to the one Jim arranged. We discussed his vision for conducting applied research that advances theory and practice, and we talked about how much he values relationships with his students. Those two topics of conversation foretold a truly fortunate education under an excellent adviser, for which I cannot express enough thanks.

I would also like to thank several others. First, Stephanie Al Otaiba and her colleagues at the Florida Center for Reading Research, along with the Society for the Study of School Psychology, helped me collect and analyze the data in this dissertation. The study would not have happened without them. Second, I want to thank my fellow graduate students, especially Amanuel Medhanie and my incredible cohort. The relationships and camaraderie we developed will not end with graduate school. Third, I want to offer a special thank you to Sarah Scullin Stokes for co-authoring an important chapter in my life. Finally, I wish to thank my entire family. My wife Katie and I are blessed to share their love and support.

## Dedication

This dissertation is dedicated to all children, whose natural joy for life and learning is a constant source of inspiration.

## Abstract

The writing performance of many school-aged children is consistently below levels necessary to produce positive outcomes. Early intervention frameworks are designed to help remedy this problem before it becomes more severe, but a key feature of early intervention frameworks—the use of data for targeting interventions—has not been sufficiently researched. This study examined the role of data from curriculum-based assessment for instructional design (CBA-ID) for targeting interventions in early writing.

Participants included 147 first grade students from a larger study conducted in the Southeastern United States. Each student was identified as struggling in early literacy using measures of beginning reading. Students were placed into a supplementary writing intervention with tiers of support that were more (4 times per week for 45 min) or less (2 times per week for 30 min) intensive depending on screening scores. Students responded to early writing progress assessments during intervention. The resulting data were used to derive novel CBA-ID criteria and for longitudinal analyses that investigated students' growth patterns. Additional analyses examined the importance of data from CBA-ID, intervention intensity, and other demographic factors for predicting growth patterns of students.

Results showed that the current CBA-ID criteria (i.e., 11-17 correct word sequences on an early writing measure) overlapped some, but not completely, with previous derivations of CBA-ID criteria (i.e., 8-14 correct word sequences). Results also suggested that students followed more than one type of growth pattern, and a solution with three separate patterns was identified based on empirical data and interpretability. The results from additional analyses suggested that CBA-ID data, in addition to gender, were significant predictors of growth pattern; but intervention intensity did not predict growth patterns.

The current results were contextualized within previous research, and implications for practice and theoretical models of writing development were discussed. Specifically, the results of the study were discussed in terms of their contribution to (1) the role of instructionally-relevant assessment data in targeting interventions within an early intervention framework, and (2) the understanding of how early writing skills develop independently and in concert. Lastly, limitations to the study and future directions for early intervention research were outlined.

# Table of Contents

## List of Tables

## List of Figures

# CHAPTER 1

# INTRODUCTION

In 2009, a project of the National Commission on Writing (2009) requested students to send letters to the President. Over 6,500 students responded, expressing concerns from the cost of health care to African genocide. The letters were powerful. The students integrated facts and knowledge to make strong arguments in favor of their positions. The letters were also well-written. Each student used accurate grammar (e.g., Lunsford, 2002) and effective style (Strunk & White, 2000) to focus attention on their position rather than distract or confuse the reader. Together, the letters illustrate how writing can be used for communication and influence, and they demonstrate why writing is a cultural imperative that is taught in schools beginning at a young age (Deno, 2009).

Writing begins to emerge when young children start learning the complex phonological and orthographic systems of language (Ritchey, 2008). Later in life, effective writing skills contribute to a variety of positive outcomes, from improved learning in content courses (Bangert-Drowns, Hurley, & Wilkinson, 2004), to better prospects for post-secondary education and employment (National Commission on Writing, 2004; 2005). Unfortunately, many students struggle or fail to acquire effective writing skills, preventing them from experiencing these positive outcomes.

In the most recent National Assessment of Educational Progress (Salahu-Din, Persky, & Miller, 2008), about two-thirds of $8^{th}$ and $12^{th}$ grade students failed to meet proficiency standards for writing. Results from the previous assessment, which included students in $4^{th}$ grade, were similar: about two-thirds of students failed to demonstrate proficiency in writing (Persky, Daane, & Jen, 2003). These results show it is still necessary to give writing special consideration among efforts to improve education (National Commission on Writing, 2003).

Research suggests that improving the skills of young writers may be especially important. Poor compositional skills at a young age are related to subsequent writing problems (Berninger, Mizokawa, & Bragg, 1991), but early intervention may be able to remediate early writing problems and prevent subsequent difficulties for many students (Graham, Harris, & Larsen, 2001; Leinemann, Graham, Leader-Janssen, & Reid, 2006). For example, research with young students found that interventions as early as first grade have been shown to improve compositional skills (Berninger et al., 1997; Berninger et al., 1998; Jones & Christensen, 1999). These interventions suggest early intervention in writing is promising, but educators face challenges when targeting interventions to the students who need them.

## Statement of the Problem

A key principle for early writing intervention is tailoring instruction to meet student needs (Graham et al., 2001), but tailoring instruction requires assessment that produces instructionally-relevant data. Instructionally-relevant assessment data are a critical component of early intervention frameworks within educational settings (Batsche et al., 2005; Ysseldyke, Burns, Scholin, & Parker, 2010). These frameworks are often referred to as response to intervention (RTI). Research shows that system-wide implementation of RTI reduced incidences of learning disabilities (VanDerHeyden, Witt, & Gilberston, 2007) as well as improved learning outcomes across academic skills (e.g., Bollman, Silberglitt, & Gibbons, 2007; VanDerHeyden & Burns, 2005). Within RTI, data help educators (a) identify students who need additional resources to develop skills (e.g., Burns, Deno, & Jimerson, 2007; Burns & VanDerHeyden, 2006; Glover & Albers, 2007; Johnson, Jenkins, & Petscher, 2010; Kovaleski & Black, 2010), (b) monitor the progress of those students who are receiving additional interventions (Deno, 1985; Fuchs, 2004), and (c) match interventions with student needs (Batsche et al., 2005; Graham et al., 2001).

The use of data within RTI for early writing, however, is not as well understood as in other academic areas. Several promising measures for identifying struggling young writers and assessing their progress during interventions have been identified (e.g., Coker & Ritchey, 2010; McMaster, Du, & Petursdottir, 2009), but a better understanding of using data for targeting interventions is needed. One promising approach is to use data from curriculum-based assessment for instructional design (CBA-ID; Gravois & Gickling, 2008). The purpose of CBA-ID is to match interventions to the current skill level of students to create a task-skill difficulty level that is instructional, meaning the task demands of intervention do not exceed student skills and student skills do not exceed the task demands. The use of CBA-ID data to target and inform intervention effectiveness has been successful in reading (e.g., Burns, 2007; Burns & Parker, in press; Treptow, Burns, & McComas, 2007) and math (Burns, Codding, Boice, & Lukito, 2010; Codding et al., 2007), but much less is known about the use of CBA-ID data within early writing.

## Study Purpose

The purpose of this study is to broaden the research base for the use of data within early intervention for writing difficulties. Specifically, the study examined the role of CBA-ID data in early writing with two primary objectives. First, the study used a novel data set to derive criteria for use in CBA-ID for early writing and compared the results with previous research. Second, the study used a statistical approach that is appropriate for modeling longitudinal data (Muthén, 2004) to examine patterns of writing growth during the context of an RTI framework for early writing. Students who struggled with early literacy skills all received a research-based intervention (Williams, Stafford, Lauer, Hall, & Pollini, 2009), and their progress was assessed using a promising measure of early writing progress (McMaster et al., 2009). The patterns of growth for different students were then predicted using relevant variables, including data derived from CBA-ID.

## Significance of Study

This study is designed to inform practice and research related to early writing assessment and intervention. Practitioners who seek to intervene early with writing problems require instructionally-relevant data, and the results of this study are intended to help clarify which data can be used to target interventions. Research that seeks to improve outcomes for struggling writers needs to be based on sound theoretical models of early writing development. The current study is intended to broaden understanding of how early writing skills are acquired within current models of early writing development. The ultimate goal for successful early intervention is to provide students the support they need to improve a targeted skill, and this study is designed to contribute to meeting that goal in early writing.

## Research Questions

The following research questions guided the study:

1. What is the correspondence between instructional level criteria that are derived from the writing performance of students in different locations and under different conditions?
2. What are the growth patterns of students receiving a research-based early writing intervention?
3. What are the differential effects of pre-intervention instructional difficulty and intervention intensity on growth patterns of early writing?

**Definitions**

Curriculum-based assessment for instructional design (CBD-ID): An assessment approach that seeks to match interventions to the current skill level of students to create an appropriate level of task difficulty. CBA-ID was originally defined by Gickling and Havertape (1981).

Curriculum-based measurement (CBM): An assessment approach developed by Deno (1985) that uses brief, technically adequate measures of academic performance that can be administered repeatedly across time.

Growth mixture modeling (GMM): A type of longitudinal data analysis that relaxes the assumption that growth patterns of student performance represent a single population with a single set of growth parameters. Instead, it allows for identification of multiple growth trajectories, each with a set of growth parameters.

Instructional hierarchy (IH): A framework for understanding academic skill responding proposed by Haring and Eaton (1978) in which skill development passes through stages of (a) acquisition, (b) proficiency, (c) generalization and maintenance, and

(d) adaption. The framework outlines a corresponding instructional approach for each stage.

Instructional level (IL): The level of task difficulty at which student skills will best respond to instructional demands.

Response to intervention (RTI): A framework for providing high-quality instruction and matching interventions to students needs that uses learning rates and level of performance to make important educational decisions (Batsche et al., 2005).

Simple view of writing: A model of writing development proposed by Juel, Griffith, and Gough (1986) and elaborated by Berninger & Amtmann (2003) that is appropriate for early writing. It depicts early writing as having three components: transcription, self-regulation, and text production.

Skill-by-treatment interaction: A paradigm that targets interventions based on student skill performance rather than on underlying aptitudes (Burns, Codding, Boice, & Lukito, 2010).

Transcription skills: The skills consisting of spelling and handwriting that form a foundation for writing and likely take temporal precedence in early writing development.

## Delimitations

The following limitations were placed on the study:

(a) Early writing was studied only for first grade students in their spring semester.

(b) Study participants only included a sample of at-risk students in the Southeastern United States representing only seven schools.

(c) Early writing progress was assessed using only the picture-word measure for CBM-W (McMaster, Du, & Petursdottir, 2009).

(d) Early writing progress was assessed only weekly or bi-weekly.

(e) The intervention targeted only self-regulatory skills (Williams, Stafford, Lauer, Hall, & Pollini, 2009).

## Organization of the Dissertation

This dissertation is organized around four additional chapters. Chapter 2 provides an overview of the literature relevant to (a) using data in frameworks of intervention, (b) early writing development and assessment, (c) targeting interventions in early writing, and (d) longitudinal analysis that is appropriate for assessing growth trajectories during interventions. Chapter 3 outlines the methodology used in the current study. It describes the characteristics of the 147 participants, the measures used for screening and assessing progress, the intervention that was used and its implementation procedures, and the data analysis. Chapter 4 reports the results for each research question and includes several tables and figures to help interpret the data. Chapter 5 discusses the results within the context of previous research. It also discusses the results in terms of implications for practice, writing development, and future research. The chapter concludes with limitations for interpreting the current data.

# CHAPTER 2

# LITERATURE REVIEW

## Organization of the Chapter

This chapter provides an overview of research that is relevant for better understanding the use of data to target and inform early writing interventions. It is organized into four sections. The first section broadly discusses the use of data within educational interventions. It covers the use of data to identify struggling students and monitor their progress during interventions. It also briefly reviews past and current uses of data for targeting intervention. The next section focuses on early writing development and assessment. It identifies a useful model for early writing development and provides an overview of research related to assessing early writing skills. The third section discusses the systematic use of data for targeting interventions in early writing. It synthesizes research from two useful frameworks for interpreting assessment data: curriculum-based assessment for instructional design (CBA-ID; Gravois & Gickling, 2008) and the instructional hierarchy (IH; Haring & Eaton, 1978). The chapter concludes with a section that introduces and describes the analysis used to answer the current research questions. The final section was included because the analysis is relatively new to longitudinal educational research. It offers a novel approach to analyzing longitudinal data that can identify separate growth patterns within a single study sample, which is appropriate for examining intervention response.

## The Use of Data within Intervention

Research suggests that student and school-level outcomes improve when schools deliver intervention earlier. For example, the Minneapolis Public Schools used a four-step problem-solving model for decision-making to provide earlier intervention for struggling students, resulting in higher achievement levels for participating students (Marston, Lau, & Muyskens, 2007). In another study, data were used to validate a wide range of student academic difficulties and provide corresponding interventions, and results showed significant gains in mathematics scores for at-risk students (VanDerHeyden & Burns, 2005), as well as a marked decrease in the number of referrals for special education evaluations (VanDerHeyden, Witt, & Gilberston, 2007).

Early intervention also holds promise for preventing and remediating the problems students experience with writing (Graham et al., 2001; Leinemann et al., 2006), which continue to be widespread across geographic and demographic variables (Persky et al., 2003; Salahu-Din et al., 2008). Response to intervention (RTI) is a framework for early intervention that can be defined as "the practice of providing high-quality instruction and interventions matched to students needs and using learning rate over time and level of performance to make important educational decisions" (Batsche et al., 2005, p. 5). This definition implies three primary uses of data within RTI frameworks: (1) identifying, or screening, which students need additional interventions; (2) assessing the progress of students who receive additional interventions; and (3) targeting interventions that are 'matched' to students' needs.

**Screening**

A substantial amount of research has examined the use of data within RTI for screening decisions. Screening can be defined as the process of assessing all students in a given classroom, grade, or school with measures that help school personnel determine which students may require supplemental instruction (Ikeda, Neessen, & Witt, 2008).

Accuracy and efficiency are primary concerns when conducting screening (Glover & Albers, 2007; Johnson, Jenkins, Petscher, & Catts, 2009). Usability (Glovers & Albers, 2007), which is related to efficiency, is often the determining factor in educators' choices of screening measures, leading to the use of data with unknown or poor technical characteristics such as data from informal reading inventories (Parker et al., 2012). However, research suggests that using multiple measures (Johnson, Jenkins, & Petscher, 2010; Speece et al., 2010) or multiple steps (Compton et al., 2010) may be the most accurate and cost-effective approaches to screening (Fuchs, Fuchs, & Compton, 2012).

Identified students receive more intensive support within RTI frameworks regardless of the screening approach used to identify them. Additional levels of support are typically provided in what are referred to as Tier II or Tier III interventions (Burns, Deno, & Jimerson, 2007). Research has not definitively determined the procedures for providing students Tier II or Tier III support (Fuchs et al., 2012), but in general the intensity of the intervention (i.e., frequency, duration, and problem solving) increases from Tier II to Tier III levels (Al Otaiba, Schatschneider, & Silverman, 2005; Barnett, Daly, Jones, & Lentz, 2004; Burns et al., 2007; Tilly, 2008).

**Assessing Progress**

Data are also necessary for assessing the progress of students once they are identified as requiring additional support within an RTI framework. Data from curriculum-based measures (CBMs) are useful for this purpose because they are brief, technically adequate indicators of student performance (Deno, 1985; Fuchs, 2004). Progress assessments are conducted with regular, often weekly, administrations of CBMs that produce data that can be used to determine whether students are making desirable progress (e.g., Deno, 1985; Deno & Mirkin, 1977; Marston, 1989). Moreover, commercial systems have been developed that incorporate progress data from CBM assessments into graphing software for easier interpretation (e.g., AIMSweb; Pearson, 2010).

Progress assessments monitor growth by assessing general outcomes, such as oral reading fluency (Wayman, Wallace, Wiley, Ticha, & Espin, 2007), numerical reasoning (e.g., Chard et al., 2005), computation (e.g., Christ, Johnson-Gros, & Hintze, 2005; Fuchs, Hamlett, & Fuchs, 1998), problem solving (e.g., Fuchs, Hamlett, & Fuchs, 1999), or written production (e.g., Coker & Ritchey, 2010; McMaster, Du, & Petursdottir, 2009; McMaster et al., 2011). General outcome measurements are a composite of academic subskills that provide an indicator of progress in acquiring important learning outcomes (Fuchs & Deno, 1991) and allow for idiographic evaluation of intervention effects when administered over time (Shapiro, 2004).

## Targeting Interventions

The role of data in matching student needs to appropriate interventions (Batsche et al., 2005; Graham et al., 2001) has not been as thoroughly researched as the role for data in screening and assessing progress. The fields of psychology and education have a history of attempting to use cognitive diagnostic data for targeting interventions, but results were conclusively negative (Arter & Jenkins, 1979; Cronbach & Snow, 1977; Kavale & Forness, 1987, 2000). Cronbach and Snow (1977) spent several years attempting to find an aptitude by treatment interaction (ATI) for targeting interventions, but the ineffectiveness of the approach coupled with the proven results of direct approaches led them to the conclusion that targeting interventions based on cognitive aptitudes had little to no practical value. Kavale and Forness (2000) employed meta-analysis in subsequent research to determine that direct instruction approaches were consistently more effective for remediating skill deficits than were ATI approaches, which in their research showed negligible effects on outcomes.

Although ATI approaches to designing interventions have not been proven effective, other approaches that focus on a skill-by-treatment interaction have been more promising (Burns, Codding, Boice, & Lukito, 2010). A skill-by-treatment interaction focuses on assessing specific skills that are malleable with intervention and are directly linked to increased academic performance (Burns et al., 2010). Curriculum-based assessment for instructional design (CBA-ID; Gravois & Gickling, 2008) has been considered promising for identifying appropriate interventions in reading and math (Burns, Dean, & Klar, 2004), and was suggested as an effective framework for skill-by-treatment interaction decisions (Burns et al., 2010). Early research in reading showed that interventions based on CBA-ID produced gains in student on-task behavior and reading comprehension (Gickling & Armstrong, 1978). Subsequent research in reading found that CBA-ID data were sufficiently reliable (Burns, Tucker, Frame, Foley, & Hauser, 2000) and valid (Burns, 2004) for use in making instructional decisions. Similar results were observed for the reliability and validity of CBA-ID data in math (Burns, VanDerHeyden, & Jiban, 2006). Within an RTI context, studies investigating the use of CBA-ID data to inform intervention found that students demonstrated improved outcomes in reading (Burns, 2007; Burns & Parker, in press; Treptow, McComas, & Burns, 2007) and math (Shapiro, & Ager, 1992; VanDerHeyden & Burns, 2005), but no research has investigated the use of CBA-ID data in early writing.

**Synthesis**

Frameworks of early intervention use data to make several instructionally-relevant decisions (Batsche et al., 2005). Research has determined a role for data in screening and assessing progress of struggling students within early intervention frameworks, but less research is available for the systematic use of data to target interventions. Aptitude-by-treatment approaches that targeted interventions on cognitive or other indirect data were not successful (Cronbach & Snow, 1977; Kavale & Forness, 1987, 2000), whereas skill-by-treatment approaches have been more promising (Burns et al., 2010). The skill-by-interaction framework uses CBA-ID data for targeting interventions, but very little is known about the role of CBA-ID data in matching interventions to students' early writing needs. Additional research to determine the use of CBA-ID data to help target early writing interventions should be informed by current understanding of early writing development and assessment.

## Early Writing Development and Assessment

Although several models of writing development exist, including cognitive/motivational models (e.g., Hayes, 1996) and sociocultural models (e.g., Schultz & Fecho, 2000), each was largely oriented toward adult writing and composing written text. Young children must first develop the component skills of writing (i.e., spelling, transcription, planning) before composing complex texts. Scardamalia and Bereiter (1987) discussed the difference between mature and immature writers and noted that immature writers typically use a knowledge-telling approach in which content and discourse knowledge account for writing quality. But models for both types of writers did not consider the importance of basic compositional skills such as syntactic fluency

(Bereiter & Scardamalia, 1987), which subsume other early writing skills that are important for successful writing.

**Simple View Model of Early Writing**

The ‘simple view’ model of writing development is more relevant for understanding the writing of younger students (Juel, 1988; Juel, Griffith, & Gough, 1986), because it focuses on a period when basic compositional skills are still developing. Berninger and Amtmann (2003) expanded the simple view of writing and suggested that text production, which is the goal of writing, relies on self-regulation strategies such as goal setting and self-monitoring to guide the writing process, as well as transcription skills such as accurate and fluent (automatic) production of text (via handwriting or typing). Within the simple view model of writing development, these three processes (i.e., text production, self-regulation, and transcription) are hypothesized to be constrained by limited working memory resources such that when more cognitive resources are required for transcription, less are available for self-regulation or text generation (McCutchen, 1996). The simple view model recognizes that skills involved in basic composition develop over time and affect broader writing outcomes, which helps researchers identify early writing skills that are important for assessment and intervention.

Research with young students has investigated hypotheses derived from the simple view model of writing development. Abbott, Berninger, and Fayol (2010) found that text production and spelling skills (i.e., transcription skills) were significant predictors of later writing skills. Other work found that transcription skills were a good predictor of compositional quality and length in elementary school, explaining up to 25%

and 66% of the variance, respectively (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Transcription skills were also found to continue developing even into college: the predictive strength of fluent letter writing with quality of text was maintained throughout high school and college (Connelly, Campbell, MacLean, & Barnes, 2006; Peverly, 2006). Related basic research found that children's recall in working memory tasks was inhibited when the recall was in the written versus oral medium (Bourdin & Fayol, 1994), which provided support for the hypothesis that transcription skills tax the working memory system (McCutchen, 1996). Finally, intervention studies showed that instruction in transcription skills (e.g., handwriting or spelling) transferred to compositional skills for struggling first grade writers. Students who participated in scripted lessons designed to improve letter-writing automaticity or basic spelling skills showed improved compositional fluency over students in control conditions (Berninger et al., 1997; Berninger et al., 1998; Graham, Harris, & Fink, 2000; Jones & Christensen, 1999).

The simple view model of writing offers a useful framework for understanding early writing development. It provides hypothesized links between foundational writing skills (i.e., transcription and self-regulatory skills) and their importance for later writing skills (i.e., text production). The model also recognizes that transcription skills alone do not lead to successful writing. The ultimate goal for writing, both practically and according to the simple view model, is text production that carries meaning (Berninger & Amtmann, 2003). The simple view model of writing suggests that some struggling young writers may have deficits in transcription skills that inhibit their overall writing quality. It also recognizes that some struggling students may have sufficient transcription skills but could benefit from intervention that focuses on other writing skills (e.g., self-regulatory strategy instruction; Graham & Harris, 2009). By providing a model for how early writing skills develop, the simple view of writing clarifies which assessment data will hold the greatest potential for targeting intervention for struggling young writers.

## Early Writing Assessment

Assessments for educational decision making should produce data with adequate psychometric (Thorndike, 2005) and practical characteristics (Choate, Enright, Miller, Poteet, & Rakes, 1992). Educational assessments should also be evaluated based on the consequential aspects of using and interpreting scores obtained from assessments (Messick, 1995). To this end, formal assessments (i.e., norm-referenced, standardized assessments) possess several inherent strengths including standardized administration directions and the capacity to compare students to normative groups (Penner-Williams, Smith, & Gartin, 2009). However, the data from formal assessments provide minimal direction for instructing struggling students (Stiggins, 2005) and lack practical characteristics of tests that are usable within an educational context such as simple administration, cost-effectiveness, and efficiency (Choate et al., 1992). The relative instructional deficiencies of formal assessments limit their capacity to provide instructionally-relevant data (Salvia, Ysseldyke, & Bolt, 2010) and suggest that alternative writing assessments should be considered.

Writing assessments that can be used formatively assist educators in making instructional changes that lead to improved writing skills (Graham, Harris, & Hebert, 2011). Curriculum-based measurement (CBM; Deno, 1985) is a formative assessment approach that has been suggested as an alternative to standardized assessments (Deno, 1985; Fuchs, 2004). CBM for writing (CBM-W) has variants across one or more of three primary features: type of task, duration of sample, and scoring procedures (McMaster & Campbell, 2008). The responses to CBM-W prompts can be scored with different metrics, including total words written, words spelled correctly, or a count that includes grammatical and structural accuracy such as correct word sequences (Videen, Marston, & Deno, 1982) or correct minus incorrect sequences (Espin et al., 2008).

Marston (1989) found that CBM-W generally yields scores with sufficient psychometric properties. Subsequent reviews of CBM-W research found that existing measures were likely appropriate for students beginning in third grade (McMaster & Campbell, 2008; McMaster & Espin, 2007). As a result, researchers developed new CBM-W tasks to use with younger students. Many of these measures aligned with the simple view model of early writing development (Berninger & Amtmann, 2003; Berninger et al., 1992), because their scoring procedures produced metrics of early writing skill including transcription fluency and spelling. These measures included word copying, sentence copying, and dictation tasks (Ritchey, 2006; 2008; Lembke, Deno, & Hall, 2003). Research found promising reliability ($r$ = .68 to .85) and validity coefficients ($r$ = .32 to .70) for 3 to 5 min samples of these measures using scoring procedures such as correct word sequences or correct letter sequences (McMaster et al., 2009; Ritchey, 2006; 2008). Other CBM-W prompts accounted for ideation in early writing (e.g., sentence writing based on picture prompts or simple verbal prompts), and also had promising technical characteristics (McMaster et al., 2009; Coker & Ritchey, 2010). Moreover, these CBM-W prompts produced data that could potentially be used for assessing young students' writing growth over time (Coker & Ritchey, 2010; McMaster et al., 2011), with some prompts and scoring procedures producing reliable growth slopes in as few as four time points (e.g., picture-word measures scored with correct word sequences; McMaster et al., 2011).

**Synthesis**

The simple view model of writing offers a useful model for understanding the writing development of young students. It also helps identify the component skills of early writing that should be targeted in early writing assessment in order to produce instructionally-relevant data. Recent research has identified early writing measures that show promise for producing data that can assist in identifying struggling writers and monitoring their progress (Coker & Ritchey, 2010; McMaster et al., 2009, McMaster et al., 2011). Each measure provides data regarding young students' skills for accurately and proficiently creating text, which are important skills in early writing development (Berninger & Amtmann, 2003). These measures can be used to identify students who are deficient in early writing skills and require additional intervention, but targeting the appropriate intervention to meet their needs (Batsche et al., 2005; Graham et al., 2001) requires additional assessment procedures.

## Targeting Interventions with Data

Previous research has typically differentiated interventions based solely on intensity, such as modifying frequency and duration (e.g., Torgesen et al., 2001) or group size (e.g., Vaughn et al., 2003), or without systematically using data to target student needs. Research that directly examined the effects of different levels of intervention intensity found that additional intensity did not result in better outcomes (Denton et al., 2011; Wanzek & Vaughn, 2008). Denton et al. (2011) found that different levels of duration and scheduling did not lead to differences in intervention effectiveness. Vaughn et al. (2003) examined the effects of delivering intervention in one-to-one formats, small groups of three students, or groups of 10 students. Results showed that although students in the one-to-one and groups of three formats made significantly greater gains in comprehension compared to the students in the groups of 10 format, no differences were noted between students in the one-to-one and groups of three format (Vaughn et al., 2003). These results were similar to the conclusion reached by a meta-analysis of reading outcomes during interventions that found one-on-one formats were no more effective than small groups for improving reading outcomes (Erlbaum, Vaughn, Hughes, & Moody, 2000).

Relatively few studies have modified instructional components of intervention by explicitly or systematically using data. Several studies have examined intervention effects or modifications with students who were non-responders to previous intervention, but the modifications in subsequent intervention were not explicitly discussed in terms of how data were used to target instructional components (Berninger et al., 2002; McMaster, Fuchs, Fuchs, & Compton, 2005; Vadasy, Sanders, Peyton, & Jenkins, 2002). For example, McMaster et al. (2005) randomly assigned non-responders to one of three conditions: continued peer tutoring, modified peer tutoring, or one-on-one adult tutoring. The students who received modified peer tutoring and adult tutoring received intervention with additional instruction in foundational skills, but data were not used to place students in the conditions and results indicated nearly 50-80% of the students continued to be non-responders (McMaster et al., 2005).

Targeting interventions strictly in terms of intensity, or without systematically using data, seems to rest on implicit assumptions about student needs and the interventions available to meet those needs. Targeting interventions based strictly on intensity implies that student needs are a result of insufficient practice. Such an assumption makes sense at an intuitive level, and is supported by findings that insufficient practice is an instructional variable that may explain poor academic performance (Daly, Witt, Martens, & Dool, 1997) and that repeated practice interventions produce positive effects (e.g., repeated reading; Samuels, 1979; Therrien, 2004).

Targeting interventions without systematically using data seems to rest on an implicit assumption that it is impossible to know with certainty which interventions will be effective for struggling students. Such an assumption is supported by results showing the persistent non-response of students, even during research-based interventions (Berninger et al., 2002; McMaster et al., 2005; Vadasy et al., 2002), as well as by problem-solving approaches to intervention that use an experimental teaching method to identifying effective intervention (Deno & Mirkin, 1977; Fuchs, Fuchs, & Stecker, 2010). There are, however, several potential issues with these assumptions for targeting interventions.

The first issue is that insufficient practice is only one of several potential hypotheses for why students may be struggling with an academic skill; students may also struggle due to insufficient instruction, insufficient motivation, or inappropriate material, (Daly et al., 1997). The second issue is that emerging research provides a strong indication that data can be used to target interventions using a skill-by-treatment interaction approach (e.g., Burns et al., 2010; Burns & Parker, in press; Codding et al., 2007). Burns et al. (2010) and Codding et al. (2007) systematically used data to better understand intervention effects. Burns & Parker (in press) systematically used data to target and design reading interventions, resulting in stronger skill growth. The third issue is that many students receiving intervention will *not* be persistent non-responders in the context of the appropriate intervention, and systematically using data to target interventions can help ensure that those students are responders. Finally, although non-responders will likely be present in all intervention studies, the systematic use of data is inherent to the problem solving approach itself (Deno & Mirkin, 1977; Fuchs et al., 2010). Thus, while some students will continue to be non-responders, the systematic use of data to target interventions should be part of intervention design, as is noted by common conceptualizations of RTI frameworks (Batsche et al., 2005).

**Intervention Frameworks**

Fuchs and Fuchs (1999) noted that for assessment data to inform instruction, educators "must deliberately incorporate methods that help connect assessment information to instructional decisions" (p. 662). Early efforts to connect assessment data with instructional decision-making relied on providing practitioners feedback designed to help interpret CBM data (for an example in spelling, see Fuchs, Fuchs, Hamlett, & Allinder, 1989; in reading, see Fuchs, Fuchs, Hamlett, & Ferguson, 1992; in math, see Fuchs, Fuchs, Hamlett, & Stecker, 1990). However, these efforts did not inform instruction prospectively or systematically. Instead, they interpreted CBM data that had already been collected to help practitioners identify specific subskills that could be the focus of future instruction. In each of these methods, practitioners made subjective decisions regarding which skills to teach and how they would be taught. Other methods have the potential to more deliberately connect assessment information to instructional decisions by incorporating research-based knowledge about instructional difficulty and learning processes.

**Instructional level.** Curriculum-based assessment for instructional design (CBA-ID; Gravois & Gickling, 2008) uses assessment data for the goal of carefully controlling the difficulty of instructional material or task so that the task demands and skills of the student align. In reading and math, CBA-ID is used to determine whether the task is (a) appropriately challenging (i.e., instructional level material), (b) too difficult (i.e., frustration level material), or (c) too easy (i.e., independent level material). The purpose is to use the data to design an intervention in which the task is at the instructional level. When used to inform intervention efforts within an RTI context, CBA-ID resulted in improved outcomes for both reading (Burns, 2007; Burns & Parker, in press; Treptow et al., 2007) and math (Shapiro, & Ager, 1992; VanDerHeyden & Burns, 2005).

In early writing, however, research with CBA-ID is limited. Parker, McMaster, and Burns (2011) empirically derived instructional level criteria for young writers using methodology established by Burns, VanDerHeyden, and Jiban (2006). The approach involved longitudinal data analysis to establish instructional level criteria by identifying ranges of initial skill that corresponded to the most growth over a given period of time. Results showed that instructional level criteria for early writing produced significant reliability ($\kappa$ = .37 to .47) and validity coefficients ($r$ = .21 to .50) (Parker et al., 2011). These results were promising in terms of identifying instructional level criteria for CBA-ID procedures in early writing, but the application and interpretation of CBA-ID data within early writing intervention requires additional research.

**Instructional hierarchy.** The instructional hierarchy (IH; Haring & Eaton, 1978) is a framework for understanding skill development that has potential for application in early writing because it conceptualizes academic responding in a way that leads to specific instructional suggestions. Within the IH conceptual framework, academic responding passes through stages of acquisition, fluency, generalization/maintenance, and adaption (Haring & Eaton, 1978). Performance in the acquisition stage is slow and inaccurate. In the fluency stage, the student becomes accurate in the skill, but still performs slowly. Academic responding in the fluency stage should become faster and more automatic. The next two stages, generalization/maintenance (applying the skill across settings, tasks, and time) and adaption (applying the skill in different situations to solve problems), are developed once the student can perform the skill with accuracy and sufficient speed.

The IH conceptual framework specifies the focus for instruction depending on a student's current skills (Haring & Eaton, 1978). For example, if a student performs a skill slowly and inaccurately (i.e., acquisition stage), the instructional focus is on modeling and immediate error correction. The instructional focus for a different student who performs the skill accurately but slowly (i.e., fluency stage) will be on frequent and repeated opportunities to respond with less immediate error correction (Haring & Eaton, 1978). The IH clarifies that responding changes as academic skills are acquired, and as academic responding strengthens, different instructional approaches facilitate progression within and between the hierarchical levels (Ardoin & Daly, 2007).

The IH has facilitated understanding of the effects of different instructional procedures for academic skills (Martens & Eckert, 2007). In one study, different reading interventions had instructional elements that made them more or less effective depending on the learner's stage of skill development according to the IH (Daly, Lentz, & Boyer, 1996). For example, reading interventions with a higher degree of modeling were found to be differentially effective depending on whether students were considered at the acquisition or fluency stages (Daly et al., 1996). In math, researchers found that initial math fluency skills differentiated the effectiveness of interventions targeting either acquisition or proficiency of skills (Codding et al., 2007).

**Combining Instructional Level and Instructional Hierarchy**

The IH could also be used in early writing to better understand intervention effectiveness. Absent, however, from the IH conceptual framework are criteria for determining the stage of a student's current skills. In reading and math, this issue has been addressed using instructional level criteria that are part of CBA-ID. The study by Codding et al. (2007) used instructional level criteria that were not empirically derived (Deno & Mirkin, 1977), but results indicated that instructional level criteria showed promise for understanding the effects of IH stage on intervention effectiveness. Subsequent meta-analytic research used empirically-derived instructional level criteria (Burns et al., 2006) to examine the effectiveness of math interventions, and proposed a skill-be-treatment interaction framework using CBA-ID data to determine the level of skill development within the IH (Burns et al., 2010). Results showed that students whose initial skill level was at the frustration level and received an acquisition intervention had larger effect sizes than students who were at the instructional level and received an acquisition intervention (Burns et al., 2010). Moreover, research in reading used CBA-ID data to target accuracy- and proficiency-focused interventions, and results indicated greater growth in proficiency-focused interventions after acquisition-focused interventions facilitated reading accuracy to within the instructional level (Burns & Parker, in press).

**Synthesis**

Research suggests that targeting interventions based on assumptions of insufficient practice or a limited probability of identifying an effective intervention may not be effective for some students (Denton et al., 2011; McMaster et al., 2005; Vadasy et al., 2002). Applying CBA-ID data for early writing (Parker et al., 2011) along with the IH framework (Haring & Eaton, 1978) holds potential for understanding the effects of early writing interventions in a way that could help target interventions more systematically. Using CBA-ID procedures, a student with initial performance below the instructional level in early writing could be considered in the acquisition stage of the IH, requiring an accuracy-focused intervention that targets transcription skills. A student performing within the instructional level criteria could be considered in the proficiency stage of the IH, needing additional practice to become a faster, more automatic writer. A student whose skills are above the instructional level could be considered in the generalization or adaption stages of the IH, needing strategy instruction or continued development of discourse knowledge. Using instructional level criteria from CBA-ID to identify a students' stage within the IH framework might therefore be a way to match interventions to students' early writing needs (Batsche et al., 2005; Graham et al., 2001).

## Modeling Early Writing Growth

Research suggests that targeting interventions within an RTI context can be informed by data from CBA-ID, and that interventions aligning with a student's stage within the IH result in greater effectiveness (Burns & Parker, in press; Burns et al., 2010; Codding et al., 2007; Daly et al.,. 1996). The implication of these findings is that within a group of students receiving an early writing intervention, different growth trajectories might be observed depending on each student's initial skills as assessed by CBA-ID. However, research has not previously modeled potential differences in growth trajectories for young writers, and as a result it is unknown whether certain classes within a given sample of students follow different or similar growth trajectories in terms of acquiring early writing skills. Moreover, it is unknown whether factors such as intervention intensity or CBA-ID data affect growth trajectories, which prevents educators from identifying potentially useful data for targeting interventions.

Most longitudinal research in early writing has been conducted to determine the technical characteristics of early writing measures used to assess growth. Findings indicate that such measures are generally reliable and valid indicators that are also sensitive to writing growth (Coker & Ritchey, 2010; McMaster et al., 2011). These measures could therefore be used to model growth in the context of an early writing intervention to determine whether groups of students follow different growth trajectories.

More commonly used traditional longitudinal modeling approaches such as latent curve growth modeling may have limitations for modeling early writing growth. Latent curve growth modeling procedures assume that an entire sample comes from the same population, producing single parameter estimates for intercepts and slopes. While useful for determining sensitivity to growth (e.g., Coker & Ritchey, 2010; Parker, McMaster, Medhanie, & Silberglitt, 2011), latent growth curve modeling procedures are limited in terms of testing hypothesized differences in growth trends across subjects from a single sample. Single parameter estimates have a risk of obscuring patterns of growth that are differentiated across subgroups within the observed sample. For example, Figure 1 shows a simplified case in which two subgroups in the observed sample had markedly different intercepts and slopes. Traditional latent curve growth modeling produces estimates that fit the overall data with a single set of growth parameters (see the thick, dashed line), which does not accurately describe the growth of either subgroup depicted in Figure 1 and would result in a misspecified model.

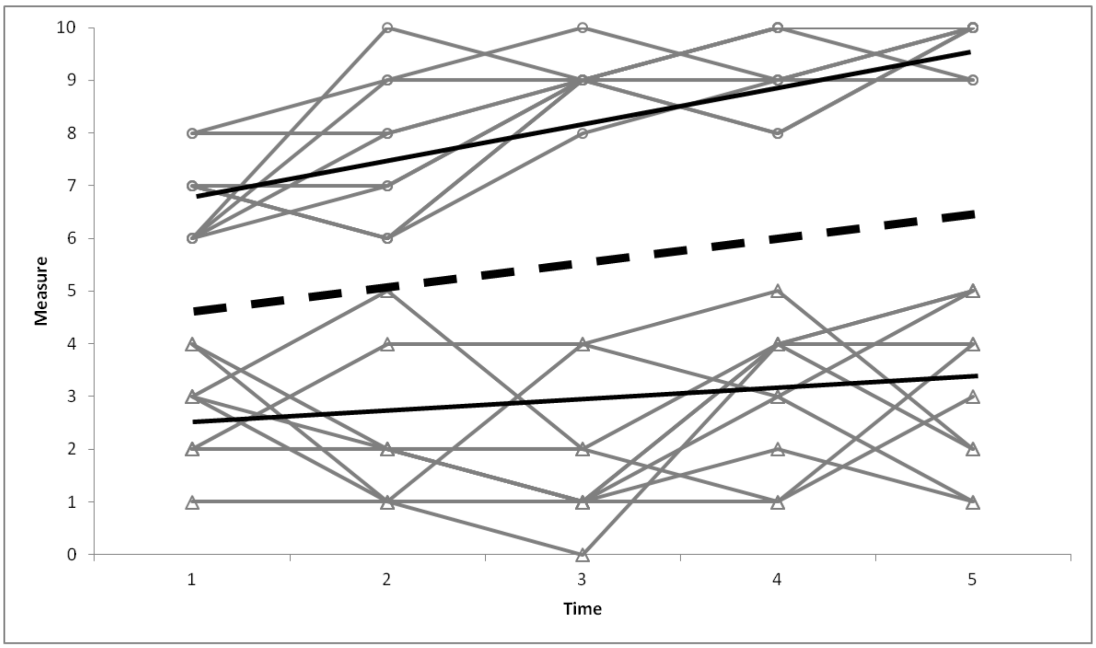


*Figure 1.* Simplified example showing differentiated patterns of individual growth trajectories. The two solid black lines indicate best-fitting growth trajectories for the groups of students depicted by circles and triangles, respectively. The dashed line shows the best-fitting growth trajectory if estimated by a single set of parameters.

Growth mixture modeling (GMM) relaxes the assumption that all individuals are from a single observed population (Muthén, 2004; Muthén & Muthén, 2000). Instead, GMM takes into consideration heterogeneity within the sample and assumes there may be multiple unobserved subpopulations of growth trajectories. Moreover, GMM allows for the modeling of effects of time-invariant covariates on unobserved subpopulation membership or growth patterns, which is helpful for identifying intervention targets and protocols (Boscardin, Muthén, Francis, & Baker, 2008). Applying GMM to longitudinal research of early writing growth permits the testing of hypotheses related to writing development, such as how data from CBA-ID inform writing growth within an RTI context, and the testing of underlying developmental hypotheses is a recommended application of the GMM approach (Wang & Bodner, 2007).

## Summary and Research Questions

The current study used longitudinal analysis to answer research questions relevant to the current literature review. The use of data within RTI frameworks includes the targeting of interventions matched to student needs (Batsche et al., 2005; Graham et al., 2001). The systematic use of CBA-ID data to understand and target interventions has support in other academic domains (Burns et al., 2010; Burns & Parker, in press; Codding et al., 2007), but research is lacking in early writing.

The following research questions guided the study:

1. What is the correspondence between instructional level criteria that are derived from the writing performance of students in different locations and under different conditions?
2. What are the growth patterns of students receiving a research-based early writing intervention?
3. What are the differential effects of pre-intervention instructional difficulty and intervention intensity on growth patterns of early writing?

# CHAPTER 3

# METHOD

## Setting and Participants

Data for the current study were collected as part of a larger, ongoing study examining the effectiveness of alternate procedures for placing students into RTI tiers (Al Otaiba, 2012). The larger study took place across multiple years and focused on reading. During the spring term of one of the academic years the larger study included a writing intervention adapted from a research-based reading intervention with a writing component (Williams, Stafford, Lauer, Hall, & Pollini, 2009).

### Schools

Participants in the larger study were recruited from a school district in a mid-size city in northern Florida, which nominated seven schools to be included in the study. All principals and all 34 first grade teachers within these schools agreed to participate. A majority of the teachers (22 teachers, 64.7%) were Caucasian, 9 (26.5%) were African American, 1 (3%) was Hispanic, 1 (3%) was Asian, and 1 (3%) was multiracial. Nine teachers held graduate degrees (26.5%) and the rest held Bachelor's degrees (77.3%). Teachers averaged about 15 years of experience (M = 14.54; SD = 9.74). The seven schools served an economically diverse area, ranging from schools with 15.8% of students who participated in the federal government Free or Reduced-Price Lunch (FRPL) program to schools with 89.9% of students participating in FRPL.

A total of 518 students participated in the larger study. Of the total sample, less than 1% were American Indian, 2% were Asian, approximately 50% were black, 43% were white, and about 4% were multiracial. Approximately 55% of the students were male.  About 1.5% of the students were considered English Learners (EL), and about 7% of the students had been retained a grade level in school.  Overall, about 57% of the students participated in the FRPL program, with 48% receiving a free lunch and 9% receiving a reduced-price lunch.  The average age of the students was 7.2 (SD = .37) years old, and ranged from 6.4 to 8.7 years.

**Criteria for Participation**

At-risk students were identified based on their performance on four standardized reading assessments.  The assessments consisted of (a) a teacher rating scale of student reading skills (Speece & Case, 2001), (b) a 1 min letter sound fluency assessment (Pearson, 2010), (c) a word identification fluency assessment (Fuchs, Fuchs, & Compton, 2004), and (d) the Sight Word Efficiency and Phoneme Decoding Efficiency subtests of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999). A student was identified as requiring Tier II intervention if they were rated below grade level on the teacher rating scale (Speece & Case, 2001), or if they scored below the $40^{th}$ percentile of the school-based local normative data for three out of four of the other measures.  To be identified as requiring Tier III intervention, a student needed to be rated below grade level on the teacher rating scale and also score below the $40^{th}$ percentile on the local normative data on all four of the other measures.

**Final Sample**

Out of the larger sample, 147 students were identified as needing Tier II or Tier III interventions, which constituted the sample for the current study.  The demographic information for the final sample is shown in Table 1.  The average age of the students in both tiers was 7.3 (SD = .41) years old, and ranged from 6.4 to 8.6 years.

Table 1. *Demographic Characteristics of Final Sample.*

| Characteristic | Tier II (n = 94) | Tier III (n = 53) |
|---|---|---|
| Ethnicity | | |
| African American | 54.3% | 73.6% |
| Asian | 4.3% | -- |
| White | 34.0% | 24.5% |
| Multiple Ethnicities | 6.4% | 1.9% |
| Gender | | |
| Male | 63.8% | 60.4% |
| Female | 36.2% | 39.6% |
| English Learner | | |
| Yes | -- | -- |
| No | 100.0% | 100.0% |
| Federal Free and Reduced-Price Lunch | | |
| Neither | 39.4% | 11.3% |
| Reduced | 9.6% | 17.0% |
| Free | 51.1% | 71.7% |
| Previous Grade Retention | | |
| Yes | 14.6% | 25.5% |
| No | 85.4% | 74.5% |

# Measures

## Screening Measures

Each screening measure was a standardized, individually-administered assessment. One was a teacher rating scale, and three were administered directly to the student in a brief session lasting about 5 min for each child. The screening measures examined students' phonics skills and early site-word reading skills, which are generally considered to be strong predictors of early reading performance (e.g., Bus & van Ijzendoorn, 1998; Snow, Burns, & Griffin, 2000). Table 2 shows the scores on the screening measures for students in Tier II and Tier III.

**Teacher rating scale.** The first screening measure was an adaptation of a teacher rating scale used in previous research (Speece & Case, 2001), in which teachers rated a student's skills on a 5-point scale consisting of 1 (well below grade level), 2 (below grade level), 3 (at grade level), 4 (above grade level), and 5 (well above grade level). Ratings were based on the Academic Competence domain of the Social Skills Rating System for Teachers, which has moderate to high reliability ($r > .80$) and acceptable content, criterion-related, social, and construct validity (Gresham & Elliot, 1990). Teachers rated each student's overall reading performance, considering the student's performance on several subskills of reading, including decoding, fluency, & comprehension. Table 2 shows the frequency and percentage with which teachers rated students at a certain level on the 5-point scale. No students were rated as well-above grade level (a score of 5), about 47% of students in Tier II compared to 13% in Tier III were rated at grade level (a score of 3), and many more students (87% vs. 42%) in Tier III were rated below or well-below grade level than students in Tier II.

Table 2. *Scores on Screening Measures for Students in Final Sample.*

| | Tier II (n = 94) | | Tier III (n = 53) | |
|---|---|---|---|---|
| Measure | Mean | SD | Mean | SD |
| Letter Sound Fluency | 52.83 | 13.13 | 50 | 13.76 |
| Word Identification Fluency* | 33.33 | 20.05 | 20.34 | 15.93 |
| Phoneme Decoding Efficiency* | 10.91 | 5.47 | 6.87 | 5.44 |
| Sight Word Efficiency* | 27.88 | 10.98 | 20.51 | 9.43 |
| Teacher Rating Scale[a,*] | Frequency | Percent | Frequency | Percent |
| Well Below GL | 12 | 12.8% | 19 | 35.8% |
| Below GL | 27 | 28.7% | 27 | 50.9% |
| GL | 44 | 46.8% | 7 | 13.2% |
| Above GL | 11 | 11.7% | -- | -- |
| Well Above GL | -- | -- | -- | -- |

GL = Grade Level; SD = Standard Deviation;
[*] = significant difference between Tiers ($p < .05$)
[a] = based on Gresham & Elliot (1990)

**Test of Word Reading Efficiency.** The first of the direct student measures was the Test of Word Reading Efficiency (TOWRE; Torgesen et al., 1999), which is a standardized assessment of an individual's skills in pronouncing words accurately and efficiently. The Sight Word Efficiency and Phoneme Decoding Efficiency subtests were administered and raw scores are reported. Test-retest, alternate-form, and interrater reliability values are reported as generally high in the TOWRE manual, ranging from .86 to .99. Validity coefficients with other standardized assessments of early reading skills were reported to be high, ranging from .75 to .90. Table 2 shows that students in Tier III had lower sight word and phonemic decoding skills than students in Tier II. Students in Tier III averaged a raw score of about seven on the phoneme decoding task compared to about 11 for students in Tier II. Students in Tier III averaged a raw score of 20.5 on the sight word task compared to about 28 for students in Tier II.

**Word identification fluency.** The second direct screening measure used was the word identification fluency task (WIF; Fuchs et al., 2004). The WIF task is a curriculum-based measure in which students read 50 first grade sight words that were selected randomly from lists of high-frequency first grade words (e.g., Zeno, Ivens, Millard, & Duvvuri, 1995). Testing is discontinued for students who cannot identify any of the sounds in the first 10 words, and scores are recorded as a zero. The WIF task developers report alternate-form reliability estimates of .97, concurrent validity estimates with standardize assessments of reading ranging from .52 to .93, and predictive validity estimates with standardized reading assessments of .45 to .80 (Fuchs et al., 2004). Table 2 shows the student scores on this measure. Students in Tier III had lower skills in this task with approximately 20 words identified, compared to 33 for students in Tier II.

**Letter sound fluency.** The last direct screening measure was the letter sound fluency (LSF) task from the AIMSweb assessments for early literacy (Pearson, 2010). The LSF task is a curriculum-based measure in which students read rows of letters and respond by identifying the most common sound the letter makes. Testing is discontinued for students who fail to identify one correct sound in the first 10 letters attempted.

Alternate-form reliability estimates for the LSF task are reported in the technical manual as .83, and predictive validity estimates range from .66 to .82 (Good et al., 2004). As shown in Table 2, student scores in Tier III and Tier II were more comparable on this measure, with students in Tier III performing slightly lower than students in Tier II.

**Writing Progress**

Student progress was assessed using a picture-word task, which is a curriculum-based measure for early writing (McMaster et al., 2009; McMaster et al., 2011). The picture-word task consists of words and pictures printed on the left side of the page with two rows of lines printed beside them (see Figure 2). A single picture-word task includes multiple pages, each with three picture-word combinations that serve as the writing stimulus. Most words were chosen from the high-frequency word bank from the Houghton-Mifflin curriculum (Cooper & Pikulski, 2005), but other high-frequency words (e.g., *window*, *bunny*) were added because some words in the curriculum word bank are difficult to represent with pictures. The picture-word task has multiple forms, each with different high-frequency words, and administration of forms is typically randomized.

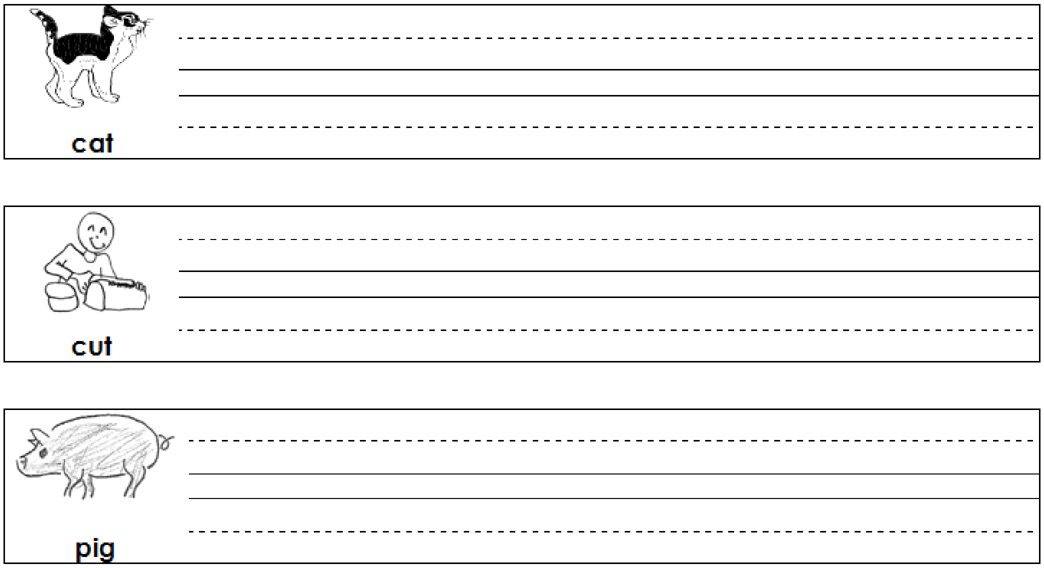


*Figure 2.* Sample picture-word prompt.

Administration of the picture-word task began with the students completing a practice task. The examiner then asked the students to continue writing sentences using the words and pictures in their packets. After 3 min, the examiner instructed participants to stop.

Each picture-word writing task was scored for correct word sequences (CWS; Videen, Marston, & Deno, 1982). A CWS is scored for any two adjacent, correctly spelled words that are also used correctly within context of written English conventions (i.e., a native speaker would judge the word sequences as syntactically and semantically correct). Incorrectly spelled or used words, missing or incorrect punctuation, and incorrect capitalization for proper nouns or the beginning of sentences were scored as incorrect word sequences, and did not count toward the total CWS. Previous research using the picture-word prompts scored for CWS reported alternate-form reliability coefficients ranging from .70 to .77, validity coefficients with a norm-referenced standardized writing assessment from .23 to .49, and sensitivity to growth in writing (McMaster et al., 2011; Parker et al., 2011).

## Procedure

### Intervention

Students who met criteria for participation in the current study received intervention in either a Tier II or Tier III condition. The intervention was the same across both conditions, and it was adapted from a text structure intervention that was previously found to improve reading comprehension and written responses with young students (Williams et al., 2009). The only difference across intervention conditions was the intensity with which the intervention was delivered. Students in Tier II participated in the intervention 2 days per week for 30 min in groups of five to seven students, and students in Tier III participated in the intervention 4 days per week for 45 min in groups of one to three students. Both intervention conditions were delivered in a small-group setting.

The intervention followed a 4 day cycle, and all procedures were based on the principles of direct instruction (Carnine, Silbert, & Kame'enui, 1997). On day 1, students read a selected story, reviewed important words, previewed a graphic organizer, and were introduced to a specific writing strategy they would use in subsequent days. Students all learned a keyword writing strategy using the words ‘First’, ‘Next’, and ‘Last’ as keywords to assist in structuring their written text. On day 2, teachers reviewed the First, Next, Last writing strategy, and students read the story in partners and completed a graphic organizer based on the writing strategy. On day 3, teachers reviewed the keyword strategy for writing again, and students read the story independently and wrote a summary of the text using the graphic organizer. On day 4, students completed a writing assessment using the First, Next, Last writing strategy and a display of their completed graphic organizer. Students also used a paragraph frame and sight word lists as visual aids. Throughout each day of the intervention, the interventionists served to introduce and reinforce use of the writing strategy and also assist students in completing the independent writing components of the intervention.

**Assessing Student Progress**

Student progress during the intervention was assessed at regular intervals using the picture-word measure. Students participating in Tier II intervention completed a picture-word probe every other week, which was necessary to ensure completion of an entire cycle of the 4-day intervention. Students in Tier III intervention completed a picture-word probe weekly. All probes were scored by the researchers, and students were not graded or evaluated directly based on their performance. The administration sequence of the probes was randomized prior to beginning the study. All students completed the same picture-word probes; students in Tier II did not complete the even-numbered probes in the sequence. Students in Tier II completed four repeated administrations using the picture-word task and students in Tier III completed seven repeated administrations. Each picture-word probe was scored by a trained scorer. Scorers included an associate professor of special education, and a doctoral graduate student who had been trained in advanced practices for curriculum-based measurement.

## Data Analysis

### Missing Data

Prior to addressing the research questions, the data set was checked for missing data. Given that students in Tier II were assessed for writing progress every other week, and students in Tier III were assessed for progress weekly, two repeated measures data matrices were examined, each with J x i rows, where J is the number of students and i is the number of repeated measures. For Tier II there were 94 students by four repeated measurement occasions. The resulting matrix contained 376 cells (94x4), and there were 45 cells with missing data. Therefore, there was approximately 12% missing data. For the students in Tier III there were 53 students by seven repeated measurement occasions. This matrix contained 371 cells (53x7), and there were 55 cells with missing data. Therefore, there was approximately 15% missing data. Combined, there were 747 cells, with approximately 13% missing data.

Students with missing data were compared to students with complete data on key demographic variables such as ethnicity, FRPL, gender, and EL status. Results of chi-square tests indicated that patterns of missing data were consistent with chance expectations for ethnicity, gender, and EL status. Results for FRPL revealed a non-random distribution only for the students in Tier II, $\chi^2(2, n = 94) = 8.98$, $p = .01$, with those who were receiving free lunch less likely to have complete data. To account for potential bias resulting from data that were missing at random (Shafer & Graham, 2002), FRPL status was included as a covariate in the final analysis.

An additional analysis compared students who had missing data with students who had complete data on the screening measures. Results showed no significant differences between the groups on the LSF and teacher rating scale. In the Tier II group, students without missing data were significantly higher in sight-word reading skills than students with missing data. On the Sight Word Efficiency task of the TOWRE, students without missing data read more words (M = 29.75, SD = 11.40) than students with missing data (M = 23.69, SD = 8.80), $t(92) = 2.55$, $p = .01$. On the WIF task, students without missing data read more words (M = 36.43, SD = 20.73) than students with missing data (M = 26.38, SD = 16.77), $t(92) = 2.30$, $p = .02$. To account for potential bias due to differences in sight-word reading skill across students with and without complete data, student scores on these screening measures were added to the propensity score analyses, which were in turn added to the final analyses.

Overall, the results of the missing data analyses indicated that students with lower socio-economic status, as measured by participation in the FRPL program, were more likely to have missing data. Moreover, although no differences in initial skills were observed between students with and without complete data for measures of phonics or as assessed by their teachers, students who had missing data had weaker sight-word reading skills. Although these results suggest a need for caution when interpreting findings in the current study, each covariate that was related to differences between students with and without missing data was accounted for in the final analyses. Moreover, the procedures used in the current study employed maximum likelihood estimation, which produces less biased estimates of parameters when the missing data mechanism can be ignored (Enders, 2001); therefore it was believed that the estimates for the longitudinal models could be considered acceptably unbiased.

**Descriptive analysis**

The first research question was in regard to comparing instructional level criteria derived from different samples and in different conditions. The computation of instructional level criteria followed methodology used in previous research of the instructional level in early writing (Parker et al., 2011) and math (Burns et al., 2006). First, slopes of growth over the duration of the intervention were calculated for CWS scores on the picture-word measure using the ordinary least squares method. Next, the students whose growth slopes equaled or exceeded the 66$^{th}$ percentile of all slopes were identified, and were considered the high-growth group. The mean score on the first progress assessment of the high-growth group was considered an estimate of the instructional level, and a range for this instructional level estimate was identified by adding and subtracting two standard errors of the mean from the instructional level estimate. Scores that exceeded the resulting range were considered in the independent level and those that fell below the range were considered in the frustration level. Once the instructional level categories were computed for each student, they were added into the final GMM analyses as a categorical, time-invariant predictor variable.

**Propensity Scores**

Given that the students in Tiers II and III had preexisting differences on the observed screening measures as part of the larger study methodology, propensity scores were computed to reduce potential bias on the longitudinal parameter estimates. Propensity scores were calculated for the predicted probability of a student being assigned to Tier III intervention. For each student, propensity scores were calculated by fitting a logistic regression model to the Tier data using scores on the teacher rating scale, AIMSweb screener, TOWRE Phoneme Decoding Efficiency and Sight Word Efficiency screeners, and WIF as the covariates. The analyses produced a propensity score for each student that estimated the logistic probability of the student being in the Tier III group, conditional on the prediction model.

**Growth Mixture Modeling**

The second and third research questions were in regard to patterns of latent class growth and the potential effects of intervention intensity and instructional level on writing growth. To answer these questions, GMM was conducted to determine best fitting models of writing growth in terms of numbers of classes, and predicted class membership probabilities were computed using intervention intensity and instructional level as time-invariant covariates. Given that GMM can be used to identify the number of classes in the best-fitting model (Muthén, 2004), the predicted class membership probabilities provided information regarding the effect of both intervention intensity and instructional level while also controlling for the effect of additional variables.

Using GMM, individuals are allowed to be in one of *K* classes that have unique growth profiles. Individual growth trajectories within a class are allowed to vary around the class profiles, both for initial status (i.e., intercepts) and rate of growth (i.e., slopes). For individual *i* in latent class *K* at time point *t*, the growth mixture model can be expressed as follows using the form of Wang and Bodner (2007):

$$Y_{Kit} = \eta_{IKi} + \eta_{SKi}\lambda_{SKt} + \varepsilon_{Kit} \tag{1}$$

$$\eta_{IKi} = \mu_{IK} + \zeta_{IKi} \tag{2}$$

$$\eta_{SKi} = \mu_{SK} + \zeta_{SKi} \tag{3}$$

where $Y_{Kit}$ is the dependent variable observed for individual *i* in class *K* at time *t*, $\eta_{IKi}$ represents the intercept for individual *i* in class *K*, $\eta_{SKi}$ represents the rate of change over time for individual *i* in class *K*, $\lambda_{SKt}$ represents the time-varying covariate reflecting time of measurement for individual *i* in class *K*, and $\varepsilon_{Kit}$ represents the error associated with individual *i*'s estimate at time *t* within class *K*. These equations illustrate that factor loadings ($\lambda_{SKt}$) and measurement error ($\varepsilon_{Kit}$) may vary across different classes, and that classes may also differ across mean ($\mu_{IK}$, $\mu_{SK}$) and variance-covariance ($\zeta_{IKi}, \zeta_{SKi}, \zeta_{ISK}$) estimates.

To incorporate a conditional model to the unconditional model described above, time-invariant covariates are added to the equations as predictors. For conditional models, the time-invariant predictor covariates are added to Equations 2 and 3 in the form

$$\eta_{IKi} = \mu_{IK} + \gamma_{IK} x_{Ki} + \zeta_{IKi} \tag{4}$$

$$\eta_{SKi} = \mu_{SK} + \gamma_{SK} x_{Ki} + \zeta_{SKi} \tag{5}$$

where $\gamma_{IK}$ represents the regression coefficient of *x* on the intercepts and slopes within class *K*. As shown in the conditional Equations 4 and 5, the effect of the regression coefficients for the predictor covariates may differ across classes. The predictive effect of the *x* covariates on class membership is then computed using multinomial logistic regression (Muthén, 2004). In the form of Wang and Bodner (2007), the resulting equation is

$$P(c_i = K \mid x_i) = \frac{e^{a_k + b_k x_k}}{\sum_{c=1}^{K} e^{a_c + b_c x_i}} \tag{6}$$

where *a* and *b* represent the logit intercept and slope. Therefore, relative to the $K^{th}$ latent class, $b_1$ is the increase in log odds of being in the first latent class versus the $K^{th}$ latent reference class for a given unit increase in *x*. The resulting GMM model is depicted in Figure 3.

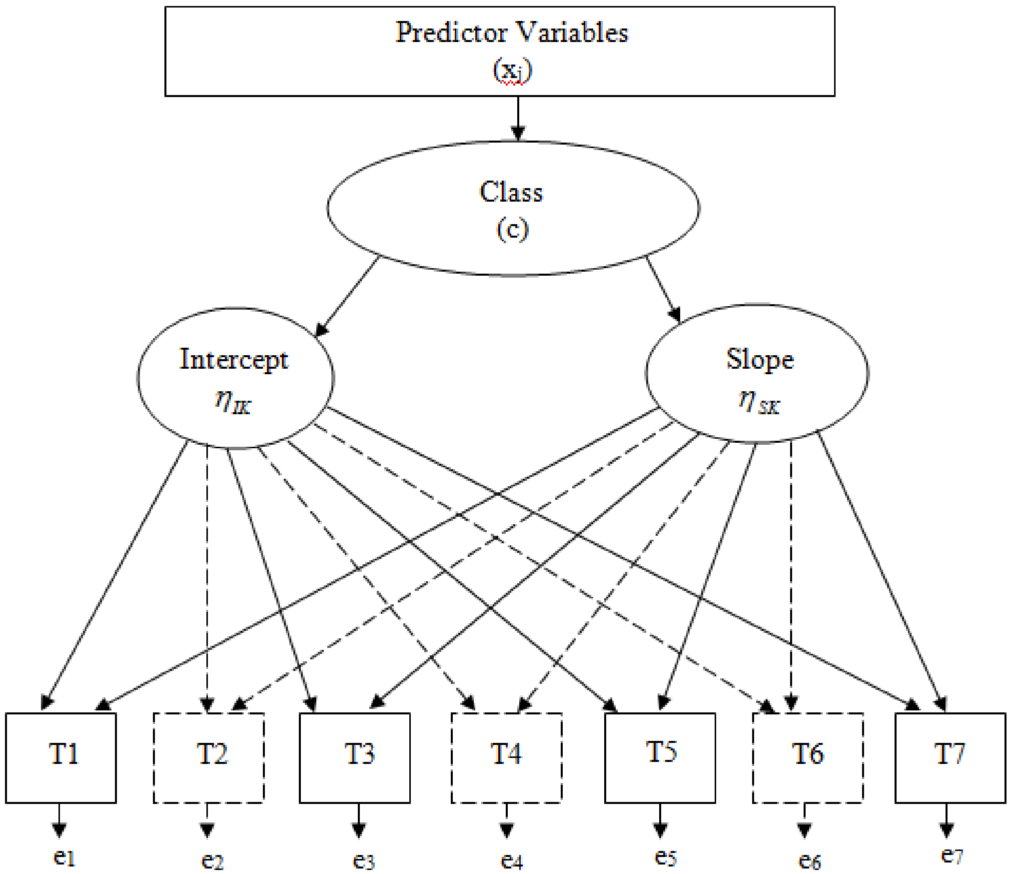


*Figure 3.* Observed and latent variables used to identify different trajectories of early writing growth. Dashed lines indicate data collected only for students in Tier III.

The next step in the GMM procedures is to identify the optimal number of classes that fit the observed longitudinal data, which should take into consideration both empirical and theoretical factors (Wang & Bodner, 2007). Empirical factors considered in the current study included the log likelihood values, Akaike's information criterion (AIC; Akaike, 1974) Bayesian information criterion (BIC; Shwartz, 1978). Smaller values were interpreted as indicators that a model with a given number of classes fit the observed longitudinal data better than models with different number of classes. The overall quality of classification was also computed for each model by using entropy (Jedidi, Ramaswamy, & Desarbo, 1993), which produces values ranging from 0.00 to 1.00. Values of entropy above .80 have been suggested to indicate good classification (Muthén, 2004).

In addition to empirical factors, substantive theory and interpretability should also be considered when determining latent classes (Wang & Bodner, 2007). On a theoretical level, the best-fitting model to the longitudinal data was considered in the context of how task demands interact with student skill (Burns et al., 2010; Haring & Eaton, 1978). It was hypothesized that students would have varying degrees of early writing skills necessary for optimal growth during intervention, which would be represented by initial skill levels (i.e., intercepts). Therefore, it was expected that the number of classes and the respective slopes would differ based on patterns of initial skill level.

## Treatment Fidelity and Interrater Agreement

### Treatment Fidelity

The interventionists implementing the writing intervention consisted of trained graduate assistants and a former elementary teacher that were part of the larger study staff. Training occurred over several days in conjunction with a senior consultant who had extensive expertise with the interventions and as a university instructor. Ongoing support and fidelity observations were also conducted during the intervention. Training was based on the text structure intervention procedures (Williams et al., 2009), and consisted of development guides, video clips, and modeling. Fidelity of intervention implementation was assessed across 36% of the intervention sessions using observational checklists, and results showed that 90% (SD = 6%) of the steps were implemented accurately (range = 88% to 93%).

**Interrater Agreement**

Interrater agreement was also computed for the picture-word measures. Approximately 15% of the sample was randomly selected to be scored by a second independent scorer. All scorers were graduate students in special education who also participated in a total of 10 hours of training for scoring the picture-word measure of early writing. Training consisted of three waves. In each wave, scorers received didactic training with models and discussion, followed by independent scoring of a subset of writing probes, followed by explicit feedback on their independent scoring performance. When all scorers had obtained an agreement of .90 with the trainer, they proceeded with independent scoring of the picture-word writing prompts. Agreement was calculated by determining the total number of ratings that produced an agreement between the original and second scorer, and dividing that number by the total number of agreements plus disagreements (McMaster et al., 2011). Overall interrater agreement exceeded .92, and was comparable to previous studies that calculated interrater agreement for CWS using the picture-word measure (McMaster et al., 2011).

# CHAPTER 4

# RESULTS

## Purpose and Research Questions

The purpose of this chapter is to provide the results of the research conducted. The chapter begins with a review of the research questions and proceeds to review the data pertaining to each question. The research questions that guided the study were:

1. Research question 1: What is the correspondence between instructional level criteria derived from the writing performance of students in different locations and under different conditions?
2. Research question 2: What are the growth patterns of students receiving a research-based early writing intervention?
3. Research question 3: What are the differential effects of pre-intervention instructional difficulty and intervention intensity on growth patterns of early writing?

## Research Question 1

The first research question examined how the instructional level criteria derived in the current study compared with previous derivations of instructional level criteria. Prior to calculating instructional level criteria, descriptive statistics were computed and are displayed in Table 3. Following previous methodology (Burns et al., 2006; Parker et al., 2011), the slope representing the 66$^{th}$ percentile was found to be 1.3 CWS growth per week of intervention. Students who had growth at or above 1.3 CWS per week were considered to be in the group with the highest slope. The mean initial CWS score for the group of students with the highest slopes was 13.96 (SEM = 1.21). The resulting instructional level range was found by creating a 95% confidence interval around the point estimate of 13.96 (13.96 ± 1.96*1.21), which resulted in an instructional range of 11.60 to 16.32. After rounding, the instructional level range for the current study was found to be 11 to 17 CWS.

Table 3. *Correct Word Sequences across Weeks of Intervention.*

| | Tier II | | Tier III | | Total Sample | |
|---|---|---|---|---|---|---|
| Week | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 1 | 14.75 | 8.59 | 11.13 | 7.09 | 13.42 | 8.23 |
| 2 | | | 13.24 | 9.08 | 13.24 | 9.08 |
| 3 | 20.96 | 12.67 | 16.11 | 10.69 | 19.24 | 12.19 |
| 4 | | | 16.1 | 8.87 | 16.1 | 8.87 |
| 5 | 18.73 | 11.74 | 15.41 | 10.67 | 17.53 | 11.43 |
| 6 | | | 18.4 | 10.83 | 18.4 | 10.83 |
| 7 | 21.18 | 12.57 | 17.35 | 11.12 | 19.86 | 12.18 |

Previous derivations of instructional level criteria for early writing were found to be 8 to 14 CWS using the picture-word prompt (Parker et al., 2011). Table 4 displays the characteristics of the previous study compared with the characteristics of the current study. Although both studies examined the writing of first grade students, notable differences in sample characteristics were observed across a variety of variables, including gender, ethnicity, location, time during the school year, and instructional conditions.

Table 4. *Characteristics of Studies that Empirically Derived Instructional Level Criteria.*

| Study Characteristic | Parker, McMaster, & Burns (2011) | Current Study |
|---|---|---|
| N | 85 | 147 |
| Grade | First | First |
| Percent male | 51.0% | 62.6% |
| Ethnicity | | |
| White | 41.2% | 30.6% |
| African American | 28.2% | 61.2% |
| Hispanic | 25.9% | 4.1% |
| Other | 4.8% | 2.7% |
| FRPL | 57.0% | 70.7% |
| EL | 21.0% | 0.0% |
| Location | Upper Midwest United States | Southeast United States |
| Community | Major urban area | Mid-size city |
| Month of first Data collection | January | April |
| Instructional conditions | Uncontrolled | Research-based intervention |
| Number of Assessments | 12 | 4 (Tier II) or 7 (Tier III) |
| Instructional Level Range* | 8 to 14 | 11 to 17 |

* in Correct Word Sequences

The instructional level range was then used to code each student's initial picture-word assessment score as below (frustration), within (instructional), or above (independent) the instructional level. The resulting categories were added to the final analyses to determine the effect of instructional level on subsequent growth trajectories modeled using GMM procedures. To do this, the CBA-ID data were added as a dummy-coded covariate that indicated the level of student skill prior to starting intervention.

## Research Question 2

The second research question examined whether students followed different latent class growth patterns during the context of the early writing intervention. Answering this question required determining the number of latent classes within the best-fitting model for the data. Table 5 shows the fit indices for each model. Models with more classes generally fit the data better, suggesting that students likely did not follow a single growth trajectory during the intervention. The greatest improvement in fit was between the one- and two-class models and a moderate improvement in fit was observed between the two- and three-class models. Models with four and five classes resulted in minimal improvement of fit, and the BIC value increased between the four- and five-class model, suggesting a decrease in fit for the five-class model versus the four-class model. Table 5 also shows that the two and three class models each had good values for entropy ( > .80; Muthén, 2004), with smaller values for the four and five class models.

Table 5. *Fit Indices for Growth Mixture Models.*

| Growth Mixture Model | Log Likelihood | AIC | BIC | Entropy |
|---|---|---|---|---|
| One-class | -2447.30 | -2450.30 | -2454.78 | -- |
| Two-class | -2119.87 | -2131.87 | -2149.03 | 0.90 |
| Three-class | -2055.50 | -2076.50 | -2106.53 | 0.89 |
| Four-class | -2022.48 | -2052.48 | -2095.38 | 0.78 |
| Five-class | -2007.34 | -2046.34 | -2102.11 | 0.82 |

Note: *N* = 129. AIC = Akaike information criterion; BIC = Bayesian information criterion

The three-class model was selected due to higher interpretability and its alignment with how the skill-by-treatment framework (Burns et al., 2010) would hypothesize student writing growth to occur during the research-based intervention. Given the CBA-ID results, students were either (1) in the frustration level (i.e., needing acquisition-focused intervention), (2) in the instructional level (i.e., needing proficiency-focused intervention), or (3) in the independent level (i.e., needing generalization-focused, or strategy development, intervention). A three-class solution facilitated interpretation with these potential CBA-ID results.

Table 6 shows the parameter estimates for the three-class model as well as the number and percentage of students in each class. Class 1 included 45 (35.7%) students, who started intervention with skill to write about 6.5 CWS, and who showed little to no growth (i.e., slope not significantly greater than zero). Class 2 included 67 (51.1%) students, who started intervention with being able to write about 17 CWS, and who gained about 1 CWS per week of intervention. Class 3 included 17 (13.2%) students, who started intervention being able to write about 28 CWS and gained twice as many words per week as Class 2. Figure 4 shows that Class 1 had the lowest performance on the first picture-word writing assessment and made little or no growth during the writing intervention, whereas Class 3 had the highest initial performance and grew the most during the intervention.

Table 6. *Number, Percentage of Students, and Parameter Estimates of Growth Trajectories for Each Class in Three-Class Model.*

| | Students | | Intercept | | Slope | |
|---|---|---|---|---|---|---|
| Class | n | % | *M* | *SE* | *M* | *SE* |
| 1 | 45 | 35.7 | 6.64 | 0.92 | 0.41* | 0.24 |
| 2 | 67 | 51.1 | 16.71 | 0.77 | 1.13 | 0.20 |
| 3 | 17 | 13.2 | 28.29 | 1.52 | 2.17 | 0.41 |

* slope not significantly greater than zero (*p* = .08)

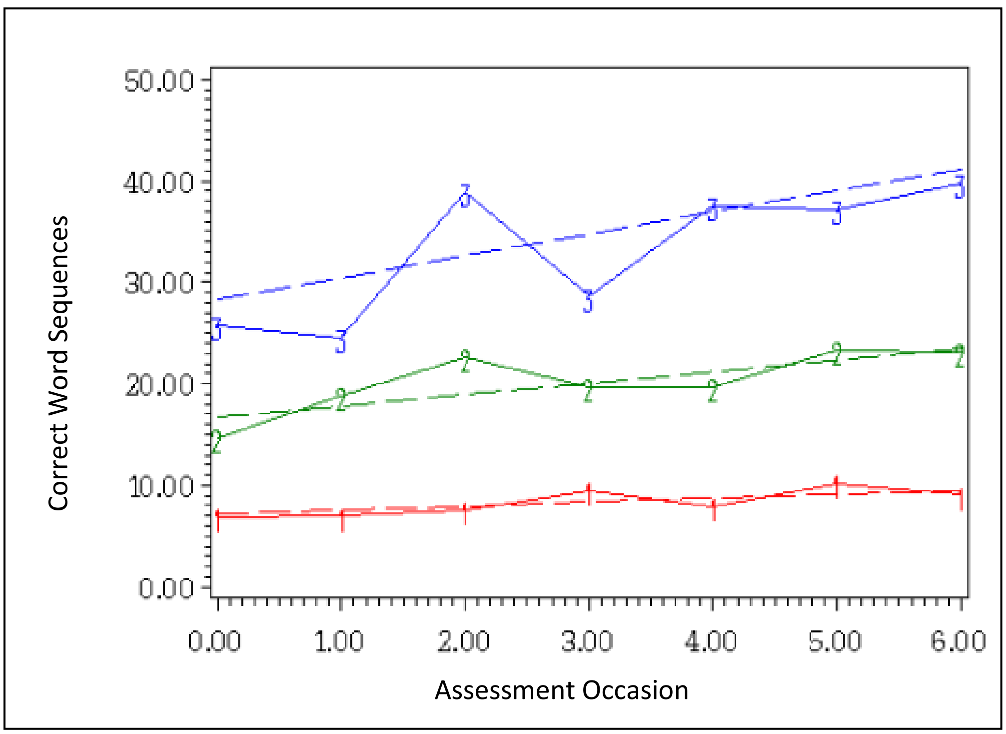


*Figure 4.* Fitted early writing growth trajectories for each of the three classes.

## Research Question 3

The third research question asked about the differential effects of pre-intervention instructional difficulty and intervention intensity on latent growth patterns of early writing. Table 7 displays the estimated multinomial logistic regression coefficients and corresponding standard errors of the time-invariant predictor variables on latent class membership. The reference group for the log-odds of class membership is Class 1, the low initial score and low growth class. Gender, which was coded 1 for males and 0 for females, was significantly and negatively related to class membership, in that boys were significantly less likely than girls to be in Classes 2 or 3 relative to Class 1. None of the other demographic variables were significant predictors of class membership.

Table 7. *Logistic Coefficient Estimates for Predictors of Class Membership.*

| | Latent Class | | | |
|---|---|---|---|---|
| | Class 2 | | Class 3 | |
| Predictor | Estimate | SE | Estimate | SE |
| Federal Free and Reduced-Price Lunch | -0.04 | 0.36 | 0.44 | 0.55 |
| Gender (Male = 1) | -1.20* | 0.59 | -2.91* | 1 |
| Propensity Score | -0.97 | 1.51 | -2.23 | 2.64 |
| Intervention intensity | -0.31 | 0.69 | -0.45 | 1.15 |
| CBA-ID Data | | | | |
| Frustration vs. independent | -3.73* | 1.03 | -21.46 | 2035.64 |
| Instructional vs. independent | -1.28 | 1.24 | -3.35* | 1.64 |

Note: The reference class for the multinomial estimates is Class 1.
* = $p < .05$

Intervention intensity was not a significant predictor of class membership, but CBA-ID data were found to be significant predictors of class. Students whose CBA-ID data indicated frustration level skills compared to students whose data indicated independent level skills were significantly less likely to be in Class 2 versus being in Class 1 (log odds = -3.73, $p < .01$). Similarly, students whose CBA-ID data indicated instructional level skills versus independent level skills were significantly less likely to be in Class 3 versus being in Class 1 (log odds = -3.35, $p < .05$).

# CHAPTER 5

# DISCUSSION

## Organization of the Chapter

The purpose of this chapter is to synthesize the results of the study. It starts by interpreting the results of each research question. Next, it contextualizes the current data within previous research and discusses implications for practice and models of early writing. The chapter concludes with a description of directions for future research and the study limitations.

## Review of Study Purpose

Students with successful writing skills are likely to experience a variety of positive outcomes, from increased learning in other academic areas (Bangert-Drowns et al., 2004) to a better likelihood of getting into college or getting job promotions (National Commission on Writing, 2004; 2005). Unfortunately, large numbers of students do not have successful writing skills (Persky et al., 2003; Salahu-Din et al., 2008). Early intervention holds promise for remediating the writing difficulties of students (Graham et al., 2001; Leinemann et al., 2006), but intervention should be matched to student needs (Batsche et al., 2005; Graham et al., 2001).

Data from CBA-ID in other academic skills have been helpful in understanding (Burns et al., 2010; Codding et al., 2007) and targeting (Burns, 2007; Burns & Parker, in press; Shapiro & Ager, 1992) interventions. The role of CBA-ID as a source for instructionally-relevant assessment data in early writing has not been well researched. The purpose of this study was to derive and compare novel instructional level criteria with previous instructional level criteria for writing (Parker et al., 2011), as well as to investigate the role of CBA-ID data in understanding the effectiveness of early writing intervention.

## Research Question 1:

## Correspondence between Different Instructional Level Criteria

The results of the current study found that between 11 and 17 correct word sequences (CWS) represented an instructional level range for early writing, which overlapped some, but not completely, with results of the previous study that derived instructional level criteria for early writing (8-14 CWS; Parker et al., 2011). Differences between the two instructional level criteria could be due to a number of factors. The time of year during which the study occurred could have affected the results. The previous study began in January and lasted a total of 12 weeks, whereas the current study took place in April and lasted about 8 weeks. The later start date of the current study may have been related to different levels of performance at the beginning of the two studies. Given that students typically gain nearly 15 letters per minute of handwriting speed from first to second grade (Graham, Berninger, Weintraub, & Schafer, 1998), it is likely that at the beginning of the respective studies students in the current study had slightly higher transcription skills than students in the previous study.

Other differences between the two studies may have contributed to the differences between the instructional level criteria. First, demographic characteristics differed across the two studies, including a higher proportion of English Learner (EL) students in the previous study, which may have contributed to differences in overall performance and growth during the context of the study. Second, the two studies also took place in different geographic locations of the United States and in different sized communities, and therefore different writing curricula and other contextual influences could have affected the students' performance. Third, the students in the current study were considered identified as at-risk for struggling with early literacy, whereas students in the previous study consisted of students whose risk status was not identified (Parker et al., 2011). Fourth, the instructional conditions in the first study were uncontrolled (McMaster et al., 2011; Parker et al., 2011), whereas the instructional conditions in the current study were highly controlled as part of the research-based writing intervention (Williams et al., 2009). Thus, systematic delivery of writing instruction and opportunities for students to respond may have affected student growth slopes during the study period.

Differences in the instructional contexts in each study also require different interpretations of the CBA-ID data. In the first study, the instructional level criteria constitute a range of initial skills that corresponded to the most growth in the context of minimal systematic instruction in early writing skills. In the second study, the instructional level criteria constitute a range of initial skills that corresponded to the most growth when instruction systematically targeted self-regulatory writing skills. However, it is important to note that neither the previous nor current study systematically targeted instruction for transcription skills, which take on particular importance for struggling young writers (Berninger et al., 1997; Berninger et al., 1998; Jones & Christensen, 1999; Graham & Harris, 2009). Therefore the instructional level criteria in both studies correspond to growth when at least some transcription skills were sufficiently developed.

Despite some differences between the instructional level criteria and their interpretation, the similarities suggest general trends in early writing development. The overlap in the CBA-ID data indicate that in the spring term of first grade between 8-17

CWS may represent an instructional level when a metric accounting for grammar, spelling, and usage is used to score picture-word prompts (McMaster et al., 2009). These data imply that at least some transcription skills are necessary for optimal growth when instruction does not systematically target transcription skills. In other words, when students have very low transcription skills (e.g., < 8 CWS in a 3 min early writing assessment) they may not be as likely to make acceptable progress in writing skills when instruction is uncontrolled or focuses on self-regulatory skills.

## Research Question 2:

## Growth Patterns of Students Receiving Early Writing Intervention

Results of the growth mixture modeling (GMM; Muthén, 2004) indicated that more than one growth pattern was necessary to best fit the observed data. Comparing across fit indices and classification accuracy, the three class model appeared to fit the data best. The three class model also aligned with the skill-by-treatment interaction framework that suggests growth patterns would be related to how well the intervention aligned instructionally with the students' initial skills (Burns et al., 2010; Gravois & Gickling, 2008). The results indicated that the three classes of growth patterns were characterized by (1) a group of initially low-performing students who did not make significant growth during intervention, (2) a group of students who had higher initial performance and who made modest growth during intervention, and (3) a group who had the highest initial performance and who made faster growth during intervention. These results indicated that those students who had stronger initial skills were more likely to benefit from the intervention than those with weak initial skills. Conversely, those students who had the lowest initial skills showed non-significant growth during the intervention, indicating a different instructional strategy might have been more appropriate for their needs.

These results suggest the importance of taking into consideration the possibility that student performance during intervention does not follow a homogenous pattern and therefore cannot be summarized by a single set of growth parameters. Most research regarding early writing interventions finds variability in student responding. For example, in a study of various treatment approaches for improving transcription skills involved in handwriting, up to 42% of students were non-responders, but the remaining students made acceptable gains (Berninger et al., 1997). The results of this study suggested different growth patterns across students within the larger sample of students, and given that no single intervention is likely to be 100% effective (Deno & Mirkin, 1977; Fuchs et al., 2010), other intervention studies would likely benefit from considering the possibility of multiple growth patterns.

The potential for finding different patterns of growth for different sub-groups in the sample is supported by other phenomena that when measured longitudinally have multiple growth patterns. For example, satisfaction following retirement has been found to follow different patterns over time for different groups of retirees (Wang & Bodner, 2007). Another study noted that different growth patterns for kindergarten literacy skills were related to subsequent growth patterns for literacy skills in first and second grades (Boscardin et al., 2008). A third study found that students in second grade who were struggling with spelling differed in the amount of spelling skills gained during explicit phonics instruction, with students who had the middle level of skills growing more than those with relatively high or relatively low initial skills (Amtmann, Abbott, & Berninger, 2008). In each of these studies, meaningful results were found by employing an analytic strategy that relaxed the assumption that all participants represented a single, homogenous population (Muthén, 2004) and instead allowed for the identification of different patterns of growth for subgroups of participants.

## Research Question 3: Differential Effects of Pre-Intervention Instructional Difficulty and Intervention Intensity

Results from the multinomial logistic regression indicated which predictor variables were significantly related to the likelihood of a student following a particular growth pattern. Of the demographic variables included as predictors, gender was a significant predictor of growth pattern, indicating that boys were less likely than girls to be in either the middle (middle initial skills and modest growth) and high (highest initial skill and growth) groups. This finding is consistent with previous findings of gender differences in writing. Across age levels, the writing performance of girls consistently outperformed boys in first, second, and third grades (Berninger & Fuller, 1992); in fourth, fifth, and sixth grades (Swanson & Berninger, 1996); and into adulthood (Berninger, Nielson, Abbott, Wijsman, & Raskind, 2008). In addition, a recent longitudinal study found gender to be a significant predictor of initial skill but not differences in slope (Parker et al., 2011). The current study added evidence to previous findings regarding differences in single-point assessments of writing skills; however, the finding that growth patterns with initially higher skill also had subsequently greater growth suggests that more girls than boys may have had sufficient early writing skills to benefit from the current intervention. Additional research will be necessary to investigate whether girls and boys follow different trajectories based solely on their growth patterns over time or whether different growth trajectories are a function of different initial skill.

Of the other potential predictors of growth pattern, it is notable that intervention intensity, as represented by tier membership, was not a significant predictor of growth patterns. Other research found that interventions differentiated solely on aspects of intensity (i.e., frequency or duration) did not result in greater effectiveness for struggling students (Denton et al., 2011; Wanzek & Vaughn, 2008), but it could have been hypothesized that intervention intensity would predict growth. For example, students with higher overall literacy skills (i.e., those in Tier II) might have been expected to respond better to the intervention; or conversely, students who received over twice the intervention intensity (i.e., those in Tier III) might have been expected to respond better to intervention. In either case, results would have indicated a greater likelihood of a more positive growth pattern, but the current results found that student growth patterns were not related to intervention intensity.

The CBA-ID data were related to student growth patterns, in that those students whose CBA-ID data indicated frustration and instructional levels were more likely to be in the low initial performance and non-significant growth pattern. These results suggest that CBA-ID data were a relevant consideration for targeting early writing intervention. In particular, when interpreted with the IH framework (Haring & Eaton, 1978) the results indicated that as student skill moved from needing acquisition-focused instruction for transcription skills to proficiency-focused or adaption-focused instruction, growth patterns were more likely to be stronger, as shown by the stronger slope values for Classes 2 and 3 in Table 6. Thus, students who had higher initial skills within the context of the current research-based intervention (Williams et al., 2009) were more likely to benefit from the instructional strategies used during intervention.

Previous research has found that CBA-ID data interpreted through the IH framework are helpful when understanding intervention effectiveness in other skill areas. Codding et al., (2007) examined the effectiveness of two different math interventions, one that matched the acquisition level of skill development and one that matched the proficiency level of skill development. Findings showed that students whose initial skills were in the frustration level benefited more from the acquisition-focused intervention while students whose initial skills were in the instructional level benefited more from the fluency-focused intervention. Although promising results were found, the Codding et al. (2007) study used non-empirical criteria for interpreting CBA-ID data, and therefore the data that were obtained had unknown technical characteristics. A subsequent study used empirically-derived CBA-ID data in a meta-analysis designed to further investigate the use of CBA-ID data in understanding mathematics interventions. Results indicated that for students at the frustration level, acquisition-focused interventions were more effective than proficiency-focused interventions (Burns et al., 2010).

Together with studies that have applied CBA-ID for targeting interventions (e.g., Burns, 2007; Burns & Parker, 2012; Shapiro & Ager, 1992), the above and current studies provide support for the use of CBA-ID data in better targeting interventions for struggling learners (Burns et al., 2010). The extant evidence suggests that CBA-ID data may be useful for understanding a student's current needs so that intervention can be effectively matched to those needs (Batsche et al., 2005). Data from CBA-ID across reading, math, and writing hold promise for indicating if students (a) are acquiring basic skills and need an intervention focused on acquisition, (b) are ready to become proficient in the basic skills and need an intervention focused on proficiency, or (c) are ready to generalize and apply those skills and need intervention to focus on application in different contexts and for problem solving (Haring & Eaton, 1978).

**Implications for Practice**

The current study provides additional support for the use of CBA-ID data for targeting early writing interventions. Young students who are struggling with early writing skills may have a variety of skill deficits, ranging from problems with transcription skills such as handwriting or spelling, to problems with self-regulatory skills such as planning, organizing, or revising (Berninger & Amtmann, 2003). Given that early writing development is marked by variability in performance at a given point in time (Herrick, 1960; Wann & Jones, 1986) and rapid acquisition of skills across time (Graham et al., 1998), it may be particularly difficult for educators to understand early writing problems without a systematic framework for linking assessment data to interventions (Fuchs & Fuchs, 1999). When interpreted within the IH framework (Haring & Eaton, 1978), data from CBA-ID appear to hold promise for that purpose.

As currently conceptualized, data from CBA-ID have potential to help educators understand the nature of young students’ writing problems and what intervention approach might be most successful. For example, students might need additional support to become accurate in transcription skill production when CBA-ID data indicate early writing tasks are at a frustration level (i.e., < 11 CWS in the current study; < 8 CWS in Parker et al., 2011). These students could benefit from intervention designed to build accuracy in handwriting (Berninger et al., 1997) or spelling (Berninger et al., 1998) skills.

Students whose CBA-ID data indicate early writing tasks are at an instructional level (i.e., 11-17 CWS in the current study) might need additional support to develop proficient transcription skills. These students could benefit from intervention designed to promote proficiency in using transcription skills, such as repeated practice (Parker, Dickey, Burns, & McMaster, in press) or performance contingencies (Duhon et al., 2004). Finally, students whose CBA-ID data indicate early writing tasks are at an independent level (i.e., > 17 CWS in the current study) might need support to generalize or apply those skills for problem solving or meaningful text composition. These students could benefit from interventions that promote application of early writing skills for composing, such as self-regulated strategy development (Rogers & Graham, 2008).

Viewed in light of the IH framework (Haring & Eaton, 1978), the intervention used in the current study (Williams et al., 2009) was likely a self-regulated strategy development approach because of the focus on organizing and composing text using keywords. Transcription skills were therefore assumed to be accurate and proficient, but transcription skills may need to be adequately developed prior to students experiencing success with self-regulated strategy approaches (Graham & Harris, 2009). The current results indicated that those students who had CBA-ID data at the independent level, and who therefore likely had proficient transcription skills, were most likely to have more positive growth patterns, indicating the intervention was perhaps best tailored to their specific needs (Graham et al., 2001).

**Implications for Models of Writing Development**

The simple view model of writing development proposes that transcription and self-regulatory skills are both necessary for text generation (Berninger & Amtmann, 2003). The usefulness of the model is that it delineates how transcription skills and self-regulatory skills contribute to text generation. However, aside from a recommendation that sufficient transcription skills are likely necessary in order for self-regulatory interventions to be effective (Graham & Harris, 2009), little is known about how transcription skills develop and at what point they are sufficiently developed such that self-regulatory interventions will be effective. The current interpretive framework using the IH adds to the understanding of early writing development offered by the simple view of writing (Berning & Amtmann, 2003). Specifically, results from the current study help advance understanding of the development of transcription skills and their relationship with self-regulatory skills within the simple view model of writing.

According to the IH, early writing skills begin developing with accurate creation of text, which in the simple view model of writing consists of initially learning transcription skills related to handwriting and spelling. Subsequent development of early writing skills within the IH requires learning to create accurate text proficiently, which consists of applying transcription skills with increasing speed. However, the simple view of writing does not explicitly account for proficiency of transcription skill development. The IH framework allows for a more nuanced interpretation of how such skills precede other stages of skill development. The IH framework helps to clarify that proficient use of transcription skills is likely necessary prior to learning how to apply those skills in learning self-regulatory skills for writing.

This understanding of early writing development advances the simple view of writing by contextualizing the importance of proficiency within transcription skill development. The current interpretation is supported by some of the research that has investigated the theoretical underpinnings of the simple view of writing, in that proficiency may be a necessary prerequisite to shift working memory resources away from transcription skills and toward self-regulatory skills (McCutcheon, 1996). In other words, as students are able to spend fewer cognitive resources on creating accurate text, they will have more resources available for organizing, planning, and revising text. Once students can accurately and proficiently create text, they will be better suited to learn and employ self-regulatory skills. Such an understanding of early writing is largely parallel with how accurate and fluent reading permits more cognitive resources for text comprehension (LaBerge & Samuels, 1974), and is consistent with the current findings indicating that students with the strongest initial transcription skills made the greatest growth during intervention.

**Future Research**

Several directions for future research arise from the current findings. A priority for future research should be the use of CBA-ID data to target interventions prospectively. The current study applied the data retrospectively to better understand intervention effects, but future research is necessary to test hypotheses that result from the current data. For example, single-subject studies could employ methodology similar to previous work in reading (Burns & Parker, in press) that used CBA-ID data to target acquisition and proficiency interventions across multiple-baseline designs. By using CBA-ID criteria to determine when to switch from acquisition- to proficiency-focused interventions, causal data could be collected to determine the effectiveness of the CBA-ID approach. Additional studies could employ group designs that control intervention as an independent variable. In these studies, students with CBA-ID data indicating frustration, instructional, or independent levels could be provided interventions matched to each corresponding stage of the IH, and the results would provide potentially important evidence for how CBA-ID can be used to target early writing interventions (Batsche et al., 2005; Graham et al., 2001).

Other directions for future research can be drawn from the current study. The GMM analytic strategy used to determine growth patterns likely has additional uses in other skill areas. It is likely that the phenomenon of differentiated growth patterns based on task-skill alignment is not unique to longitudinal data collection during early writing interventions. Given that the testing of hypotheses regarding longitudinal research is an appropriate use for such an analysis (Wang & Bodner, 2007), it would be important to broaden the current approach to other skills. The effectiveness of reading and math interventions could be further understood by incorporating CBA-ID data to predict student growth patterns based on whether the interventions matched or mismatched with students' initial skills.

The GMM analysis could also play a role in validating methods for deriving CBA-ID criteria across skill areas. Current methods of deriving CBA-ID data use retrospective methodology to determine the starting skill of students who had highest growth (Burns et al., 2006; Parker et al., 2011), but an assumption is made that students who did not experience the highest growth during intervention were either at an independent or frustration level, such that the task demands were mismatched with student skill. Analysis using GMM follows a more empirical approach allowing for modeling of actual growth for different groups, and could be used to derive CBA-ID criteria to be compared with previous methods.

A final direction for future research could expand the current study methodology beyond current demographic and temporal conditions. Early writing assessments extend down to kindergarten (Coker & Ritchey, 2010), and transcription skills begin to emerge prior to school entrance (Ritchey, 2008). Given the importance of intervening with writing difficulties as early as possible (Graham et al., 2001), the current results could be used to inform assessment and intervention research for students younger than first grade. Additionally, core instructional practices at the elementary level vary widely across schools and regions (Cutler & Graham, 2008), which suggests that the larger writing context within which the current assessment and intervention practices occurred may be a relevant variable to investigate.

Additional research could examine the time of year in which future studies occur. A study in the early fall of first grade might be expected to produce different results from studies in late spring, when students have likely acquired considerably more writing skills. Other research could examine the study duration to determine if the length of time or number of assessment points affects results. Moreover, additional longitudinal research could include investigations of potential non-linear growth patterns, which have been identified in previous writing research (Parker et al., 2011).

## Limitations

The results and implications for the current study need to be considered in light of several limitations. First, the CBA-ID data need to be interpreted within the context of the study characteristics. These include the timing of the study, the nature of the instruction that was occurring, and the demographics of the students. The current study included a relatively high proportion of males, students receiving free or reduced-price lunches, and students who were retained at least one grade. Each of these factors could have influenced the results and should be considered for future research. In addition, the students were all identified as at-risk for poor literacy outcomes. Research investigating normative growth has determined that it may be necessary to expect different growth outcomes for students at different initial levels (Betebenner, 2009), including different growth expectations for different deciles of students below the 40$^{th}$ percentile (as was used in the current study). Thus, future work examining use of CBA-ID data for targeting intervention may need to consider the potential for different instructional level criteria for different subgroups of at-risk students.

There are also limitations inherent to conducting the analyses with two groups of students who had preexisting differences in literacy skills. Students in the Tier II intervention group had stronger early literacy skills than students in Tier III, and research has identified a relationship between reading and writing skills at an early age (Abbott et al., 2010), which may have affected the results. The calculation of propensity scores functioned as a way to account for preexisting differences between the groups (Leow, Marcus, Zanutto, & Boruch, 2004), but optimal designs would balance differences using random assignment or stronger matching procedures. The two groups also had different schedules for progress assessments and intervention delivery, which could have affected the results. Given that one group had four assessment occasions and the other had seven, each group likely had different standard errors of the slope, which are an important consideration for longitudinal research of early literacy skills (e.g., Christ, 2006).

Moreover, using early reading skills to determine intervention intensity confounded early reading skills with intervention intensity, which potentially obfuscates interpretation of the effects of initial literacy skills and intervention intensity. Future research should consider the relationship of reading and writing skills (Abbott et al., 2010) when identifying potential screening measures for targeting intervention intensity in early writing. Studies should attempt to control early literacy skill and intervention intensity individually, such as having students of the same initial literacy skill level receive different intervention intensity levels. Controlling both factors would allow for a stronger interpretation of each factor.

There are also limitations with the GMM analytic approach. First, given that CBA-ID data were collected at a single time-point, several absent students did not have data and were therefore excluded from the analysis. The missing data analysis results permitted considering the data as missing at random, but the GMM analysis could not account for missing data for the CBA-ID predictor, which was collected at the first time point. Moreover, the analyses of missing data determined that students with complete data tended to perform better on word-level skills than those missing data, which could have affected the results of the study. Second, there are general limitations that need to be considered when conducting GMM analyses (Wang & Bodner, 2007). The approach needs to be used cautiously and always driven by theory rather than by ad-hoc exploratory analyses (Bauer, 2007; Hoeksma & Kelderman, 2006). Additional limitations consist of a lack of consensus regarding indices for determining model fit, the actual existence of latent classes that are derived from data, and convergence issues (Jung & Wickrama, 2008). Although the first two issues are more an issue of professional convention, the latter is a mathematical issue representing challenges with modeling subdistributions of growth patterns within a larger sample distribution that can affect parameter estimation (Hipp & Bauer, 2006).

**Conclusion**

Early intervention can help prepare successful writers (Graham et al., 2001; Leinemann, 2006), but effective early intervention requires matching interventions to student needs (Batsche et al., 2005). The current study examined the use of CBA-ID data to better understand early writing intervention, which is vital if frameworks of early intervention are to fulfill their promise of improving student outcomes (Glover, 2010). Moreover, every 4-8 years, there is a new President who needs to know the concerns of students around the country.

**References**

Abbott, R., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*, 281-298.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control, 19,* 716–723.

Al Otaiba, S. (2012). To wait or to intervene immediately: A randomized experiment of first grader response to intervention. Manuscript in preparation.

Al Otaiba S, Schatschneider C, & Silverman E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality, 13,* 195–208.

Amtmann D, Abbott R, Berninger V. W. (2008). Identifying and predicting classes of response to explicit, phonological spelling instruction during independent composing. *Journal of Learning Disabilities, 41*, 218-234.

Ardoin, S. P., & Daly, E. J. (2007). Introduction to the special series: Close encounters of the instructional kind—how the instructional hierarchy is shaping instructional research 30 years later. *Journal of Behavioral Education, 16*, 1-6.

Arter, J. A., & Jenkins, J. R. (1979). Differential diagnosis prescriptive teaching: A critical appraisal. *Review of Educational Research, 49*, 517-555.

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,* 29-58.

Barnett, D.W., Daly, E. J., III, Jones, K. M., & Lentz, F. E., Jr. (2004). Response to intervention: Empirically-based special service decisions from increasing and decreasing intensity single case designs. *Journal of Special Education, 38*, 66−79.

Batsche, G., Elliott, J., Graden, J., Grimes, J., Kovaleski, J., Prasse, D., … Tilly, W. D. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research, 42,* 757–786.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Erlbaum.

Berninger,V. W.,Abbott, R. D.,Vermeulen, K., Ogier, S., Brooksher, R., Zook, D., et al. (2002). Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disability Quarterly*, *25*, 59-76.

Berninger, V., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345– 363). New York Guilford.

Berninger, V., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for diagnosis of writing disabilities in primary grade children. *Journal of School Psychology, 30,* 363−382.

Berninger, V., Mizokawa, D., & Bragg, R. (1991). Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology*, *29,* 57-97.

Berninger V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology, 46,* 151-172.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., Brooks, A. et al.(1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*, 652-666.

Berninger, V. W., Vaughn, K. B., Abbott, R. D., Brooks, A., Abbott, S., Reed, E. et al. (1998). Early intervention for spelling problems: Teaching spelling unit of varying size within a multiple connections framework. *Journal of Educational Psychology*, *90*, 587-605.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing, 4*, 257-280.

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Education Measurement: Issues and Practice, 28*(4), 42–51.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The St. Croix River education district model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 319-330). New York: Springer.

Boscardin, C. K., Muthén, B., Francis, D.J., & Baker, E.L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology, 100*, 192–208.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, *29*, 591-620.

Burns, M. K. (2004). Using curriculum-based assessment in the consultative process: A useful innovation or an educational fad. *Journal of Educational and Psychological Consultation, 15,* 63-78.

Burns, M. K. (2007). Reading at the instructional level with children identified as learning disabled: Potential implications for response-to-intervention. *School Psychology Quarterly, 22,* 297-313.

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review, 39*, 69-83.

Burns, M. K., Dean, V. J., & Klar, S. (2004). Using curriculum-based assessment in the responsiveness to intervention diagnostic model for learning disabilities. *Assessment for Effective Intervention, 29*, 47-56.

Burns, M. K., Deno, S., & Jimerson, S. R. (2007). Toward a unified model of Response to Intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 428-440). New York: Springer.

Burns, M. K. & Parker, D. C. (in press). Using the instructional level to determine the appropriate reading fluency intervention. *Reading and Writing Quarterly*.

Burns, M. K., Tucker, J. A., Frame, J., Foley, S., & Hauser, A. (2000). Interscorer, alternate-form, internal consistency, and test-retest reliability of Gickling's model of curriculum-based assessment for reading. *Journal of Psychoeducational Assessment, 18,* 353-360.

Burns, M. K., & VanDerheyden, A. M. (2006). Using response to intervention to assess learning disabilities: Introduction to the special series. *Assessment for Effective Intervention, 32,* 3–5.

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. L. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35,* 401-418.

Bus, A. G., & van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: a meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*, 403–414.

Carnine, D. W., Silbert, J., Kame'enui, E. J., & Tarver, S. G. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Chard, D., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Instruction, 30*, 3–14.

Choate, J., Enright, B., Miller, L., Poteet, J., & Rakes, T., (1992). *Curriculum-based assessment and programming.* Needham Heights, MA: Allyn & Bacon.

Christ, T. J. (2006). Short-term estimates of growth using curriculum-based measurement of oral reading fluency: Estimates of standard error of slope to construct confidence intervals. *School Psychology Review, 35,* 128-133.

Christ, T. J., Johnson-Gros, K. N., & Hintze, J. M. (2005). An examination of alternate assessment durations when assessing multiple-skill computational fluency: The generalizability and dependability of curriculum-based outcomes within the context of educational decisions. *Psychology in the Schools, 42,* 615–622.

Codding, R. S., Shiyko, M., Russo, M., Birch, S., Fanning, E., & Jaspen, D. (2007). Comparing mathematics interventions: Does initial level of fluency predict intervention effectiveness? *Journal of School Psychology, 45*, 603-617.

Coker, D. L., & 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.

Compton D. L., Fuchs D., Fuchs L. S., Bouton B., Gilbert J. K., Barquero L. A., et al. (2010). Selecting at-risk first-grade readers for early intervention: eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology, 102,* 327–340.

Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of lower-order skills to the written composition of college students with and without dyslexia. *Developmental Neuropsychology*, *29*, 175–196.

Cooper, D., & Pikulski, J. (1999). *Invitations to literacy.* Boston, MA: Houghton Mifflin.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, *100*, 907 – 919.

Daly III, E. J., Lentz Jr, F. E., & Boyer, J. (1996). The instructional hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly, 11,* 369-386.

Daly, E. J., III, Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review, 26,* 554–574.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

Deno, S. L. (2009). Problem-solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10-40). New York: Guilford.

Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual.* Reston, VA: Council for Exception Children.

Denton, C. A., Cirino, P. T., Barth, A. E., Romain, M., Vaughn, S., Wexler, J., … Fletcher, J. M. (2011). An experimental study of scheduling and duration of "Tier 2" first-grade reading intervention. *Journal of Research on Educational Effectiveness, 4,* 208-230.

Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review, 33*, 429-443.

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at-risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605-619.

Enders, C.K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*, 128-141

Espin, C. A., Wallace, T., Campbell, H., Lambke, 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.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*, 188-192.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57,* 488-500.

Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review*, *28*, 659-671.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, *27*, 216–227.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children*, *78,* 263–279.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991). The contribution of skills analysis to curriculum-based measurement in spelling. *Exceptional Children, 57,* 443-452.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436-450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1990). The role of skills analysis to curriculum-based measurement in math. *School Psychology Review, 19*, 6–22.

Fuchs, D., Fuchs, L.S., & Stecker, P. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children, 76*, 301-323.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1998). *Monitoring basic skills progress: Basic math computation* (2nd ed.) [computer program manual]. Austin, TX: PRO-ED.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1999). *Monitoring basic skills progress: Basic math concepts and applications* [computer program manual]. Austin, TX: PRO-ED.

Gickling, E. E., & Armstrong, D. L. (1978). Levels of instructional difficulty as related to on-task behavior, task completion, and comprehension. *Journal of Learning Disabilities*, *11*, 559-566.

Gickling, E. E., & Havertape, S. (1981). *Curriculum-based assessment*. Minneapolis, MN: School Psychology Inservice Training Network.

Glover, T. A. (2010). Supporting all students: The promise of response to intervention. In T. A. Glover, & S. Vaughn (Eds.), *The promise of response to intervention: Evaluating current science and practice* (pp. 1-6). New York: Guilford.

Glover T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45,* 117-135.

Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., & Laimon, L. (2004). *Technical adequacy and decision making utility of DIBELS* (Technical Report No. 7). Eugene: OR: University of Oregon.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*, 170-182.

Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). The development of handwriting fluency and legibility in grades l through 9. *Journal of Educational Research, 92,* 42–52.

Graham, S., & Harris, K. (2009). Almost 30 years of writing research: Making sense of it all with The Wrath of Khan. *Learning Disabilities Research, 24*, 58-68.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92*, 620-633.

Graham, S., Harris, K., & Hebert, M. A. (2011). *Informing writing: The benefits of formative assessment. A Carnegie Corporation Time to Act report.* Washington, DC: Alliance for Excellent Education.

Graham, S., Harris, K. R., & Larsen, L. (2001). Prevention and intervention of writing difficulties for students with learning disabilities. *Learning Disabilities Research & Practice, 16,* 74-84.

Gravois, T. A. & Gickling, E. (2008). Best practices in instructional assessment. In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology V* (pp. 503-518). Bethesda, MD: National Association of School Psychologists.

Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System- Teacher*. Circle Pines, MN: American Guidance Service.

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional technology: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom*. Columbus, OH: Merrill.

Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Erlbaum.

Herrick, V. E. (1960). Handwriting and children's writing. *Elementary English, 37,* 248-258.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods, 11,* 36–53.

Hoeksma, J. B., & Kelderman, H. (2006). On growth curves and mixture models. *Infant and Child Development, 15,* 627–634.

Ikeda, M. J., Neesen, E., & Witt, J. C. (2008). Best practices in universal screening. In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology V* (pp. 103-114). Bethesda, MD: National Association of School Psychologists.

Kavale, K. A., & Forness, S. R. (1987). Substance over style: Assessing the efficacy of modality testing and teaching. *Exceptional Children, 54*, 228-239.

Kavale, K. A., & Forness, S. R. (2000). Policy decisions in special education: The role of meta-analysis. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.), *Contemporary special education research: Synthesis of the knowledge base on critical instructional issues,* (pp. 281–326). Mahway, NJ: Lawrence Erlbaum.

Kovaleski, J. F., & Black, L. (2010). Multi-tier service delivery: Current status and future directions. In T. A. Glover, & S. Vaughn (Eds.), *The promise of response to intervention: Evaluating current science and practice* (pp. 23-56). New York: Guilford.

Jedidi, K., Ramaswamy, V., & Desarbo, W. S. (1993). A maximum likelihood method for latent class regression involving a censored dependent variable. *Psychometrika, 58,* 375–394.

Johnson, E. S., Jenkins, J. R., & Petscher, Y. (2010). Improving the accuracy of a direct route screening process. *Assessment for Effective Intervention, 35,* 131-140.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research & Practice, 24*, 174-185.

Jones, D., & Christensen, C. A. (1999). Relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology, 91,* 44-49.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255.

Jung, T., & Wickrama, K.A.S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass, 2,* 302–317.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Leinemann, T. O., Graham, S., Leader-Janssen, B., & Reid, R. (2006). Improving writing performance of struggling writers in second grade. *Journal of Special Education, 40*, 66-78.

Lembke, E., Deno, S. L., & Hall, K. (2003). Identifying an indicator of growth in early writing proficiency for elementary school students. *Assessment for Effective Intervention*, *28*, 23-35.

Leow, C., Marcus, S., Zanutto, E., & Boruch, R. (2004). Effects of advanced course-taking on math and science achievement: Addressing selection bias using propensity scores. *American Journal of Evaluation, 25*, 461-478.

Lunsford, A. A. (2002). *The everyday writer* (2nd ed.). Boston: Bedford/St. Martin's.

Marston, D. (1989). A curriculum-based approach to assessing academic performance: What is it and why do it. In M. R. Shinn (Ed.). *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.

Marston, D., Lau, M. & Muyskens, P. (2007). Implementation of the problem-solving model in the Minneapolis public schools. In Jimerson, S. R., Burns, M. K, &

VanDerHeyden, A. M. (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 279-287). New York: Springer Science Inc.

Martens, B. K., & Eckert, T. L. (2007). 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,* 82-90.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8,* 299–324.

McMaster, K. L., & Campbell, H. (2008). New and existing curriculum-based writing measures: Technical features within and across grades. *School Psychology Review, 37,* 550-556.

McMaster, K. L., Du, X., & Petursdottir, A. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities, 42*, 41-60.

McMaster, K. L., Du, X., Yeo, S., Deno, S. L., Parker, D., & Ellis, T. (2011). Curriculum-based measures of beginning writing: Technical features of the slope. *Exceptional Children, 77,* 185-206.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education, 41*, 68-84.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*, 445-463.

Messick, S. (1995). Validity of psychological assessment. *American Psychologist, 50,* 741-749.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Thousand Oaks, CA: Sage.

Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882-891.

National Commission on Writing (2003). *The neglected "R:" The need for a writing revolution.* Retrieved December 10, 2011, from http://www.writingcommission.org/.

National Commission on Writing (2004). Writing: A ticket to work . . . or a ticket out: A survey of business leaders. Retrieved December 10, 2011, from http://www.writingcommission.org/prod_downloads/writingcom/writing-ticket-to-work.pdf.

National Commission on Writing (2005). Writing: A powerful message from state government. Retrieved December 10, 2011, from http://www.writingcommission.org/prod_downloads/writingcom/powerfulmessage-from-state.pdf.

National Commission on Writing (2009). Letters to the president: Student voices. Retrieved December 11, 2011, from http://www.collegeboard.com/prod_downloads/writingcom/letters-to-president-natl-comm-writing.pdf.

Parker, D. C., Dickey, B. N., Burns, M. K., & McMaster, K. L. (in press). An application of brief experimental analysis with early writing. *Journal of Behavioral Education*.

Parker, D. C., McMaster, K. L., & Burns, M. K. (2011). Determining an instructional level for beginning writing skills. *School Psychology Review, 40*, 158-167.

Parker, D. C., McMaster, K. L., Medhanie, A., & Silberglitt, B. (2011). Modeling early writing growth with curriculum-based measures. *School Psychology Quarterly, 26*, 290-314.

Parker, D. C., Zaslofske, A. F., Burns, M. K., Kanive, R., Hodgson, J., Scholin, S., et al. (2012). Comparison of diagnostic accuracy of oral reading fluency and reading inventory levels for reading failure risk among second and third grade students. Manuscript submitted for publication.

Pearson (2010). *AIMSweb progress monitoring and improvement system*. Available from http://www.aimsweb.com/.

Penner-Williams, J., Smith, T. E. C., & Gartin, B. C. (2009). Written language expression: Assessment instruments and teacher tools. *Assessment for Effective Intervention*, *34*, 162-169.

Persky, H. R., Daane, M. C., & Jin, Y. (2003). The Nation's Report Card: Writing 2002. Retrieved December 1, 2009, from http://nces.ed.gov/nationsreportcard/writing/.

Peverly, S. (2006). The importance of handwriting speed in adult writing. *Developmental Neuropsychology*, *29*, 197– 216.

Ritchey, K. D. (2006). Learning to write: Progress-monitoring tools for beginning and at-risk writers. *Teaching Exceptional Children*, *39*, 22-27.

Ritchey, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. *Reading and Writing: An Interdisciplinary Journal, 21*, 27-47.

Rogers, L., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100,* 879‑906.

Salahu-Din, D., Persky, H., & Miller, J. (2008). The Nation's Report Card: Writing 2007. Retrieved December 1, 2009, from http://nces.ed.gov/nationsreportcard/writing/.

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). *Assessment: In special and inclusive education* (11th ed.). Boston, MA: Houghton-Mifflin.

Samuels, S. J. (1979). The method of repeated reading. *The Reading Teacher, 32*, 403-408.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Vol. 2. Reading, writing, and language learning* (pp. 142-175). Cambridge, MA: Cambridge University Press.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147-177.

Schultz, K., & Fecho, B. (2000). Society's child: Social context and writing development. *Educational Psychology, 35*, 51–62.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford Press.

Shapiro, E. S., & Ager, C. (1992). Assessment of special education students in regular education programs: Linking assessment to instruction. *Elementary School Journal, 92,* 283-296.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academies Press.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93,* 735–749.

Speece, D.L., Ritchey, K. D., Silverman, R., Schatschneider, C., Walker, C. Y., & Andrusik, K. N. (2010), Identifying children in middle childhood who are at risk for reading problems. *School Psychology Review, 39*, 258-276.

Stiggins, R. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan, 87,* 324-328.

Strunk, W. & White, E. B. (2000). *The elements of style* (4th ed.). Needham Heights, MA: Longman Publishing.

Swanson, H. L. & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of Experimental Child Psychology, 63,* 358 – 385.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial and Special Education, 25*, 252-261.

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

Tilly, W. D., III (2008). The evolution of school psychology to science- based practice: Problem solving and the three-tiered model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 17–36). Bethesda, MD: National Association of School Psychologists.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*, 33-58.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency.* Austin, TX: Pro-Ed.

Treptow, M. A., Burns, M. K., & McComas, J. J. (2007). Reading at the frustration, instructional, and independent levels: The effects on student's reading comprehension and time on task. *School Psychology Review, 36*, 159-166.

Vadasy, P. F., Sanders, E. A., Peyton, J. A., & Jenkins, J. R. (2002). Timing and intensity of tutoring: A closer look at the conditions for effective early literacy tutoring. *Learning Disabilities Research and Practice*, *17*, 227-241.

VanDerHeyden, A. M., & Burns, M. K. (2005). Using curriculum-based assessment and curriculum-based measurement to guide elementary mathematics instruction: Effect on individual and group accountability scores. *Assessment for Effective Intervention*, *30,* 15–29.

VanDerHeyden, A. M.,Witt, J. C., & Gilbertson, D. A. (2007). Multiyear evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*, 225–256.

Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education, 24*, 301-315.

Videen, J., Marston, D., & Deno, S. L. (1982). Correct word sequences: A valid indicator of proficiency in written expression (Vol. IRLD-RR-84). University of Minnesota, Institute for Research on Learning Disabilities.

Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, *10*, 635-656.

Wann, J. P., & Jones, J. G. (1986). Space-time invariance in handwriting. *Human Movement Science*, *5*, 275-296.

Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students with low response to intervention. *Journal of Learning Disabilities, 41*, 126–142.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education, 41*, 85-120.

Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology, 101*, 1-20.

Ysseldyke, J., Burns, M.K., Scholin, S.E., & Parker, D.C. (2010). Instructionally valid assessment within Response to Intervention. *Teaching Exceptional Children, 42*, 54-61.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide.* New York: Touchstone, Applied Science.