A COMPARISON OF NARRATIVE AND EXPOSITORY TEXT COMPREHENSION FOR STUDENTS AT VARYING LEVELS OF SES: A LATENT GROWTH CURVE ANALYSIS

By

Laura Clark Briggs

A Dissertation Presented in Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in Literacy Studies

Middle Tennessee State University

May 2017

Submitted for Review to:

Jwa K. Kim, Ph.D., Committee Chair

Amy M. Elleman, Ph.D., Committee Member

Ying Jin, Ph.D., Committee Member

ACKNOWLEDGMENTS

Thank you, Jim, for your encouragement and support you've given me all these years. You have shouldered a huge amount of the burden. You are a wonderful dad and you've been a great Mr. Mom. You really do deserve a medal. I know it is difficult, and I cannot adequately express my gratitude for all that you've done to help me achieve this goal. I know it's been hard. I love you always and appreciate you so much. Tyler and Emily, I won the lottery when you two were born. Thank you for understanding that this was something that I had to do. You're the best, and I'm so honored to be your Mom. To my mother, Norma Prestwood, you have never for one minute doubted that I would eventually do this. I appreciate every single thing you've done for me over the years and I am so fortunate to have you in my corner. You are the one who reminds me that I told you more than thirty years ago that this was the plan. You believed it would happen even when I didn't. Thank you.

I would like to thank my dissertation committee members for their guidance, expertise, patience, encouragement, and wisdom. Dr. Kim, thank you for absolutely everything. You have been kind and generous with your time. Thank you for your sense of humor, especially at times when I really needed a good laugh. You will disagree, but there *is* crying in statistics. I will miss you. Thank you for being my committee chair. Dr. Elleman and Dr. Jin, thank you for serving on my committee. Your input and encouragement have been so valuable to me. I am in awe of your knowledge and expertise. You both have been so helpful at every step of this process. I am grateful and appreciative that you agreed to serve on my committee. Thank you to my professors in the Literacy Studies Program. It has been an honor to learn from so many intelligent,

dedicated, interesting, and passionate people. My heartfelt thanks and appreciation goes to Angela Morell. There are not enough words to describe just what you mean and to all of the students and faculty of the LITS program. My most sincere gratitude and appreciation is extended to the reference librarians at MTSU's James E. Walker Library. To Jacqueline Dowdy and to Leigh Watson, I thank you so much for all of the assistance you've given me these past four years.

I thank my fellow classmates. I appreciate your support and encouragement along the way. I value your friendship. Stacy, Summer, Zack, Weon, Perry, Anne, Tess, and Leticia, your time is coming soon! I cannot wait to cheer for you as you celebrate your success. Thank you to Joanne Veatch Coggins and to Tyra Williams Pickens. The Three Musketeers we are. I love you both.

Thank you to new friends made along the way. Marc and Mindy Liebman, thank you for opening up your home to Joanne and me while we completed our practicum in Utah. I miss y'all and I even miss the pointy mountains. I cannot wait to see you two again soon and to enjoy a good meal and great conversation with lots of laughter.

I thank the faculty, staff, current, and former students at The Academy at Hickory Hollow. You all have been the best support system a gal could ever hope for. I would especially like to thank my principal, William P. Fellman. You have always done everything you could to make it possible for me to balance full-time work with my studies, meetings, and research. As well, you have been a great source of encouragement for me. I want you to know how much I appreciate you for everything. I also thank one special graduate, Mohammad Anwar. What a happy coincidence that Jim's former middle school student became my high school student. Thanks to you, your mom, and

your sisters for babysitting Emily on those long Saturdays when Jim was in Philadelphia and I was in class all day long. Watching your self-discipline and your commitment to fitness and bodybuilding over the years has been an inspiration to me. All of the time you put in before the competitions. All the meal prep. All the workouts on days when you didn't really feel like doing it. I know you don't realize this, but you have taught your old teacher a lot about setting a goal and sticking to it.

Lastly, I want to thank my friends – my childhood friends, my Panama friends, my former-students-who-are-all-grown-up-now friends, my grown-people friends, my around-the-world friends, all of y'all. I have no idea how I ever got to be such a lucky girl to have the support, encouragement, well-wishes, and goodwill of so many people. None of this journey has been easy, but knowing that you were cheering for me every step of the way means more to me than you will ever know, and I love and thank you for that.

ABSTRACT

Research on secondary student reading comprehension performance is scant, yet demands for improved literacy at college and career levesl indicate that an understanding of trends and growth patterns is necessary to better inform teaching and learning for high school students. To improve understanding of reading performance at the secondary level, reading growth trajectories were investigated for 9^{th} (n = 5752) and 11^{th} (n = 3754) grade students. Free or reduced lunch membership (FRL) served as a proxy variable for student socioeconomic status (SES). Item performance of narrative and expository text was examined based on SES status by trend analysis and by latent growth curve anaylsis (LGCA) to determine if SES impacts initial starting point and growth on reading comprehension. Results revealed linear and quadratic trends of reading comprehension growth for 9^{th} and 11^{th} grade students. The dominant linear trend for 9^{th} grade performance suggests that performance improved throughout the academic year. The dominant quadratic trend for 11^{th} grade performance indicates that student performance declined at the second test administration before improving at the third test. Performance on English I expository tests showed a negative intercept-slope relationship indicating that students who scored lower initially performed better on subsequent exams compared to those who scored higher initially. The positive SES-intercept impact suggests that SES is correlated with performance on initial test administration. The negative SES-slope impact suggests that, during the academic year, student SES does not correlate with comprehension growth, possibly owing to the equalizing effects of the school environment on student achievement (Alexander, Entwistle, & Olson, 2007; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Entwistle & Alexander, 1992, 1994;

Heyns, 1978; Jamar, 1994; Pfost, Hattie, Dörfler, & Artelt, 2014). Performance on narrative items revealed a positive relationship on the intercept and slope as well as a positive impact for both SES-intercept and SES-slope. For English III scores, results indicated positive relationships on intercept-slope, SES-intercept, and SES-slope for expository items. Due to the poor model fit for the narrative models, impact and relationship among these variables could not be determined.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The National Assessment of Education Progress (NAEP), also referred to as the Nation's Report Card, has monitored student reading comprehension progress and achievement in national trend assessments of 9-, 13-, and 17- year-old students every four years since 1971. In the reading series, students are assessed based on their performance on multiple choice items of comprehension tasks that require them to evaluate and locate information from both narrative and expository texts. Since the inception of the NAEP, reading scores on these national long-term trend assessments have revealed of gains for 9- and 13-year-old students; however, results for 17-year-olds have reflected no significant change (National Center for Education Statistics, 2013). A reading component of the NAEP began in 1992 and is administered every two years, assessing students using a variety of question formats including multiple choice, short answer, and extended answer (National Center for Education Statistics, 2013). The most recent results from 2015 show declines in scores for eighth graders across all demographic groups. Among twelfth graders, only 37% of test takers achieved a score at or above 302 – the magic cut score that is indicative of a student's ability to handle the reading demands of college according to the National Assessment Governing Board's preparedness research (NAEP, 2015). Despite the report's optimistic language, emphasizing the fact that 34% of eighth grade students and that 37% of twelfth grade students are at or above proficient, these percentages must be of little comfort for the 66% of eighth grade and 63% of twelfth grade students that are *not* proficient.

Poverty and the Persistence of the Performance Gap

Within grade groups, a performance gap between students based on parental income or on free- and reduced-price lunch status (referred to as FRP or FRL) stubbornly persists and has in fact widened (NAEP, 2015). FRL is commonly used as a proxy for student and family socioeconomic status information despite concerns about its accuracy as a measure (National Forum on Education Statistics. 2015; Snyder & Musu-Gillette, 2015). Among eighth grade students, only 21% of those eligible for FRL scored proficient or better compared to a proficient/at proficient score for 53% of students who were not FRL eligible. In twelfth grade, NAEP data showing parental level of education are available and are suggestive of student and family socioeconomic status. Achievement scores from 2015 reveal that only 18% of students whose parents did not finish high school scored at or above proficient, compared to 49% scoring at or above proficient for students whose parents were college graduates (NAEP, 2015). For the majority of the nation's eighth and twelfth grade students, reading remains a struggle; but for students who come from lower-income homes, the achievement gap between them and those from higher-income households is staggering.

The Coleman Report (Coleman et al., 1966), ordered as a part of the Civil Rights Act, marked a beginning in the systematic collection and use of testing data to identify gaps in achievement between groups of students (Viadero, 2006). One of the most striking and disturbing assertions was that the school itself played very little role in equalizing the academic outcomes for students who were disadvantaged as a result of race of poverty (Ravitch, 1981). The importance of the Coleman Report cannot be understated in that it was the genesis of sharing educational survey data that would then

go on to shape policy. The numbers were available for all to see, and this ushered in an era of reporting achievement data on a large scale, including NAEP data, first released in 1971. The data from NAEP paint a picture of students who still struggle with reading – especially in the later grades, and the gap between students in poverty and those who are not has persistently remained despite decades of research and valiant efforts to effect positive change.

The Performance Gap, Standardized Testing, and Benchmark Assessments

Although standardized testing in the United States has existed since the mid- 19th century (U.S. Congress, 1992), the Coleman Report's findings pointed out the great disparity among students from higher- and lower-income families. It also highlighted the importance of standard achievement assessments for measuring skills that would afford graduates sound career opportunities (Coleman, 1966). A year before the release of the Coleman Report, the Elementary and Secondary Education Act (ESEA) instituted programs backed by federal dollars including Title I, with mandates that these programs be evaluated; and with this, the use of standardized assessments to measure progress became commonplace (Hamilton & Koretz, 2002). The advent of the NAEP in the early 1970s and the release of *A Nation at Risk* in 1983 revealed stark facts about functional illiteracy among a large percentage of the nation's youth and about the country's fall from its reputation as a powerful, preeminent nation as indicated by several criteria including literacy rate, performance on national assessments such as the SAT, and the need for college students to take remedial coursework (Gardner, 1983). Later, as a result of the No Child Left Behind Act (NCLB), standards-based reform ushered in high-stakes

testing as an accountability measure with new attention paid to subgroup performance, an area that often did not receive as much examination under state accountability systems (Lauen & Gaddis, 2011). Benchmark assessments have emerged as a result of increased emphasis on high-stakes testing. These assessments are typically administered at several time points throughout the school year, focusing on areas of reading and math to determine student progress through the curriculum or standards (Olson, 2005).

An Argument for National Standards: Common Core State Standards

At the secondary school level, reading typically falls squarely on the shoulders of the English Language Arts (ELA) teacher with the presumption being that students have already learned the requisite skills to be successful readers throughout their elementary school years. However, in order for students to successfully handle the reading demands placed on them at the college level, they must have strong vocabulary knowledge attained from subjects aside from ELA courses such as science, math, and social studies (ACT, 2005). Many students who have not been challenged by more rigorous and complex texts find themselves without the tools or skills to handle the demands of "text-heavy" courses and thus are precluded from access to more academically challenging courses. As such, literacy behaves as a gatekeeper for many. The Common Core State Standards (CCSS) were developed, in part, to address the literacy crisis in the United States with the launch of the initiative in 2009. At its inception, leaders from 48 states as well as leaders from two U.S. territories and the District of Columbia were involved in the launch of the initiative (National Governors Association, 2010) which aimed to raise student performance expectations significantly in order for students to be sufficiently prepared

for the reading demands that would be placed on them in both the college and the workplace.

The CCSS design for ELA incorporated an increasingly heavier load of informational[1] (also referred to as *explanatory* or *expository*) text, but stressed that the responsibility for teaching with informational text was a shared one among disciplines and that the increase in exposure to informational text was in keeping with the National Assessment of Education Progress (NAEP) framework (National Governors Association, 2010). Citing the NAEP framework, the National Governors Association displayed the distribution for literary/informational text ratio for grades 4, 8, and 12 as follows: for fourth-grade students, 50%-50%; for eighth-grade students, 45%-55%, and for twelfth-grade students, 30%-70%.

Narrative and Expository Texts

Text varies according to type. In its most general form, the terms narrative and expository refer to a binary of sorts that distinguishes text whose purpose is to narrate from text whose purpose is to convey information or to explain. Narrative text is typically thought of as fiction which contains elements such as setting, characters, inciting event(s), conflict, and resolution. Examples of expository text are many and can include the type of content found in textbooks, manuals, cookbooks, or reports. The prevalence of literature-based classrooms in the 1980s and 1990s (Applebee, 1993; Brody, DeMilo, & Purves, 1989; Close, 1992; Jipson & Paley, 1992; Langer, 1990; 1991; 1992; Morrow, 1992; Walmsley & Adams, 1993), the insufficient instructional time

[1] Informational text is a subcategory of nonfiction text that informs the reader about the natural or social world (Duke & Bennett-Armistead, 2003). Informational text is also referred to as *expository* or *explanatory* text.

devoted to teaching with expository text (Duke, 2000a; Maloch & Bomer, 2013), and the persistence of a decades-old call to increase the amount of expository text used in classrooms (Applebe, 1984; Chall & Jacobs, 2003; Kamil & Lane, 1997; Maloch & Bomer, 2013; Newkirk, 1989; Pappas, 1993) precluded the deliberate focus on teaching with expository text found in the CCSS (Maloch & Bomer, 2013; National Governors Association, 2010). There is not consensus that expository text is more difficult than narrative text (N. K. Duke, personal communication, October 29, 2016; Williamson, 2008; Williamson, Fitzgerald, & Stenner, 2014); however, characteristics including text structure, content-specific specialized vocabulary, and concept density that may appear within expository text can present challenges for struggling readers or for readers who do not possess strong background knowledge about text topics (Beyer, 2007; Hall, 2004; Jennings, Caldwell & Lerner, 2006; McCormick, 2007). Additionally, the scarcity of expository texts in classrooms results in limited opportunities for practice and familiarity with the wide range of texts, particularly for students who are low-SES (Duke, 2000a; 2000b). These limited opportunities hinder expansion of student background knowledge which playes a crucial role in reading (Marzano, 2004; Recht & Leslie, 1988).

Measuring Reading Comprehension

Reading comprehension is complex, multifaceted, and comprised of multiple intertwined elements (Ruddell, 2001; Scarborough, 2002). This complexity results in various approaches to measurement and assessment of reading constructs at both the word and text level that are often constrained by formats dictated by high-stakes testing (Fletcher, 2006). Often large scale assessment creates the need for student test responses

to be captured in short forms to include cloze responses or in dichotomous or polytomous multiple-choice formats. Because reading comprehension is a latent trait, measuring precisely what it is that a student knows is difficult: measurement is constrained by the nature of comprehension in comparison to measurement at the word level, where processes of understanding are more directly observed (Fletcher, 2006, RAND Reading Study Group, 2002). Text type, individual differences, and test response format are among salient and prominent issues when considering assessment in reading comprehension (Fletcher, 2006).

It is important to determine not only how well students understand skills and concepts at one particular time point, but also whether students demonstrate growth over time. Benchmark assessments provide a vehicle for determining student concept mastery at various points throughout the year, and teachers and administrators gain valuable insight in to individual student progress on skill mastery (Betts et al., 2008; Graney, Missall, Martinez, & Bergstrom, 2009; Pickens, 2016). For longitudinal data such as benchmark assessments, where data are collected over several time points, repeated measures ANOVA may be used to test equality of mean, and polynomial contrasts reveal the shape of growth trajectories using trend analysis. Such contrasts check for trends between groups, revealing either linear or nonlinear trends. Linear trends reveal that the dependent variable change is proportionate across categories. Nonlinear trends may consist of quadratic, cubic, or quartic trends, and may depend on the number of time points evaluated. For tripartite data, which is frequently used in elementary and secondary assessment, results may reveal linear or quadratic trends characterized by the number of changes in the direction of the line (Field, 2013). The direction of the trend in

reading allows researchers to determine whether performance in reading comprehension improves over the academic year. The relationship between a student's starting point and rate of change is important to examine because it student initial performance is impacted by prior knowledge and experiences (Vacca & Vacca, 1999). In order to further discern the relationship between the starting point (intercept) and rate of changes (slope) of the growth, trend analysis may be followed up with latent growth curve modeling (LGCM). LGCM is an application of structural equation modeling (SEM) whose advantages include evaluating growth over time (Aycock & Li, n.d.; Duncan & Duncan, 2009). LGCM is suited for analysis within the behavioral sciences and education and allows the researcher to determine the initial starting point of the test taker (termed *intercept*) and to evaluate the growth exhibited longitudinally and between time points (termed *slope*) as well as to examine whether the starting point has effects on the shape or the steepness of the slope. Having this information allows researchers to answer questions such as whether a student's prior knowledge and learning experiences correlate with later growth throughout the academic year or whether students of varying SES are impacted differently by the noninstructional period preceeding the initial test examination. As such, LGCM is an apt model for analyzing trends of student performance on benchmark assessments.

Purpose of the Study and Research Questions

While the number of studies concerning elementary-level student growth trajectories of reading exceed those considering students at the secondary level, there is scant literature providing researchers and practitioners with information about secondary

students' growth and performance on measures of reading comprehension. It is at this age that students are making decisions about post-secondary plans, so it is therefore critical to examine how these students fare on assessments gauging reading comprehension skill. There is no disputing that students will encounter expository texts both in the college and in the workplace setting, but there is considerable debate in the field about whether differences even exist in difficulty between narrative and expository texts, although studies suggest that cultural background factors in to comprehension of text (Bell & Clark, 1998; N. K. Duke, personal communication, October 29, 2016). Socioeconomic status is widely believed to strongly correlate with academic achievement although meta-analyses have revealed correlations varying from weak to moderate (Hattie, 2008; White, 1982). The present study will examine following questions:

1. What is the shape of trajectories of high school students' performance on test items of reading comprehension?
2. What is the relationship between starting point (intercept) and rate of change (slope) for narrative text test items?
3. What is the relationship between starting point (intercept) and rate of change (slope) for expository text test items?
4. What is the impact of SES on intercept and slope for test items of narrative text?
5. What is the impact of SES on intercept and slope for test items of expository text?

Hypotheses

The hypotheses corresponding to the above research questions are:

1. Trajectories for overall reading comprehension will be nonlinear

2. There will be a significant relationship between students' initial starting point (i.e., the initial test score, *intercept*) and the rate of change (i.e., the growth, *slope*) on narrative test scores.
3. There will be a significant relationship between students' intercept and the slope on expository test items.
4. SES will have no significant impact on intercept and slope for test items of narrative text.
5. SES will have a significant positive impact on intercept and slope for test items of expository text.

CHAPTER II

LITERATURE REVIEW

This chapter provides a review of research on how elementary and secondary students perform in areas of reading comprehension. However, before delving in to research on growth patterns and trajectories of reading comprehension, it is first necessary to discuss how reading comprehension is defined and to explore theoretical models of comprehension in order to better understand the complexities of measuring this latent characteristic. Furthermore, it is helpful to understand how reading develops from being a pre-reader to a mature and skilled reader. This review includes six sections: the first section describes the difficulties of defining reading comprehension with precision, and the second section presents background on reading is assessed at the elementary and secondary levels. The third section describes the genesis and development CCSS and lists standards that pertain to ELA literature and informational texts. The fourth section of this review describes the origins of benchmark testing and how those tests are used in the school setting to determine student progress. The fifth section describes the relationship between student SES and reading achievement. The last section of this review examines trajectories of reading growth, providing a synthesis of research on growth patterns exhibited for students at both the elementary and secondary level. There are numerous studies concerning the growth trajectories of younger readers, however there are few in comparison regarding reading comprehension growth of adolescent readers. Intricacies and complexity of measuring reading comprehension may be one contributing factor to why such few studies at the secondary level exist.

Defining Reading Comprehension

What is reading comprehension? To the layperson, a suitable definition might be: *reading comprehension means that you understand what you read.* However, this may lead one to question what *understanding* means. Are there layers of understanding? Is it possible to understand something on a surface level, yet not have a deep, full understanding of a concept? Likewise, is it possible to understand the words read but not to understand the concept embodied by those words? The multifaceted nature of reading comprehension lends itself to interpretation on many different levels and with emphasis placed on various components. One frequent definition of comprehension is that it is the process of obtaining and creating meaning through interaction with text (Rand Reading Study Group [RRSG], 2002). The frequently-cited research of Gough and Tunmer (1986) stated in their Simple View of Reading (SVR) that reading comprehension was the result of decoding and language comprehension working in concert, but others lean toward a more complex view. Ruddell and Unrau (2004) speak specifically to the difficulty of reading model development: "[O]ur limited capacity to observe, measure, collect information, and describe processes precisely limits the accuracy of a reading or writing model" (p. 1116).

Assessing Reading

Assessing reading differs depending on the age of the reader with elementary-grade students assessed primarily on decoding, fluency, accuracy, word attack, story retell, word recognition, and passage recognition skills (Rathvon, 2004). As readers progress from learning to read to reading to learn (Chall, Jacobs, & Baldwin, 1990; Chall

and Jacobs, 2003), around the beginning of fourth grade, assessment typically consists of vocabulary knowledge, ability to infer meaning from text and ability to locate information within text to successfully answer questions. McKenna and Stahl (2009) describe models for reading assessment as roadmaps used "to navigate the assessment procedure" (p. 2).

Rathvon (2004) suggests that "[p]erhaps no other domain of reading encompasses such a wide range of test formats" (p. 163) as reading assessment. In elementary grades, formats focus on tasks of alphabet recognition, phonemic awareness, spelling, word recognition, phonics, comprehension, and oral reading fluency (Bell & McCallum, 2008) whereas later elementary and secondary assessment typically focuses on comprehension, a skill that is more difficult to quantify (Smith, 2004; Stahl & McKenna, 2009). By convention, levels of comprehension questions involve questions at the literal, inferential, or critical level (McKenna & Stahl, 2009). Assessment may involve various formats for assessing reading comprehension at the later elementary and secondary levels. Such formats include cloze assessment, where words within a passage are deleted requiring the test-taker to provide the appropriate word to complete the passage, maze tasks (Guthrie, Seifert, Burnham, & Caplan, 1974), and multiple choice test (MCT) formats. The present study concerns assessment via multiple choice format. Although this format has been the topic of debate (Feinberg, 1990; Heick, 2013; McCoubrie, 2004), MCTs offer advantages in their versatility, reliability, and validity (Brame, 2013; Haladyna, Downing, & Rodriguez, 2002).

Common Core State Standards

In 2010, the Council of Chief State School Officers (CCSSO) and the National Governors Association Center for Best Practices (NGA Center) jointly announced the launch of the Common Core State Standards (CCSS) for English language arts (ELA) and mathematics (National Governors Association, 2010). These standards were the written at the state level by workgroups represented by educators at the primary, secondary, and higher education levels as well as by other experts and were subject to public input at the time of standards development (National Governors Association, 2010). According to Conley (2014), the rationale of common standards was to address the disparities that resulted from the states' historical lack of consistency regarding academic expectations and rigor. The advent and adoption of common standards among states would allow for a framework that would guarantee that students could graduate high school with skills and knowledge that would benefit them in an increasingly complex and fast-changing society and economy, and the commonality of the language within the standards would guarantee that a skill learned in one part of the United States would be taught using the same language and with similar depth and breadth in another part of the nation, thus ensuring that the academic equity for students throughout the country.

English language arts standards were developed for use within social studies, science, and technical courses (National Governors Association, 2010) with the understanding being that reading takes place outside of the English class as well as within. The CCSS include College and Career Readiness (CCR) standards which are fundamental skills that a student should possess upon graduating high school. These

CCR standards are termed “anchor standards”. Within English language arts, anchor standards exist for reading, writing, speaking and listening, and language. Reading standards are focused on four aspects of reading and contain subskills within each aspect. These standards are contained in Appendix A.

In order for students to handle the reading demands of postsecondary education, CCSS encouraged students to increase depth and breadth of knowledge by reading across subject areas, emphasizing that the English classroom is not the sole source of reading, thus stressing the need to prepare students for post-secondary literacy demands including the ability to comprehend informational texts frequently encountered in the workplace, as embodied in the language of the CCR anchor standards (National Governors Association, 2010).

Benchmark Tests

In 1991, President Bush issued an imperative for transformation in America’s schools. To track progress toward national goals, the imperative called for system of strict accountability via administration of national assessments for 4^{th}, 8^{th}, and 12^{th} grade students in areas of English, math, science, history, and geography (Bush, 1991; Shavelson, Baxter, & Pine, 1992). This call for national testing differed significantly from the historical perspective that education was the purview of individual states (Shavelson et al., 1992). Coupled with the No Child Left Behind Act (NCLB) of 2001, this imperative marked a transformation in testing, leading to a market for large-scale national assessments designed to examine student achievement at the group and sub-group level (Lauen & Gaddis, 2012). Although the Bush imperative of 1991

recommended reexamining test format and measures, cost and efficiency remained a practical consideration for test construction with the economy and utility of multiple-choice test format as driving deciding factors (Shavelson et al., 1992). Although benchmark testing in elementary and secondary grades has existed since the 1970s (Tindal, 2013), the increased national emphasis on high-stakes testing has increased the demand and use of these assessments. Benchmark assessments are used to track a student's progress and are commonly administered at three time points throughout the school year, typically at the beginning, middle, and close of the school year (Graney et al., 2009; Olson, 2005, Pickens, 2016).

Socioeconomic Status and Reading Ability

Student socioeconomic status (SES) has been traditionally defined as a composite of parental occupation, education level, and income. SES has many correlates (Hart, Soden, Johnson, Schatschneider, & Taylor, 2013) including student academic achievement (Coleman et al., 1966; Cowan et al., 2012; Morgan, Farkas, & Hibel, 2008; Sirin, 2005; Thomas, 1962; White, 1982). Numerous studies have found that SES is a significant predictor of reading comprehension (Bowers, 2012; Cooper, Charlton, Valentine, & Muhlenbruck, 2000). Low SES background has been identified a one risk factor for reading difficulties (Aikens & Barbarin, 2008; Beecher, 2011; Ghosh, 2013; Rathvon, 2004; Whitehurst & Lonigan, 1998) and is one variable that is related to the likelihood of being diagnosed with a learning disability (Judge & Bell, 2011). Low-SES students who enter school as poor readers rarely catch up to their non-disadvantaged peers (Torgesen, 1997) and encounter difficulty extracting meaningful information from text (Bowers, 2012; Myers & Botting, 2008). Students from low-SES environments experience greater risk of summer academic achievement loss due to the non-instructional period referred to as summer effects or summer setbacks (Alexander, et al., 2007; Cooper, et al., 1996; Entwistle & Alexander, 1992, 1994; Heyns, 1978; Jamar, 1994).

Research Findings of Student Trajectories of Reading Comprehension

The paucity of research studies examining the growth trajectories of secondary students' reading comprehension achievement underscores the need for research specific to students preparing to enter college and the workforce. This section provides a review

of the literature surrounding student reading achievement growth models first by examining patterns for elementary-age students followed by a review of the limited available literature evaluating growth patterns for secondary level students. Students in the early elementary grades typically display a pattern of higher reading gains in comparison to students in the later elementary and secondary grades where gains decrease (Pfost et al., 2014). Three patterns emerge from examining student reading growth: the fan spread effect (referred to as the Matthew effect), the compensatory (or, developmental) lag model, and the stable achievement pattern. The Matthew effect (Stanovich, 1986; Walberg & Tsai, 1983) refers to the pattern exhibited when comparing good readers and poor readers, good readers continue to exhibit positive reading gains while poor readers continue to struggle and demonstrate negative growth, creating a cycle in which poor readers are unable to catch up to their peers (Stanovich, 1986, 2000; Walberg & Tsai, 1983; Pfost et al., 2014). Studies have shown that the Matthew effect in reading is more pronounced during non-instructional periods than during the school year, suggesting that instructional periods have an equalizing effect on student reading achievement (Alexander et al., 2007; Cooper et al., 1996; Entwistle & Alexander, 1992, 1994; Heynes, 1978; Jamar, 1994; Pfost et al., 2014). The compensatory lag model is characterized by a decrease in the achievement gap between strong and poor readers. This model suggests that poor readers will catch up to their peers as they develop and that this is the result of a delay rather than an ability deficit (Scarborough, 2002). Paris (2005) suggests that developmental patterns are present for highly constrained reading skills such as those learned in early elementary grades (i.e., letter knowledge, phonics) compared to less constrained skills such as comprehension. Patterns that assume that the

gaps between strong and poor readers are relative and unchanged among groups are referred to as stable achievement gaps (Pfost et al., 2014).

Differing trajectory shapes. Reading achievement differs depending on the age of the reader. Measurement of oral reading fluency (ORF) is frequently used to evaluate reading achievement in the early elementary grades with research supporting the correlation of curriculum-based measures (CBM), measures "indexing academic competence and progress" (Deno, Fuchs, Marston, & Chin, 2001, p. 508) of ORF with reading skill (Allinder, Fuchs & Fuchs, 1998; Pickens, 2016; Rathvon, 2004; Shinn, Good, Knutson, & Tilly, 1992). However, reading comprehension in high school is frequently assessed via cloze passages or maze assessment (McKenna & Stahl, 2009). Due to the reading skill assessed, in combination with student age and ability, trajectories of student reading achievement vary between students in elementary and secondary grades (Deno et al., 2001; Nese et al. 2013, Pickens, 2016).

Inclusion criteria. Studies examining reading growth in the elementary grades is abundant, yet research specifically investigating reading comprehension growth among secondary students is meager (Espin, Wallace, Lembke, Campbell & Long, 2010; Pickens, 2016). Therefore, a systematic examination of the literature was conducted to survey the patterns of reading comprehension across both elementary and secondary grades. Francis, Shaywitz, Stuebing, Shaywitz., and Fletcher (1996) suggest that reading ability changes non-linearly, eventually stabilizing (Beecher, 2011), underscoring the need to add to the body of knowledge by analyzing growth patterns in order to understand how reading achievement is revealed in the latter years of a student's educational experience. In order to locate relevant studies at the elementary and

secondary school level, an electronic search was conducted for quantitative, peer-reviewed studies using the following databases: ERIC, GoogleScholar, JSTOR, ProQuest, PsycINFO, ScienceDirect, Scopus, and Web of Science. Studies evaluating reading growth for languages other than English as well as studies not conducted in the Unites States, unpublished dissertations, case studies, and book reviews were excluded. Studies focused on growth patterns exclusively for second language (L2) learners were also excluded. Search terms including *Matthew(-)effect, developmental lag, cumulative and deficit, inter(-)individual differences* were used in combination with *literacy, reading* and *growth* or *traject** as well as *elementary* and *secondary* or *high school*. Finally, citations in the articles meeting inclusion criteria were also examined. All literature searches were conducted in the fall of 2016. The focus of the present study is reading trajectories of adolescent readers; as such, elementary reading growth patterns were presented chronologically to provide contextual background followed by a review of secondary reading growth trajectories.

Reading trajectories in the elementary grades. Research reveals mixed findings concerning reading growth patterns (i.e. linear and quadratic, presence and absence of Matthew effects, presence or absence of developmental lag). Juel (1988) conducted a longitudinal study involving students in grades one through four (n = 54) to examine whether students with reading problems remained poor readers year after year and to locate the source of reading difficulty among those students. Findings detected the presence of Matthew effects, determining that students entering school with poor phonemic awareness would later fail to master decoding as they entered second grade. It was found that children finishing their first grade year as poor readers had a .88

probability of remaining poor readers at the end of fourth grade compared to average readers, who only had a .12 probability of becoming poor readers at the end of fourth grade. The probability of average first grade readers remaining average readers by the end of fourth grade was .87. A later measurement study conducted by Williamson, Appelbaum, and Epanchin (1991) utilized individual growth modeling to examine trajectories of reading achievement from a student cohort (n =529) beginning in first grade and continuing through the end of their eighth grade year. Results from this study revealed high variability suggesting the presence of Matthew effects as were evidenced in Juel (1988). The Williamson et al. study employed use of growth modeling for longitudinal data for measuring academic achievement that has gained increasing popularity over the past two decades and is important for inclusion in this review.

The research of Fuchs, Fuchs, Hamlett, Walz, and Germann (1993) examined weekly reading growth among two samples of students in grades 1-6 during a two-year period using weekly CBM for oral reading in the first year of the study (n =117) and CBM for maze tasks for the second year (n =2257). In this study researchers used a least-squares regression to determine the slope, representing average weekly student growth followed by one-way analysis of variance (ANOVA) to examine the effect of grade level on the slope. ANOVAs were followed up by polynomial contrasts to determine the shape of student growth trajectories (i.e., linear or quadratic) of the two groups. Student growth for the first year of the study was exhibited by mostly linear trajectories for students, with 0-21% significant quadratic terms for CBM oral reading measures and 19-31% of quadratic terms for CBM maze measures. The second year of the study revealed an overall linear relationship for student growth on both the CBM oral

reading and maze measures. These findings contrast the fan-spread trajectories indicative of Matthew effects that were found in the Juel (1988) study.

The study of Francis et al. (1996) analyzed data from the Connecticut Longitudinal Study (CLS) the growth patterns of 403 children and stressed the importance of developing strong reading skills in the early grades. Growth trajectories revealing long-term development between grades 1-9 indicated that third grade poor readers remained poor readers as ninth graders and that growth patterns were suggestive of a deficit, rather than a developmental lag (Foorman, Fletcher, Francis, Schatschneider, & Mehta, 1998).

Several studies examined reading growth spanning the elementary and high school grades. The 2004 study conducted by Rescorla and Rosenthal is one such study. Researchers studied the reading growth students (n =328), beginning with measurement of reading ability at 3rd grade and then employing LGA to examine growth at 5th, 8th and 10th grade and to determine growth across time relative to students' initial status, grouped by high, medium, and low reading achievement). Findings from their study revealed that students in all progressed at similar rates, but displaying a fan-close effect at the 10th grade time point.

The use of CBM to evaluate growth over time has gained popularity over the past two decades and has been used to monitor student growth (Ardoin & Christ, 2008; Christ & Silberglitt, 2007; Deno et al., 2001; Fuchs et al., 1993; Good & Kaminski 2003; Graney et al., 2009; Hasbrouck & Tindal 2006; Powell-Smith, Shinn, Stoner, & Good, 2000). Findings from Fuchs et al. (1993) and Deno et al. (2001) revealed linear growth,

but other studies have detected nonlinear growth rates (Ardoin & Christ, 2008; Christ & Silberglitt, 2007; Nese et al., 2012).

Reading trajectories in the secondary grades. There are very few studies that have examined high school reading trajectories (Beecher, 2011; Petscher, Kershaw, Koon, & Foorman, 2014; Pickens, 2016; Rescorla & Rosenthal, 2004), and some of the research is embedded within studies evaluating elementary reading growth patterns (e.g., Beecher, 2011; Petcher et al., 2014; Rescorla & Rosenthal, 2004). Beecher (2011), Petscher et al. (2014), along with Rescorla and Rosenthal (2004) detected non-linear growth patterns whereas Pickens (2016) reported linear trends of reading performance. The research of Rescorla and Rosenthal, described in the previous subsection, revealed an absence of Matthew effects for 10th grade students' reading performance. Beecher (2011) evaluated archival data of longitudinal reading performance for students as they aged from 7 to 19. All students in the study (n 206) had been initially referred for preschool interventions for developmental delay. Of the 206 students, roughly two thirds received special education services following the preschool intervention. Reading achievement was measured yearly using the Peabody Individual Achievement Test (PIAT), a suitable measure for longitudinal research due to the instrument's continuous scale of increased difficulty (Beecher, 2011). Beecher used LGC analysis to evaluate student PIAT scores on the reading recognition (RR) and reading comprehension (RC) subtests along with post-intervention scores on the McCarthy Scales of Children's Abilities (MSCA)' general cognitive index (GCI) and verbal scales to determine the longitudinal growth pattern of reading achievement. Predictor variables included McCarthy GCI, McCarthy verbal, disability (represented by preschool special education intervention status), student SES

(represented by FRL membership), and gender. Findings suggested an overall decline in performance on the RR and RC subtests. Although individual growth curves revealed trends that were both linear and quadratic, the RC scores tended to be more quadratic, reflecting the differences between RR (decoding) and RC, a less-constrained skill (Paris, 2005). Results indicated a quadratic trend with the patterns of students who were in need of special education services at the onset of school (in the preschool intervention) revealing a more pronounced downward growth slope. This study also indicated a lower starting point for students receiving special education services than for those who did not, and although a difference between the two groups was evident, the gap between the two groups did not widen substantially, suggesting a deficit in skill that persists (Francis et al., 1996).

Pickens (2016) analyzed data from benchmark assessments over a two-year period for 8th through 10th grade students (n 225) to evaluate growth patterns and to evaluate the impact of race as well as SES (dichotomized on FRL membership) on reading comprehension growth. The shape of student reading trajectories was determined by use of trend analysis and then followed by LGC analysis to examine the relationship of student race as well as SES on reading ability. Results indicated differed from the Beecher (2011) study in that an overall linear growth with similar growth was displayed among groups regardless of race or socioeconomic status. To restate, the present study will examine following questions:

1. What is the shape of trajectories of high school students' performance on test items of reading comprehension?

2. What is the relationship between starting point (intercept) and rate of change (slope) for narrative text test items?
3. What is the relationship between starting point (intercept) and rate of change (slope) for expository text test items?
4. What is the impact of SES on intercept and slope for test items of narrative text?
5. What is the impact of SES on intercept and slope for test items of expository text?

CHAPTER III

METHOD

Participants

The archival data used in this study came from a total of 9506 participants in six states and collected from English I and English III students' benchmark assessments administered over three time waves during the 2014-2015 academic school year. Only scores for students who participated in assessments over the three waves were used for this analysis. Student demographic variables for SES indicated by free or reduced lunch (FRL) status were considered, so missing data in this category were excluded from the analysis. Student SES is dichotomized in to two groups, with low-poverty students represented by FRL membership. For the 2014-15 fiscal year (July 1, 2014 through June 30, 2015), students from a four-member household with an annual income of $31,005 or less qualified for free school. Students who lived in a four-member household with an annual income between $31,006 and $44,123 qualified for reduced-priced school meals (Income Eligibility Guidelines, 2014). From this, a dichotomous dummy variable was generated for the analyses. Demographic data include missing values on areas of gender, race, English as a Second Language (ESL) status, and special education status; however, scores for students with missing data in these areas were retained because the focus of the present study is the analysis of student performance based on FRL status. For English I students, there were a total of 5752 data points for gender with 2880 females (50.1 %) and 2872 males (49.9 %). Of the data analyzed, 1980 (34.4%) reported to be African American, 3133 (54.5%) Caucasian, 67 (1.2%) Native American, 150 (2.6%) Asian, 307

(5.3%) Other. There were missing race values for 115 (2.0%) of the students. For English III students, there were a total of 3754 data points for gender with 1831 (48.8%) females and 1923 (51.2%) males. Of the demographic data analyzed for English III students, 1616 (43.0%) reported to be African American, 1729 (46.1%) Caucasian, 19 (0.5%) Native American, 96 (2.0%) Asian, 224 (6.5%) Other. There were missing race values for 70 (1.9%) of the students. Table 1 provides demographic information for students participating in the English I assessment. Student demographic characteristics for English III participants are displayed in Table 2.

Table 1

Demographic Characteristics of English I Participants (n = 5752)

	N	%
Gender		
Male	2872	49.9
Female	2880	50.1
Race		
African American	1980	34.4
Caucasian	3133	54.5
Native American	67	1.2
Asian	150	2.6
Other	307	5.3
Missing values	115	2.0
FRL Status		
Yes	3301	57.4
No	2451	42.6

Note. FRL = Recipient of Free/Reduced Lunch. FRL serves as dichotomous proxy variable for student socioeconomic (SES) status. FRL: yes = low SES; FRL: no = med/hi SES.

Table 2

Demographic Characteristics of English III Participants (n = 3754)

	N	%
Gender		
Male	1923	51.2
Female	1831	48.8
Race		
African American	1616	43.0
Caucasian	1726	46.1
Native American	19	0.5
Asian	96	2.0
Other	224	6.5
Missing values	70	1.9
FRL Status		
Yes	2062	54.9
No	1692	45.1

Note. FRL = Recipient of Free/Reduced Lunch. FRL serves as dichotomous proxy variable for student socioeconomic (SES) status. FRL: yes = low SES; FRL: no = med/hi SES.

Measurement

Items assessed were developed by a for-profit testing company in the United States based on CCSS standards for English Language Arts (ELA). CCSS reading and language standards include: informational text (RI), language standards (L), literature (RL), and writing (W) strands, but test items assessing language and writing strands were excluded from this analysis as student performance on narrative and informational text are the items of interest for this study. Information regarding strands for RI and RL is provided at the Core Standards site (http://www.corestandards.org/ELA-Literacy/) and is

described in Appendix B. Standards for 9th grade are grouped with 10th grade ELA standards, and 11th and 12th grade standards are likewise grouped.

Tests were administered tri-annually at the fall, winter, and spring of the academic school year. Each test contained 32 multiple-choice items assessing a range of categories measuring strands from the CCSS ELA curriculum including items of both narrative and informational text. Individual student information was de-identified prior to releasing the data for analysis with an anonymized ID being assigned to each test-taker.

English I assessments. For English I assessments, test form A was administered in the fall of 2014. Test forms B and C were administered in the winter and spring of 2015, respectively. Cronbach's alphas for Forms A, B, and C are .84, .87, and .88, respectively. Items assessing writing (W-coded) and language (L-coded) were not included in this study and were excluded from analysis. Test form A consisted of nine expository items (RI-coded items 18, 19, 20, 21, 22, 23, 25, 27, and 29) and ten narrative items (RL-coded items 1, 2, 3, 5, 12, 13, 14, 16, 31, 32). Because form A only contained nine items, a conversion factor of 1.11 was used. Form B consisted of nine expository items (RI-coded items 9, 10, 11, 12, 17, 18, 25, 26, and 27) and ten narrative items (RL-coded items 1, 2, 4, 5, 7, 8, 23, 24, 28, and 29). As with form A, a conversion factor of 1.11 was used to account for the unequal amount of test items on form B. Test C contained ten items for both expository and narrative texts (RI-coded 5, 6, 7, 8, 9, 13, 14, 15, 27, and 28 for expository; 2, 3, 4, 10, 11, 19, 22, 30, 31, and 32 RL-coded for narrative).

English III assessments. English III assessments were administered in three waves at the following intervals: fall, 2014, winter, 2015, and spring, 2015. Cronbach's

alphas for Forms A, B, and C are .84, .84, and .88, respectively. All test forms consisted of 32 items, but items assessing writing (W-coded) and language (L-coded) were excluded from analysis. Test form A (fall administration) consisted of ten expository items (RI-coded 11, 12, 13, 14, 15, 16, 17, 25, 26, and 27) and ten narrative items (RL-coded 1, 2, 3, 4, 5, 6, 7, 30, 31, and 32). The winter administration (form B) consisted of 11 expository items (RI-coded 4, 5, 10, 17, 18, 19, 20, 21, 22, 23, and 24) and 9 narrative items (RL-coded 1, 2, 12, 13, 14, 15, 26, 27, and 28). Because the number of expository and narrative test items was not matched, a conversion factor of .99 was applied to the expository items and a factor of 1.11 was multiplied to the total score of narrative items to allow for equivalent comparison of scores.

Analysis

Descriptive statistics. Descriptive statistics were computed using IBM SPSS v. 21. The mean and standard deviation of scores were calculated for the entire sample, as well as within subgroups for SES membership (as indicated by FRL status) and by gender. Test items were based on the type of text assessed, whether expository (RI-coded) or narrative (RL-coded), and means and standard deviations were calculated for each based on student SES membership and by gender to provide information on student performance on text type based on subgroup.

Trend analysis. Polynomial contrasts revealing the shape of growth trajectories were evaluated using trend analysis. Polynomial contrasts check for trends between groups. For data measurements with three time points, trends may be linear or quadratic. For linear trends, the dependent variable change is proportionate across categories.

Quadratic trends are characterized by one change in the direction of the line (i.e., one bend in the curve of the line) (Field, 2013). Trend analyses were conducted using IBM SPSS v. 21.

Latent growth curve modeling. Test scores were analyzed using latent growth curve (LGC) modeling. LGC modeling may be viewed as a special case of structural equation modeling (SEM) (Aycock & Li, n.d.; Preacher, Wichman, McCallum, & Briggs, 2008) that permits the examination of change over time as well as the shape of the growth trajectory (i.e., linear, quadratic, cubic). For analyzing data using a structural equation modeling (SEM) approach, it is necessary to have a minimum of three time waves (Byrne, 2010). LGC modeling contains an intercept which is the initial level of the outcome measure. The intercept is similar to the intercept found in general linear models (GLMs) such as ANOVA or regression; however, LGC modeling differs from GLMs in that it allows researchers to examine how well the data fits the model as well as the relationship between intercept and slope of the model. The rate of change in LGC modeling is depicted by the slope of the trajectory.

CHAPTER IV

RESULTS

English I Results

Descriptive statistics and trend analysis. English I overall results for performance on both expository and narrative text items were computed using IBM SPSS v.21 and are shown in Figure 1. The mean and standard deviation of scores were calculated for the entire sample, then within subgroups of SES membership and by gender. Table 3 reports the observed sample means and standard deviations for performance at each time point. Student performance according to text type was examined and descriptive statistics are provided for performance on expository items (Table 4) and narrative items (Table 5). These scores indicate that student performance on expository test items improved with time, while performance on items of narrative text increased initially, but decreased at the final administration. For both expository and narrative text, the standard deviations were wide, thus indicating variability in test scores. Students who were low-SES had lower scores at each test administration than their higher-SES counterparts on both narrative and expository test items. Males scored higher than females on expository items at the fall and spring administrations, but females scored higher at all time points on items involving narrative text. The overall trend of English I comprehension scores on combined expository and narrative test items revealed a significant overall linear trend, $F(1, 2259) = 380.68$, $R^2 = .06$, $p < .01$ as well as a significant quadratic trend, $F(1, 106) = 20.55$, $R^2 = .004$, $p < .01$. Even though each trend was significant, the R^2 value is minimal. With the greater dominance of the linear

trend, findings reveal that the data more closely fit the linear model than the quadratic model, as depicted by the piecewise linear growth.

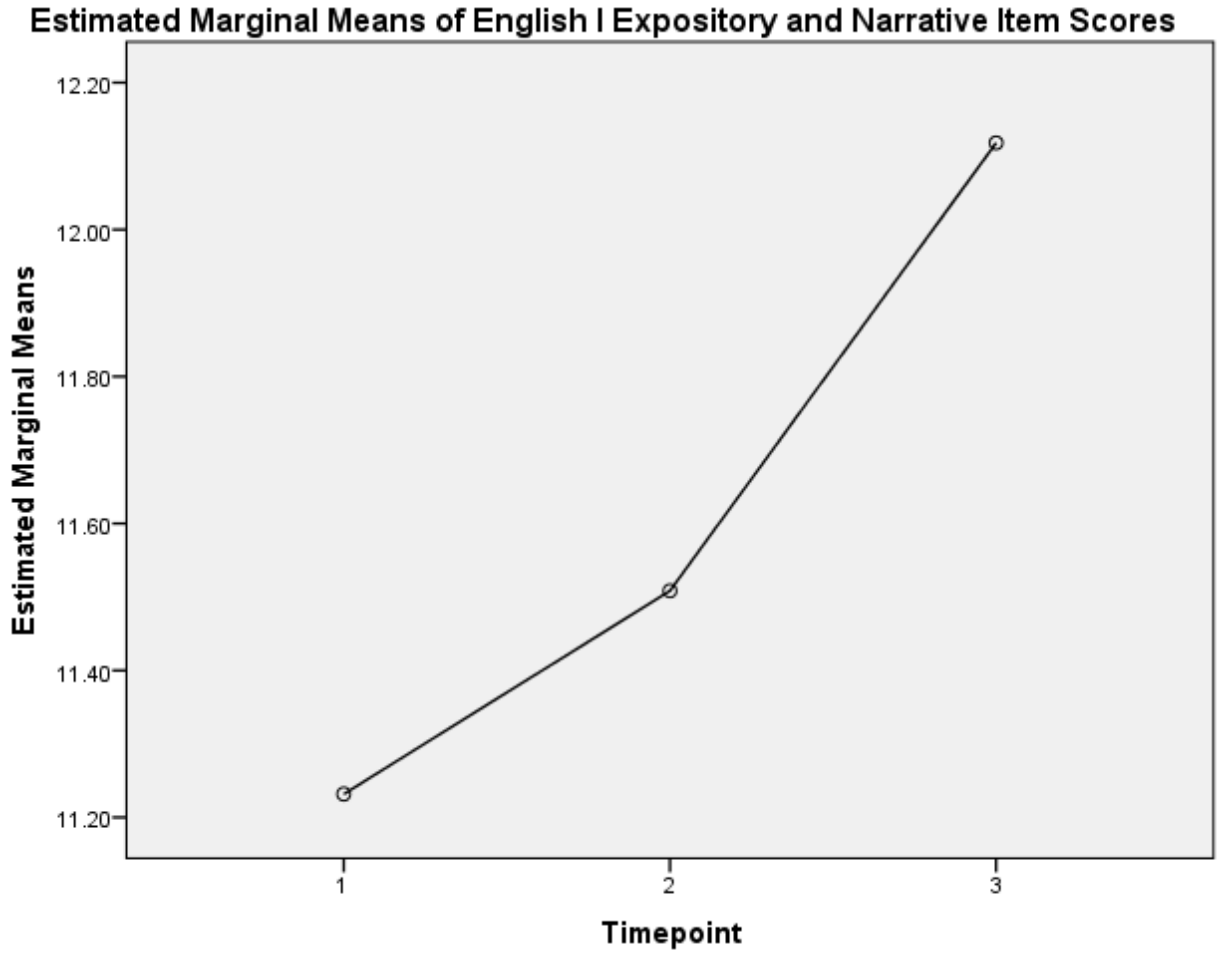

Figure 1. Trend Analysis of English I Mean Combined Scores for Expository and Narrative Test Items.

Table 3

English I Sample Means and Standard Deviations for Performance on Expository and Narrative Text Items

Overall Test Performance	*n* 5752 Full sample *M(SD)*	*n* 3301 FRL *M(SD)*	*n* 2451 no FRL *M(SD)*	*n* 2880 Female *M(SD)*	*n* 2872 Male *M(SD)*
Form A	11.23(4.66)	10.40(4.00)	12.35(4.02)	11.37(4.05)	11.10(4.19)
Form B	11.51(4.65)	10.73(4.04)	12.55(4.17)	11.64(4.17)	11.37(4.21)
Form C	12.12(4.65)	11.28(4.21)	13.25(4.01)	12.13(4.13)	12.12(4.25)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Table 4

English I Sample Means and Standard Deviations for RI-coded Items

Test Performance (Expository)	*n* 5752 Full sample *M(SD)*	*n* 3301 FRL *M(SD)*	*n* 2451 no FRL *M(SD)*	*n* 2880 Female *M(SD)*	*n* = 2872 Male *M(SD)*
Form A	5.06(2.54)	4.58(2.43)	5.71(2.55)	5.02(2.49)	5.10(2.60)
Form B	5.28(2.48)	4.87(2.38)	5.83(2.49)	5.32(2.45)	5.24(2.51)
Form C	6.12(2.32)	5.67(2.31)	6.73(2.19)	6.06(2.25)	6.19(2.39)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Table 5

English I Sample Means and Standard Deviations for RL-coded Items

Test Performance (Narrative)	*n* = 5752 Full sample *M(SD)*	*n* = 3301 FRL *M(SD)*	*n* = 2451 no FRL *M(SD)*	*n* = 2880 Female *M(SD)*	*n* = 2872 Male *M(SD)*
Form A	6.17(2.12)	5.83(2.15)	6.64(1.99)	6.35(2.08)	6.00(2.14)
Form B	6.23(2.17)	5.86(2.14)	6.73(2.12)	6.32(2.17)	6.14(2.17)
Form C	6.00(2.33)	5.61(2.34)	6.52(2.21)	6.07(2.29)	5.92(2.36)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Latent growth curve analysis. Latent growth curve analysis allows researchers to investigate change in individual growth over time (of three or more time points) when the outcome variables are of a continuous scale. The time between observations may be spaced either evenly or unevenly, but the actual number of observations and the lag between those observations must be the same for all examinees (Byrne, 2010; Byrne & Crombie, 2003; Willett & Sayer, 1994). The AMOS module in IBM SPSS v. 21 provides schematic representations of models in addition to outputs. To aid in interpreting the schematic of a model, a brief explanation is necessary: latent (or unobserved) factors are represented by ovals (or circles); observed variables are represented by squares (or rectangles); single-headed arrows depict the impact of one variable on another; and double-headed arrows are used to represent the covariance or correlation between pairs of variables (Byrne, 2010; Byrne & Crombie, 2003; Preacher, 2010).

All models presented in this study include circles with the letter "E" inside (i.e., E1, E2, E3) which represent random measurement error. When an impact variable is added (in this study, when SES is added to the model), circles with the letter "D" also appear which represent disturbance (residual) terms for latent traits. These residuals (D1 and D2) represent individual differences in intercept and growth trajectories (Byrne & Crombie, 2003). The value of the initial point of observation (in this study, student scores at the fall test administration) is deemed the intercept. In LGC, the intercept is similar to those found in general linear models (GLMs) such as ANOVA or regression; however, LGC modeling differs from GLMs in that it allows researchers to examine how well the data fit the model as well as the relationship between intercept and slope of the model. The rate of change in LGC modeling is depicted by the slope of the trajectory. For examination of change over three time points, the slope may be linear (indicating constant growth) or quadratic, depicting change in the direction of the curve between the first and the third test administration. The latent growth curve represents an average of individual student growth curves of all examinees. For English I, basic models were constructed for performance on expository and narrative items. Then a second model (the SES model) was constructed for expository and narrative items, providing information on how student socioeconomic status as indicated by free or reduced-price lunch membership impacted student performance. As such, figures depict schematics for a basic model for English I expository, a model that adds the impact of SES for English I expository, a basic model for English I narrative, and a model that adds SES for English I narrative.

Model fit. A major goal in latent growth curve modeling is the determination of how well the model represents the data, a process that is established a priori (Byrne, 2010; Preacher, 2010). Indices of model fit for this study included those of absolute, relative, and non-centrality. It is important to note that AMOS provides standardized coefficients even for fixed parameters. Fit indices are provided for default, saturated, and independence models. The default model indices provide information on the hypothesized model as well as the saturated and independence models. The saturated model suggests a model in which the number of data points equals the number of estimated parameters while the independence model is one that represents a zero value for correlations among all variables (Byrne, 2010). Byrne (2010) describes these as analogous to points along a continuum, with the saturated model found at one end of a continuum and the independence model located at the other extreme. The hypothesized model (the default model) resides somewhere in between these two extremes. The χ^2 test is used as an index for absolute fit; however, due to the sensitivity to sample size presented by χ^2 as a fit index, it is not used as frequently by researchers. The χ^2/df ratio is an early example of one of the fit statistics created to address the problems of the χ^2 sample size sensitivity (Wheaton, Muthén, Alwin, & Summers, 1977). Relative fit, which compares the model's χ^2 to one from the baseline (or null) model, is indicated by the normed fit index (NFI). Although the NFI has demonstrated the tendency to underestimate fit when sample sizes are small (Byrne, 2010), the sufficient sample size in the present study ensured accurate estimation produced by this index. A revised form of the NFI is the comparative fit index (CFI) proposed by Bentler (1990). Ranges for both the NFI and the CFI are between zero and 1.00 with acceptable limits of > .90; however,

a revised cut criterion of .95 has been proposed by Hu and Bentler (1999). The Tucker-Lewis index (TLI) another incremental fit index, also provides values that range from zero to 1.00 with values near .95 providing indication of good fit (Byrne, 2010; Hu & Bentler, 1999). Non-centrality fit indices are revealed through confidence intervals (CI) and root mean square error of approximation (RMSEA). The RMSEA is sensitive to model complexity (as indicated by the number of parameters in the model) and cut ranges for this fit index are subjective, but acceptable fit includes values as high as .08 with values lower than .05 representing good fit (Browne & Cudeck, 1993; Byrne, 2010). Table 6 summarizes acceptable fit ranges for the indices used in the present study.

Table 6

Ranges for Model Fit

Fit Index	Possible Range	Poor Fit	Acceptable Fit	Good Fit
χ^2/df	$0 - \infty$	> 3.00	2.00 – 3.00	< 2.00
CFI	0 – 1.00	< .90	.90 - .94	>.95
NFI	0 – 1.00	< .90	.90 - .94	>.95
TLI	0 – 1.00	< .90	.90 - .94	>.95
RMSEA	0 – 1.00	>.08	< .08	< .05

Note. χ^2/df = Chi-square/degrees of freedom ratio; CFI = Comparative Fit Index; NFI = Normed Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Squared Error of Approximation

English I expository models. Figure 2 displays the schematic of the basic model for English I expository text and examines how the variables influenced student scores at the three measurement time points. Slope and intercept are latent, exogenous traits in the schematic with rectangles labeled Exp1A, Exp1B, and Exp1C representing observed, endogenous variables of the mean scores on test items gauging comprehension of expository text at the three test waves. Latent variables E1, E2, and E3 represent

unobserved, exogenous traits that depict individual variance in the observed fall, winter, and spring test scores.

The basic model for expository item performance indicated good fit for three of the five major goodness-of-fit indices, suggesting that implications of theory may translate to practical use in real-world educational settings (Pickens, 2016). As stated previously, χ^2 values are sensitive to sample size and RMSEA is sensitive to complexity as indicated by the number of parameters in the model. Values for NFI, TLI, and CFI all indicated good fit ($\chi^2 = 146.10$; $df = 1$; $p < .001$; NFI = .97; TLI = .92; CFI = .97; RMSEA = .16; RMSEA 90% CI = .14 - .18).

Following model creation and verification of fit, LGC analyses were conducted for the English I expository basic model. Parameters for mean and variance were allowed to freely vary for intercept and slope. Mean (4.97) and variance (.03) of the intercept as well as the mean (1.07) and variance (.03) of the slope were both significant at the $p < .001$ level. In addition to the trend analysis which confirmed trajectory, factor loadings provided information regarding growth and change across test waves. Standardized intercept loadings to each manifest variable ranged .80 - .88 thus indicating a significant impact on each observed variable of reading comprehension performance of expository text items for the basic model. The correlation between the intercept and slope was negative (-.52). This negative estimate was statistically significant ($p < .05$) and suggests that students who had a higher initial score ("high starters") demonstrated a lower rate of increase in reading performance on expository items over the school year than those who had a lower initial score ("low starters"). As such, low starters achieved greater gains over the school year than did high starters. The negative correlation

between initial status and change of rate is indicative of the phenomenon of the law of initial values (Byrne, 2010).

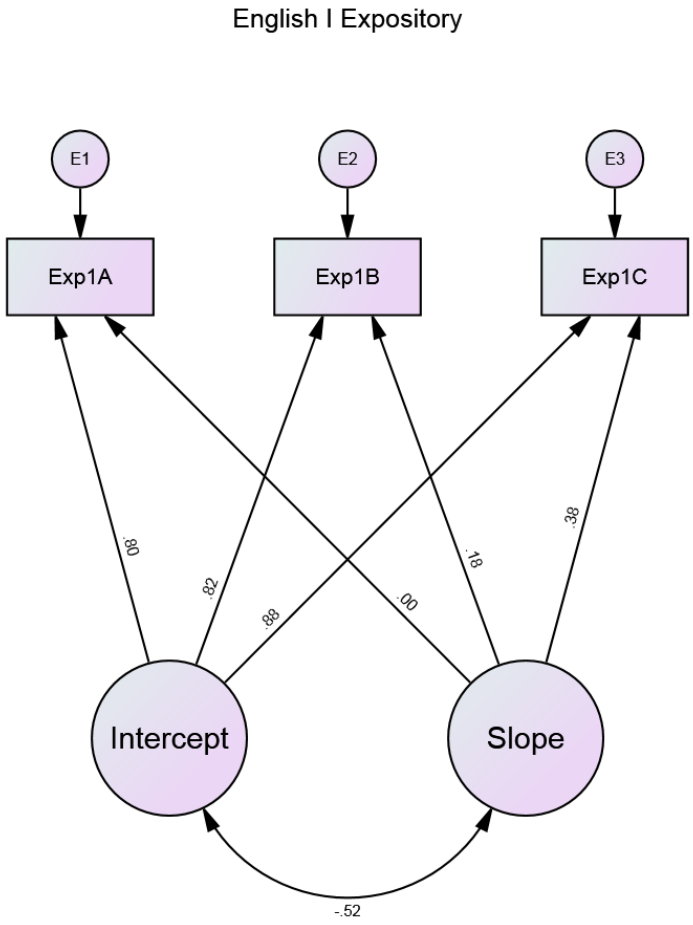

Figure 2. English I Basic Model for Expository Item Performance

Figure 3 shows the schematic of English I expository text with student socioeconomic status (SES) as an impact factor. When the impact of SES was added in to the model, two of the five indices produced good fit, one indicated acceptable fit, and

two suggested poor fit (χ^2 = 152.79; *df* = 2; *p* < .001; NFI = .97; TLI = .92; CFI = .97; RMSEA = .11; RMSEA 90% CI = .10 - .13). Following model creation and verification of fit, LGC analysis was conducted for the English I expository SES model. Adding the SES variable to the model allowed for examination of the impact of SES on the intercept and slope of the model. The impact of SES on the intercept was significant (β = .263, *p* < .001), indicating that SES membership considerably influenced the starting point of expository reading scores. The impact of SES on the slope was not significant (β = -.039, *p* < .221), indicating that SES membership did not play a substantial role in growth in scores over subsequent test waves.

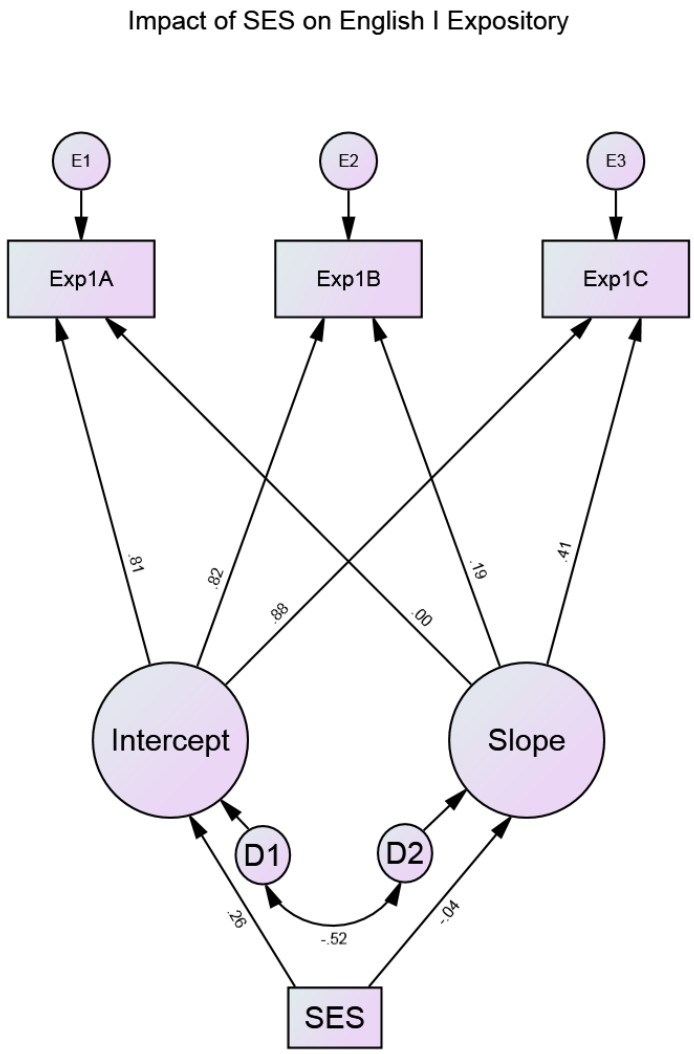

Figure 3. English I Expository Performance with the Impact of SES Added to Model

English I narrative models. The basic model reflecting performance for English I narrative text is depicted in Figure 4 with the impact of SES on narrative item performance shown in Figure 5. For both the basic and SES narrative models, the rectangles labeled E1NARA, E1NARB, and E1NARC represent the mean scores of students on narrative test items at the fall, winter, and spring testing intervals. When

testing for performance on items involving narrative text, indices for χ^2/df were 41.01. The comparative fit index (CFI) and the normed fit index (NFI), both indices of relative fit, produced close fit, with indices of .99 for both. The basic narrative model RMSEA produced a value of .08 while the narrative model with SES added produced .06, indicating acceptable fit.

Parameters were allowed to freely vary for intercept and slope for the basic model, and intercept loadings adjusted to a range of .68 - .74. Mean (4.98) and variance (.03) of the intercept as well as the mean (.14) and variance (.02) of the slope were both significant at the $p < .001$ level. The correlation between the intercept and slope was positive (.15). This positive correlation indicates that high starters' scores increased following the initial administration and that low starters increase, yet not at the same magnitude as the high starters. This estimate was not statistically significant ($p = .184$); therefore, there was no detection of the Matthew Effect or regression toward the mean for performance on these items.

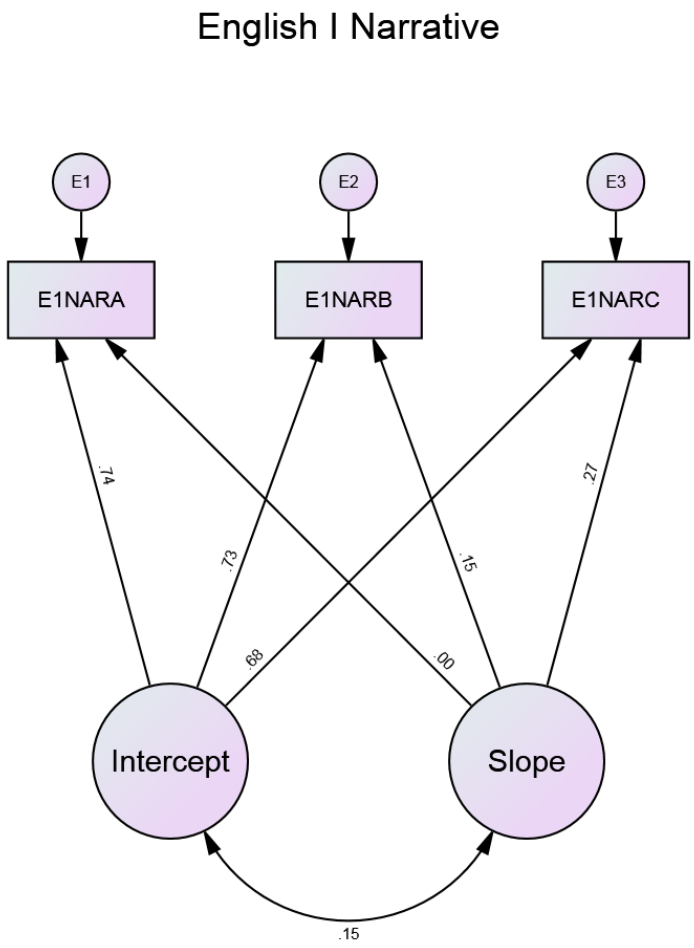

Figure 4. English I Basic Model for Narrative Item Performance

The impact of SES on the intercept was statistically significant (β = .255, p < .001) which indicated that SES membership had an impact on the initial test score as shown in Figure 5. Hence, students who were low-SES started lower than students who were higher SES. However, SES had a non-significant impact on the slope (β = .081, p =

.067), meaning that a student's rate of changes in subsequent performance on winter and spring assessments was not impacted by SES.

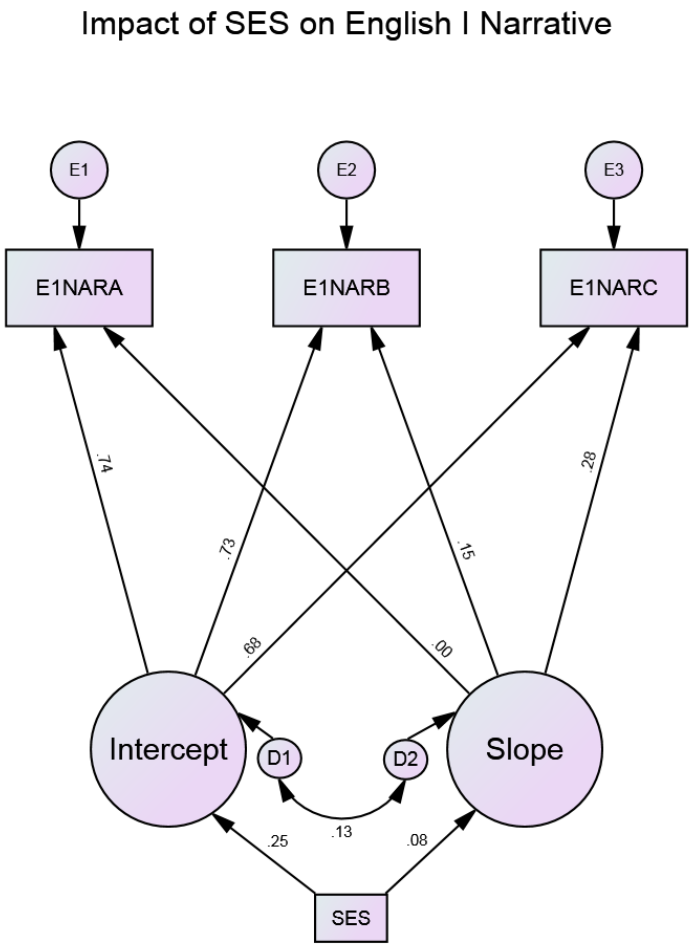

Figure 5. English I Narrative Performance with the Impact of SES Added to Model

English III Results

Descriptive statistics and trend analysis. Table 7 shows the sample means and standard deviations for combined narrative and expository item performance among students taking the English III fall, winter, and spring assessments. A visual inspection reveals a decline in performance between the initial assessment and the winter testing, followed by a sharp increase in performance at the spring test administration. Students receiving free or reduced lunch (determined as low-SES) scored lower than their counterparts at each test wave. The widest gap between combined expository and narrative scores for low- and medium/hi-SES was at the time of the spring test administration; however, the large standard deviations for both membership categories suggest wide dispersion among scores. On combined items, females outscored males at each time wave. This trend is also evident when examining performance according to item type (Tables 8 and 9); however, performance among males and females on expository items at the winter administration reveal similar means and roughly similar standard deviations. The overall trend of English III comprehension scores on combined expository and narrative test items, shown in Figure 6, revealed significant overall linear and quadratic trends, with a more the quadratic trend being more dominant. Results for the linear trend were $F(1, 1942) = 301.40$, $R^2 = .07$, $p < .01$, and quadratic trend results were $F(1, 3967) = 680.74$, $R^2 = .15$, $p < .01$.

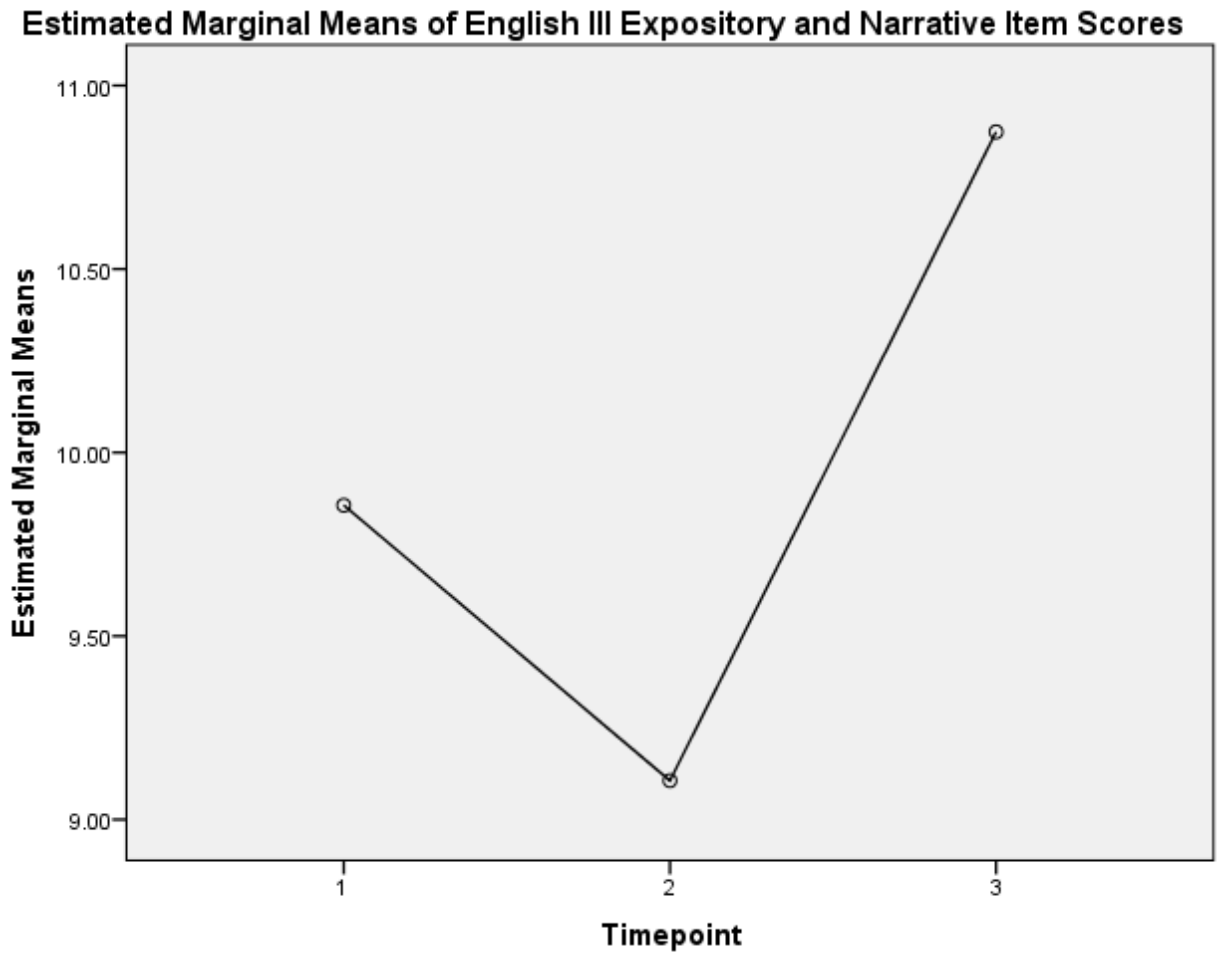

Figure 6. Trend Analysis of English III Mean Combined Scores for Expository and Narrative Test Items.

Table 7

English III Sample Means and Standard Deviations for Expository and Narrative Text Items

Overall Test Performance	*n* = 3754 Full sample *M(SD)*	*n* = 2062 FRL *M(SD)*	*n* = 1692 no FRL *M(SD)*	*n* = 1831 Female *M(SD)*	*n* = 1923 Male *M(SD)*
Form A	9.86(3.84)	8.97(3.60)	10.94(3.84)	10.04(3.88)	9.68(3.79)
Form B	9.10(3.95)	8.32(3.67)	10.06(4.07)	9.27(3.98)	8.95(3.92)
Form C	10.88(4.55)	9.89(4.29)	12.08(4.56)	11.02(4.52)	10.73(4.56)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Table 8

English III Sample Means and Standard Deviations for RI-coded Items

Test Performance (Expository)	*n* = 3754 Full sample *M(SD)*	*n* = 2062 FRL *M(SD)*	*n* = 1692 no FRL *M(SD)*	*n* = 1831 Female *M(SD)*	*n* = 1923 Male *M(SD)*
Form A	4.47(2.21)	4.03(2.09)	5.01(2.23)	4.56(2.19)	4.38(2.22)
Form B	4.69(2.30)	4.24(2.16)	5.24(2.35)	4.69(2.27)	4.69(2.32)
Form C	5.07(2.59)	4.56(2.47)	5.69(2.60)	5.14(2.58)	5.00(2.61)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Table 9

English III Sample Means and Standard Deviations for RL-coded Items

Test Performance (Narrative)	*n* = 3754 Full sample *M(SD)*	*n* = 2062 FRL *M(SD)*	*n* = 1692 no FRL *M(SD)*	*n* = 1831 Female *M(SD)*	*n* = 1923 Male *M(SD)*
Form A	5.39(2.14)	4.94(2.06)	5.93(2.10)	5.48(2.18)	5.30(2.09)
Form B	4.42(2.17)	4.08(2.05)	4.83(2.25)	4.58(2.20)	4.25(2.13)
Form C	5.81(2.41)	5.33(2.33)	6.39(2.38)	5.88(2.40)	5.74(2.42)

Note. Test Form A was administered in fall, 2014; Form B: winter, 2015; Form C: spring, 2015. *M* = mean; SD = standard deviation. FRL = Free/reduced lunch recipient (indicating low-SES)

Model fit. For the English III data, the same criteria was followed as described for English I. Models were created to include a basic model depicting performance on expository items, a model depicting performance on expository items including the impact of student SES on scores, a basic model showing performance on items assessing performance on narrative text, and a narrative model with student SES added to examine the impact of student performance according to SES membership dichotomized based on whether the student was a recipient of free/reduced lunch (FRL) or whether the student was ineligible for FRL.

English III expository models. Figure 7 displays the schematic of the basic model for English III expository text and examines how the variables influenced student scores at the three measurement time points. Slope and intercept are latent, exogenous variables in the schematic with rectangles labeled E3AEXP, E3BEXP, and E3CEXP representing observed, endogenous variables of the mean scores on test items gauging

comprehension of expository text at the three test waves. Latent variables E1, E2, and E3 represent unobserved, exogenous variables that depict individual variance in the observed fall, winter, and spring test scores. The basic model for expository item performance indicated good fit for all indices excluding χ^2 which tend to underestimate fit because of sensitivity to sample size ($\chi^2 = 6.53$; $df = 1$; $p = .011$; $\chi^2/df = 6.53$; NFI = 1.00; TLI = 1.00; CFI = 1.00; RMSEA = .04; RMSEA 90% CI = .02 - .07). The next model examined the intercept, slope, and impact of student SES (Figure 8). When the impact of SES was added, model fit was good for all indices except χ^2 which is sensitive to sample size ($\chi^2 = 7.62$; $df = 2$; $p = .022$; $\chi^2/df = 3.81$; NFI = 1.00; TLI = 1.00; CFI = 1.00; RMSEA = .03; RMSEA 90% CI = .01 - .05).

Following model creation and verification of fit, LGC analyses were conducted for the English III expository basic and SES models. Parameters were allowed to freely vary for intercept and slope. Mean (4.44) and variance (.04) of the intercept as well as the mean (.30) and variance (.02) of the slope were both significant at the $p < .001$ level. In addition to the trend analysis which confirmed trajectory, factor loadings provided information regarding growth and change across test waves. Intercept loadings adjusted to a range of .62 - .72. The correlation between the intercept and slope for the basic model was significant ($r = .356$, $p = .014$). This suggests that students' initial starting score correlated to growth on subsequent test administrations.

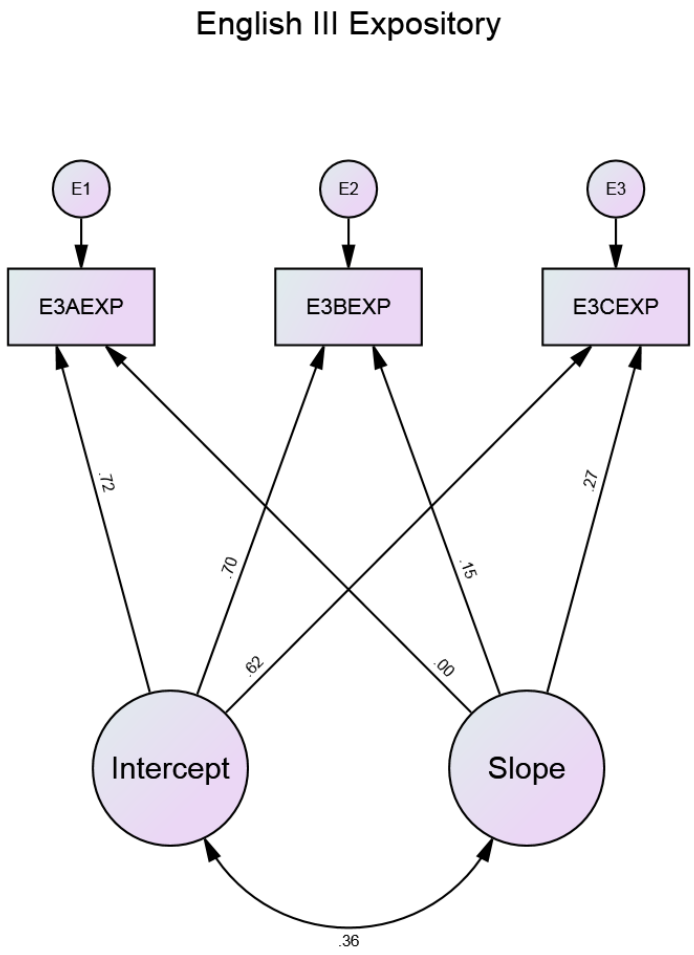

Figure 7. English III Basic Model for Expository Item Performance

Following model creation and verification of fit, LGC analyses were conducted for the English III expository SES models. Parameters were allowed to freely vary for intercept and slope. Mean (4.98) and variance (.03) of the intercept as well as the mean (.14) and variance (.02) of the slope were both significant at the $p < .001$ level. In addition to the trend analysis which confirmed trajectory, factor loadings provided

information regarding growth and change across test waves. Intercept loadings adjusted, similar to the adjustment seen in the basic model, to a range of 62 - .72. Adding the SES variable to the model allowed for examination of the impact of SES on both intercept and slope on reading scores. The impact of SES on intercept was statistically significant (β = .297, $p < .001$) indicating meaningful influence of SES membership for intitial assessment expository reading scores. The impact of SES on slope was not significant (β = .095, $p = .065$) which revealed no substantial impact to the change rate of reading scores over time.

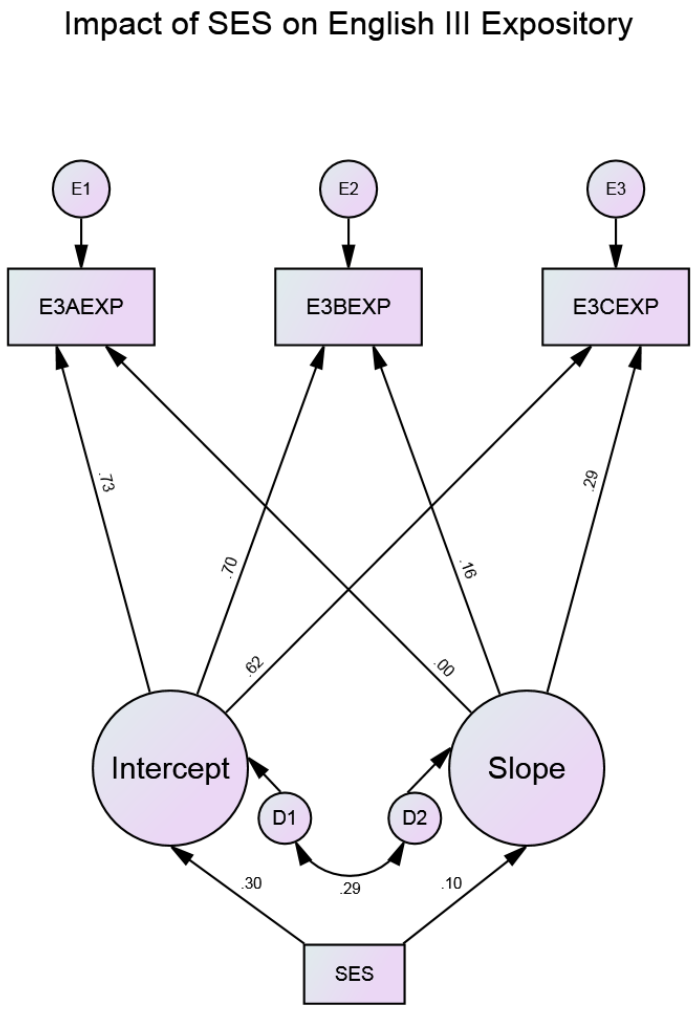

Figure 8. English III Expository Performance with the Impact of SES Added to Model

English III narrative models. The basic model reflecting performance for English III narrative text is depicted in Figure 9 with the impact of SES on narrative item performance shown in Figure 10. For both the basic and SES narrative models, the rectangles labeled E1NARA, E1NARB, and E1NARC represent the mean scores of students on narrative test items at the fall, winter, and spring testing intervals. Latent

variables E1, E2, and E3 represent unobserved, exogenous variables that depict individual variance in the observed fall, winter, and spring test scores. Slope values for the months between tests (.00, 1, and 2) included the quadratic term within both the basic and SES models. The LGC analysis was conducted for English III narrative reading scores with the identical medel specifications to the previous analyses. The model did not converge and all model fit indices were unacceptably low, (χ^2 = 1120.98; *df* = 1; *p* < .001; χ^2/df = 1120.98; NFI = .56; TLI = -.34; CFI = .55; RMSEA = .55; RMSEA 90% CI = .52 - .57). The causes for this non-fitting solution were unknown, and warranted further research. Due to non-convergence of the basic model, no further analysis for the SES model was conducted.

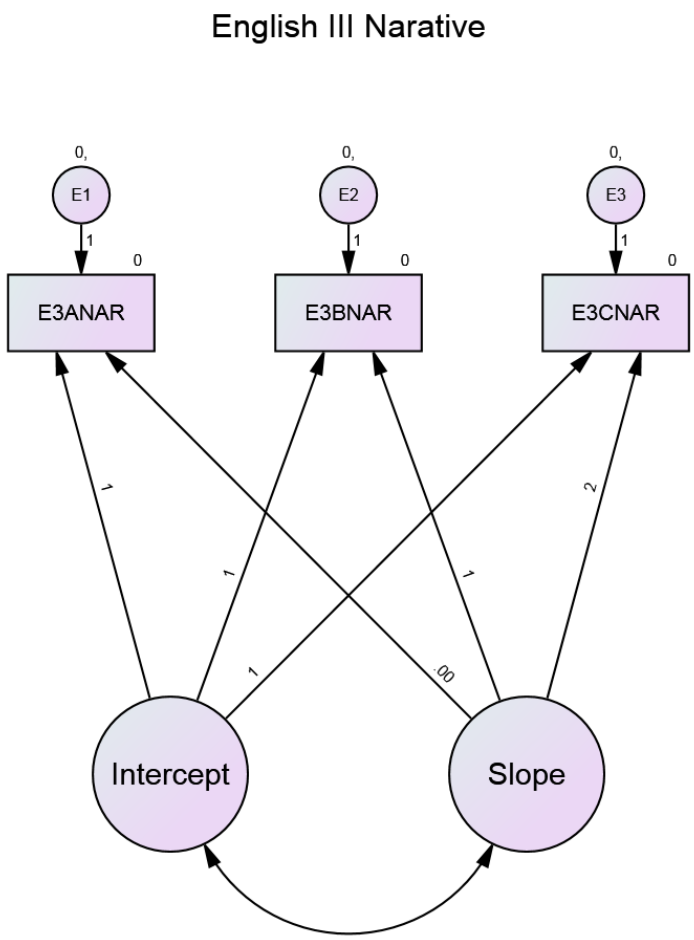

Figure 9. English III Basic Model for Narrative Item Performance

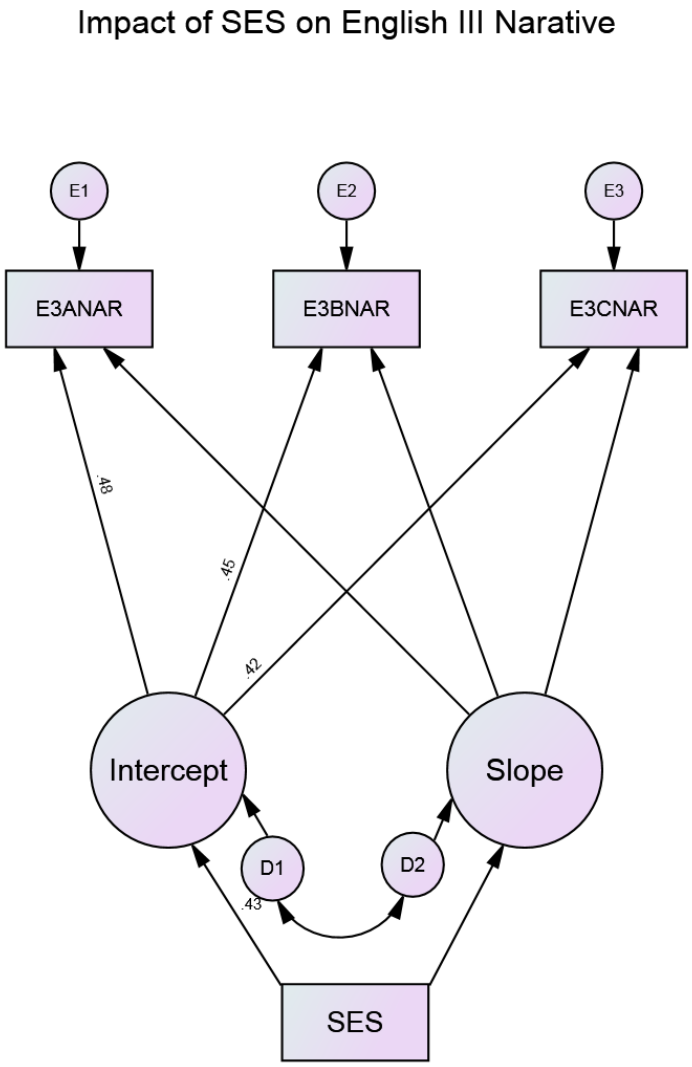

Figure 10. English III Narrative Performance with the Impact of SES Added to Model

CHAPTER V

DISCUSSION

Knowing how high school students perform in reading is crucial. Although research involving students at younger grades is abundant, we particularly need to identify what elements pose difficulty for our students as they prepare to transition from school to work or to higher education. Likewise, we need to know how factors such as poverty impact student reading comprehension and to examine whether text type presents particular issues for students. The present study sought to examine how students taking English tripartite exams fared overall on narrative and expository text items; it further considered the relationship between a student's starting test score and his or her growth on subsequent assessments throughout the year, as well as how student SES affects performance and impacts student scores. Despite more than 45 years of data revealing national educational progress (NAEP), we still fail to see substantial, significant positive changes in reading scores among our 17 year-olds (National Center for Educational Statistics, 2013). The widening gap among students in poverty (NAEP, 2015) tells us that we are missing something here, too – there is some critical piece that we are missing to the detriment of this subgroup of students. The use of latent growth curve modeling to examine change in literacy research to look at change and impact, particularly as it applies to students in low-SES environments, presents a way in which the present study offers a contribution to the body of knowledge around a highly important educational topic – poverty and difficulty in reading comprehension among students poised to begin their adult lives.

Overall Trajectories

The first aim of the study was to examine trajectories of high school students' performance on test items of reading comprehension. It was hypothesized that the overall trend would be nonlinear on both English I and English III exams based on a review of existing literature revealing curvilinear trends (Beecher, 2011; Petcher et al., 2014; Rescorla & Rosenthal, 2004). This hypothesis was partially confirmed: overall results for English I expository and narrative combined results revealed both a significant linear trend and a significant quadratic trend; however, the greater dominance of the linear trend suggests that the data more closely fit the linear model. The full sample had a mean initial fall score of 11.20 and a final Spring mean score of 12.12. The winter test administration mean score of 11.51, upon visual inspection (found in Figure 1 of the Results section), depicts a slight deviation from a true linear trend between initial and final score. The Winter test administration falls around the time of the end of the fall and beginning of the spring semester in U. S. schools. It is common for many schools to be out of session between late December and early January, so the non-instructional break may be indicative of a winter effect that is similar to effects seen from the prolonged non-instructional period frequently referred to as the *summer slump* (Alexander et al., 2007; Cooper et al., 1996; Entwistle & Alexander, 1992, 1994; Heyns, 1978; Jamar, 1994). It is also possible that the unsteady linear growth is attributable to residual effects from the transition to high school in general: although some English I students were re-taking assessments at a later point in their high school careers, the majority of students taking these assessments were "true" freshmen. Studies have suggested that this transitional period and adjustment to high school presents a difficulty for many (Alspaugh, 1998;

Barber & Olsen, 2004; Benner, 2011; Crosnoe & Huston, 2007; Neild, 2009). It is possible that after an initial "honeymoon period" of adjusting to life as a secondary student, the novelty wears off, thus revealing difficulties related to acclimatization exhibited in the results seen here.

Among English III test participants, estimated marginal means on combined narrative and expository item performance revealed a pronounced quadratic trend (as shown in the Results section in Figure 2). The trend analysis found that both linear and quadratic trends were significant but that the quadratic trend was more dominant. With the dominance of the curvilinear trend, this hypothesis was confirmed. When examining sample means and standard deviations of performance on expository items separate from narrative items, an interesting pattern emerged: when looking solely at scores on expository items, student performance increased incrementally at each test wave among both subgroups (gender and SES); however, when examining narrative items, sample means and standard deviations dropped at the winter test wave for both subgroups. Furthermore, standard deviations did not reveal any pronounced increase suggestive of outliers in the data. Because of the sharp dip in performance at the point of the second test, a visual examination of data for individual narrative item performance was conducted to look for the possibility of extreme irregularity in the test questions. The cause of this anomaly of poor model fit is unknown and must be left for further investigation. Analysis using item response theory (IRT) might be useful to determine item difficulty and discrimination characteristics for the winter test form's narrative items.

Relationship Between Intercept and Slope

The next two hypotheses were related: the study sought to examine how a student's initial score at the time of the fall administration (the intercept) was related to how performance changed over the following administrations (intercept) on test items according to type (i.e., expository or narrative items). It was hypothesized that there would be a significant relationship between the slope and the intercept on both narrative and expository test items. To elucidate, the hypotheses predicted that a student's initial score (i.e., whether he or she had a high initial score or a low initial score) would be significantly related to the growth on performance of subsequent tests for both item types. A high initial score would beget higher growth; a low initial score would beget less growth. The basis for this prediction was based on the phenomenon of accumulated advantage frequently referred to as the Matthew Effects (Stanovich, 1986, Walberg & Tsai 1983). Performance on English I expository test items revealed the opposite: students with a lower initial starting score experienced more gain than high starters as indicated by the negative intercept-slope correlation. A possible reason for this includes the phenomenon of the law of initial values (Wilder, 1962) which suggests that the direction of response (i.e., growth on reading comprehension assessment) depends highly on the initial level of the function (i.e., starting score on the Form A assessment). When investigating the relationship between initial performance on explository reading comprehension scores for English III test-takers, it was determined that students who scored high on the initial test performed higher over the subsequent test waves. Because of the poor model fit for the narrative basic and SES models, it was determined that the

relationship between slope and intercept could not be adequately explained by the models.

Impact of SES on the Initial Starting Score (intercept) and Change (slope) on Subsequent Assessments

The impact of SES on the intercept of items assessing expository text was positive for both English I and English III tests; however, the English III narrative model did not produce good fit, so it is not possible to interpret its true impact. The impact was significant for the English I narrative and English III expository items. The impact of SES on the slope for English I test items was mixed; SES accounted for a negative impact on the slope for English I expository items, but it produced a positive impact on narrative items. English III SES had no significant impact on the rate of change for expository items, and the poor fit for the English III narrative resulted in the inability to accurately detect SES impact for items of narrative text.

It was hypothesized that SES would have no significant impact on the intercept and slope for items involving narrative text, but that there would be a significant impact of SES on expository text items. As such, this hypothesis was partially confirmed by the significant relationship between SES and intercept on English I expository and narrative and English III expository item performance, but the non-significant relationship between SES and slope on English I expository and narrative performance and on English III expository performance.

Limitations of the Study, Implications for Practice, and Suggestions for Further Research

This study was limited in that examination was conducted using a single data set. Because the present study involved archival data, there was thus limited access to item information. Furthermore, as mentioned previously, there was poor model fit for the English III narrative models. Because poor fit was indicated in the construction of the basic model, no further analysis on the impact of SES was possible. An additional limitation is in the classification of SES status itself: as is common in practice, SES was determined through student participation in free/reduced price lunch (FRL). Because of the ease of obtaining FRL status, the FRL category is frequently used as a proxy for SES although FRL may not correctly represent poverty status as it is possible for some schools (or some school districts) to qualify as a whole as a high poverty school or district without accurately representing the income status of its individual students. This inclusion of students in to FRL membership could affect the values indicated in the data analysis.

The findings of this study have implications for real-world practice. The impact of SES on the student starting point at the time of the fall test wave indicated that students in poverty scored lower than those who were not low SES. This fall test administration was the initial assessment following the summer non-instructional period and suggests summer effects as described in research studies including Cooper et al. (1996) and Entwistle et al. (2007). These findings strengthen the argument for providing summer reading programs for low-SES students.

The research findings suggest need for further studies specifically addressing high school students' performance on measures of reading comprehension. With the increasing emphasis of learning through expository text, it is imperative to examine whether these items pose greater difficulty than items assessing comprehension of narrative text. While there is considerable debate as to whether expository text poses greater difficulty for students than narrative text, it is unclear if text type presents an additional level of difficulty for students (N. K. Duke, personal communication, October 29, 2016). It is recommended that further research involving analysis using IRT be conducted to determine whether item difficulty or discrimination varies according to text type.

REFERENCES

ACT. (2005). *Crisis at the core: Preparing all students for college and work.* Iowa City, Iowa: ACT.

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235–251. doi:10.1037/0022-0663.100.2.235

Alexander, K., Entwisle, D., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(1), 167–180.

Allinder, R. M., Fuchs, L. S., & Fuchs, D. (1998). Curriculum-based measurement. In H. B. Vance (Ed.), *Psychological assessment of children: Best practices for school and clinical settings* (2nd ed., pp. 106–129). New York: Wiley.

Alspaugh, J. W. (1998). Achievement loss associated with the transition to middle school and high school. *Journal of Educational Research. 92*, 20–25. Retrieved from http://www.jstor.org/stable/27542186

Applebee, A. N. (1984). Writing and reasoning. *Review of Educational Research, 54*(4), 577–596.

Applebee, A. N. (1993). *Literature in the secondary school: Studies of curriculum and instruction in the United States*. Urbana, IL: National Council of Teachers of English.

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from triannual universal screenings. *School Psychology Review, 37,* 109−125.

Aycock, A. C., & Li, F. (n.d.). *Latent growth curve analysis: A gentle introduction*. Unpublished manuscript, Department of Human Development and Family Sciences, Oregon State University, Corvallis, Oregon. Retrieved from http://oregonstate.edu/dept/hdfs/papers/lgcgeneral.pdf

Barber B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research, 19*, 3–30.

Beecher, C. C. (2011). *A latent growth curve analysis of reading achievement for an at-risk population* (Doctoral dissertation). Retrieved from https://research.libraries.wsu.edu/xmlui/handle/2376/3004

Bell, S. M., & McCallum, R. S. (2008). *Handbook of reading assessment*. Boston: Pearson.

Bell, Y. R., & Clark, T. R. (1998). Culturally relevant reading material as related to comprehension and recall in African American children. *Journal of Black Psychology, 24,* 455–475.

Benner, A. D. (2011). Latino adolescents' loneliness, academic performance, and the buffering nature of friendships. *Journal of Youth and Adolescence, 40,* 556–567. doi:10.1007/s10964-010-9561-2

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Betts, J., Reschley, A., Pickart, M., Heistad, D., Sheran, C., & Marston, D. (2008). An examination of the predictive bias for second grade reading outcomes from measures of early literacy skills in kindergarten with respect to English language learners and ethnic subgroups. *School Psychology Quarterly*, *23*, 553–570. doi:10.1037/1045-3830.23.4.553

Beyer, J. (2007). Strategies for helping struggling readers comprehend expository text. *Strategies for Comprehending Expository Texts, p. 1-18.* Unpublished manuscript.

Bowers, L. D. (2012). *Improving literacy through instruction and community experiences: The effect of a summer program on literacy outcomes of students from low SES homes* (Doctoral dissertation). Retrieved from http://dc.uthsc.edu/dissertations/29/

Brame, C. (2013). Writing good multiple choice test questions. Retrieved from https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

Brody, P., DeMilo, C., & Purves, A. C. (1989). *The current state of assessment in literature* (Report Series No. 3.1). Albany, NY: Center for the Learning and Teaching of Literature.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Browne, M. W., MacCallum, R. C., Kim, C.-T., Andersen, B. L., & Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods*, *7*(4), 403–421. doi:10.1037//1082-989X.7.4.403

Bush, G. H. W. (1991). *America 2000: An educational strategy.* Washington, DC: U.S. Department of Education.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* New York: Routledge.

Byrne, B. M., & Crombie, G. (2003). Modeling and testing change over time: An introduction to the latent growth curve model. *Understanding Statistics:Statistical Issues in Psychology, Education, and the Social Sciences, 2*, 177–203.

Chall, J. S., & Jacobs, V. A. (2003). The classic study on poor children's fourth-grade slump. *American Educator, 27*(1), 14–15, 44. Retrieved from http://www.aft.org/newspubs/periodicals/ae/spring2003/hirschsbclassic.cfm

Chall, J. S., Jacobs, V. A., & Baldwin, L. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

Christ, T. J., & Silberglitt, B. (2007). Curriculum-based measurement of oral reading fluency: The standard error of measurement. *School Psychology Review, 36*(1), 130−146.

Close, E. (1992). Literature discussion: A classroom environment for thinking and sharing. *English Journal, 81*(5), 65–71.

Coleman, J. S., Campbell, E. Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education. Retrieved from http://files.eric.ed.gov/fulltext/ED012275.pdf

Conley, D. T. (2014). *The Common Core State Standards: Insight into their development and purpose.* Washington, DC: Council of Chief State School Officers.

Cooper, H., Charlton, K., Valentine, J., & Muhlenbruck, B. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monograph Series for the Society for Research in Child Development, 65*(1).

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227–268.

Cowan, C., Hauser, R., Kominski, R., Levin, H., Lucas, S., Morgan, S., Spencer, M., & Chapman, C. (2012). Improving the measurement of socioeconomic status for the national assessment of educational progress: A theoretical foundation. National Center for Education Statistics. Retrieved from http://files.eric.ed.gov/fulltext/ED542101.pdf

Crosnoe R., & Huston, A. C. (2007). Socioeconomic status, schooling, and the developmental trajectories of adolescents. *Developmental Psychology, 43,* 1097–1110.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurements to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507−525.

Duke, N. K. (2000a). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35,* 202–224.

Duke, N. K. (2000b). For the rich it's richer: Print experiences and environments offered to children in very low- and very high- socioeconomic status first grade classrooms. *American Educational Research Journal, 37,* 441–478.

Duncan, T. E., & Duncan, S. C. (2009). The ABC's of LGM: An introductory guide to latent variable growth curve modeling. *Social and Personality Psychology Compass, 3*, 979–991.

Entwistle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review, 57*, 72–84.

Entwistle, D. R., & Alexander, K. L. (1994). Winter setback: The racial composition of schools and learning to read. *American Sociological Review, 59*, 446–460.

Entwistle, D. R., Alexander, K. L., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(2), 167–180. doi:10.1177/000312240707200202

Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long, J. (2010). Creating a progress-monitoring system in reading for middle school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research & Practice, 25*(2), 60–75.

Feinberg, L. (1990). Multiple-choice and its critics: Are the alternatives any better? *College Board Review, 157,* 12–17.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* (4th ed.) Los Angeles: SAGE.

Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading, 10*(3), 323–330.

Foorman, B. A., Fletcher, J. M., Francis, D. J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90*(1), 37–55. Retrieved from: http://www.mona.uwi.edu/cop/sites/default/files/resource/files/Foorman_Role_of_Instruction.pdf

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17. doi:10.1037/0022-0663.88.1.3

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 27–48.

Gardner, D. P. (1983). *A nation at risk: The imperative for educational reform.* An Open Letter to the American People. A Report to the Nation and the Secretary of Education. Washington, D. C.: Department of Education.

Ghosh, D. S. (2013). Socioeconomic status links to children's literacy development. *Washington University Undergraduate Research Digest, 9*(1), 10–15.

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*, 126–135.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7,* 6–10.

Graney, S. B., Missall, K. N., Martínez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology*, *47*(2), 121–142. doi:10.1016/j.jsp.2008.12.001

Guthrie, J. T., Seifert, M., Burnham, M. A., & Caplan, R. I. (1974). The maze technique to assess, monitor reading comprehension. *The Reading Teacher, 28*(2), 161–168.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309–334.

Hall, L. H. (2004). Comprehending expository text: Promising strategies for struggling readers and students with reading disabilities. *Reading Research and Instruction 44*, 75–95.

Hamilton, L. S., & Koretz, D. M. (2002). Tests and their use in test-based accountability. In L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.), *Making sense of test-based accountability in education.* (pp. 13–49). Santa Monica, CA: RAND.

Hart, S. A., Soden, B., Johnson, W., Schatschneider, C., & Taylor, J. (2013). Expanding the environment: Gene × school-level SES interaction on reading comprehension. *The Journal of Child Psychology and Psychiatry, 54*(10), 1047–1055. doi: 0.1111/jcpp.12083

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636–644. doi: 10.1598/RT.59.7.3

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

Heike, T. (2013). New reasons to dislike multiple-choice testing. Retrived from https://www.edutopia.org/blog/reasons-to-dislike-multi-choice-terrell-heick

Heyns, B. (1978). *Summer learning and the effects of schooling.* New York: Academic Press.

Hu, L-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Income Eligibility Guidelines, 79 Fed. Reg. § 43 (March 5, 2014), 12467–12469. Retrieved from https://www.fns.usda.gov/sites/default/files/2014-04788.pdf

Jamar, I. (1994). *Fall testing: Are some students differentially disadvantaged?* Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Jennings, J. H., Caldwell, J., & Lerner, J. W. (2006). *Reading problems assessment and teaching strategies* (5th ed.). Boston: Pearson Education.

Jipson, J. A., & Paley, N. (1992). Is there a base to today's literature-based reading programs? *English Education, 24*(2), 77–90.

Judge, S., & Bell, S. M. (2011). Reading achievement trajectories for students with learning disabilities during the elementary school years. *Reading and Writing Quarterly, 27*, 153–178. doi: 10.1080/10573569.2011.532722

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. doi:10.1037/0022-0663.80.4.437

Kamil, M., & Lane, D. (1997, December). *Using information text for first grade reading instruction: Theory and practice.* Paper presented at the National Reading Conference, Scottsdale, AZ.

Langer, J. A. (1990). The process of understanding: Reading for literary and informative purposes. *Research in the Teaching of English, 24,* 229–260.

Langer, J. A. (1991). *Literary understanding and literature instruction.* (Report Series no. 2.11). Albany, NY: National Research Center on Literature Teaching and Learning.

Langer, J. A. (1992). *Literature instruction: A focus on student response.* Urbana, IL: National Council of Teachers of English.

Lauen D. L., & Gaddis, S. M. (2011). Shining a light or fumbling in the dark? The effects of NCLB's subgroup-specific accountability on student achievement. *Educational Evaluation and Policy Analysis, 34*(2), 185–208.

Maloch, B., & Bomer, R. (2013). Informational texts and the Common Core Standards: What are we talking about, anyway? *Language Arts, 3*(90), 205–213.

Marzano, R. J. (2004). *Building background knowledge for academic achievement*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

McCormick, S. (2007) *Instructing students who have literacy problems* (5th ed.). New Jersey: Pearson Merrill Prentice Hall.

McCoubrie, P. (2004). Improving the fairness of multiple-choice questions: A literature review. *Medical Teacher, 26*(8), 709–712.

McKenna, M. C., & Stahl, K. A. (2009). *Assessment for reading instruction.* New York: The Guilford Press.

Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly, 31*, 187–198.

Morrow, L. M. (1992). The impact of a literature-based program on literary achievement, use of literature, and attitudes of children from minority backgrounds. *Reading Research Quarterly, 27*, 250–275.

Myers L., & Botting, N. (2008). Literacy in the mainstream inner-city school: Its relationship to spoken language. *Child Language Teaching and Therapy, 24*, 95–114.

National Assessment of Educational Progress. (2015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

National Center for Education Statistics. (2013). *The Nation's Report Card: Trends in Academic Progress 2012.* (NCES 2013 456). Institute of Education Sciences, U. S. Department of Education, Washington, DC. Retrieved from http://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf

National Forum on Education Statistics. (2015). *Forum Guide to Alternative Measures of Socioeconomic Status in Education Data Systems.* (NFES 2015-158). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: Authors.

Neild, R. (2009). Falling off track during the transition to high school: What we know and what can be done. *Future of Children, 19*(1), 53–76. Retrieved from http://www.futureofchildren.org/publications/journals/article/index.xml?journalid=30&articleid=38&sectionid=84

Nese, J., Biancarosa, G., Anderson, D., Lai, C., Alonzo, J., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading & Writing, 25*(4), 887–915. doi:10.1007/s11145-011-9304-0

Nese, J., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across Grades 1–8. *Journal of School Psychology, 51*(5), 625–642. doi:10.1016/j.jsp.2013.05.006

Newkirk, T. (1989). *More than stories: The range of children's writing*. Portsmouth, NH: Heinemann.

Olson, L. (2005). Benchmark assessments offer regular achievement. *Education Week, 25*(13), 13–14.

Pappas, C. C. (1993). Is narrative primary? Some insights from kindergartners' pretend readings of stories and information books. *Journal of Reading Behavior, 25*, 95–129.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*, 184–202. doi:10.1080/10888438.2011.536128

Petscher, Y., Kershaw, S., Koon, S., & Foorman, B. R. (2014). Testing the Importance of Individual Growth Curves in Predicting Performance on a High-Stakes Reading Comprehension Test in Florida. REL 2014-006. Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research, 84*(2), 203–244. doi:10.3102/0034654313509492

Pickens, T. W. (2016). *The impact of race and socioeconomic status on the reading comprehension growth trajectories of adolescents* (Doctoral dissertation). Retrieved from http://jewlscholar.mtsu.edu/handle/mtsu/4874

Powell-Smith, K. A., Stoner, G., Shinn, M. R., & Good III, R. H. (2000). Parent tutoring in reading using literature and curriculum materials: Impact on student reading achievement. *School Psychology Review, 29*(1), 5–27.

Preacher, K. J. (2010). Latent growth curve models. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 185–198). London: Routledge. Retrieved from http://quantpsy.org/pubs/preacher_2010.pdf

Preacher, K. P., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling. Quantitative applications in the social sciences*. Thousand Oaks, CA: Sage Publications.

RAND Reading Study Group. (2002). *Reading for understanding*. Washington, DC: RAND Education.

Rathvon, N. (2004). *Early reading assessment*. New York: The Guilford Press.

Ravitch, D. (1981). The meaning of the new Coleman report. *The Phi Delta Kappan, 62*(10), 718–720. Retrieved from https://www.jstor.org/stable/pdf/20386113.pdf

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*(1), 16–20.

Rescorla, L., & Rosenthal, A. S. (2004). Growth in standardized ability and achievement test scores from 3rd to 10th grade. *Journal of Educational Psychology, 96*(1), 85–96.

Ruddell, M. R. (2001). *Teaching content reading and writing.* New York: John Wiley & Sons, Inc.

Ruddell, M. R., & Unrau, N. J. (Eds.). (2004). *Theoretical models and processes of reading* (5th ed.) Newark, DE: International Reading Association.

Scarborough, H. (2002). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice. In S. Newman & D. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York: Guilford Press.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Educational Researcher, 21*, 22–27. doi:10.3102/0013189X021004022

Shinn, M. R., Good, R. H., III, Knutson, N., & Tilly, W. D., III. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*(3), 459−479. Retrieved from http://www.specialistedpsy.com/shinnetalSPR1992.pdf-link.pdf

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453. doi:10.3102/00346543075003417

Smith, F. (2004). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (6th ed.). Mahwah, N. J.: Erlbaum.

Snyder, T., & Musu-Gillette, L. (2015). Free or reduced price lunch: A proxy for poverty? National Center for Educational Statistics Blog, April 16, 2015. Retrieved from: http://nces.ed.gov/blogs/nces/post/free-or-reduced-price-lunch-a-proxy-for-poverty

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407. doi:10.1598/RRQ.21.4.1

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guilford Press.

Thomas, J. A. (1962). Efficiency in education: an empirical study. *Administrator's Notebook, 11*, 1–4.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education, 2013*, 1–29.

Torgesen, J. K. (1997). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator, 2*(1 & 2), 32–39.

U.S. Congress Office of Technology Assessment. (1992). Lessons from the past: A history of educational testing in the United States. In *Testing in American Schools: Asking the Right Questions* (pp. 103-131). Washington, D.C.: U.S. Government Printing Office. Retrieved from https://www.princeton.edu/~ota/disk1/1992/9236/923606.PDF

Vacca, R. T., & Vacca, J. L. (1999). *Content area reading: Literacy and learning across the curriculum* (6th ed). Menlo Park, CA: Longman.

Viadero, D. (2006). Race report's influence felt 40 years later: Legacy of Coleman study was new view of equity. *EdWeek, 41*(1), 21–24.

Walberg, H. J., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal, 20*, 359–373. doi:10.3102/00028312002003359

Walmsley, S. A., & Adams, E. L. (1993). Realities of 'whole language.' *Language Arts, 70*(4), 272–280.

Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey-Bass.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461–481.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872.

Wilder, J. (1962). Basimetric approach (law of initial value) to biological rhythms. *Annals of the New York Academy of Sciences, 98,* 1211–1228. doi:10.1111/j.1749-6632.1962.tb30629.x

Willet, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363–381. Retrieved from http://personal.psu.edu/jxb14/M554/articles/Willett%26Sayer1994.pdf

Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics, 19,* 602–632.

Williamson, G. L., Appelbaum, M., & Epanchin, A. (1991). Longitudinal analyses of academic achievement. *Journal of Educational Measurement, 28,* 61–76. doi:10.1111/j.1745-3984.1991.tb00344.x

Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2014). Student reading growth illuminates the Common Core text-complexity standard: Raising both bars. *The Elementary School Journal, 11*(2), 230–254.

APPENDICES

APPENDIX A

COMMON CORE STATE STANDARDS: ENGLISH LANGUAGE ARTS

STANDARDS, GRADES 9-12

Adapted from http://www.corestandards.org/ELA-Literacy/

Key Ideas and Details

1. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
2. Determine central ideas or themes of a text and analyze their development; summarize the key supporting details and ideas.
3. Analyze how and why individuals, events, or ideas develop and interact over the course of a text.

Craft and Structure:

4. Interpret words and phrases as they are used in a text, including determining technical, connotative, and figurative meanings, and analyze how specific word choices shape meaning or tone.
5. Analyze the structure of texts, including how specific sentences, paragraphs, and larger portions of the text (e.g., a section, chapter, scene, or stanza) relate to each other and the whole.
6. Assess how point of view or purpose shapes the content and style of a text.

Integration of Knowledge and Ideas:

7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.
8. Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.
9. Analyze how two or more texts address similar themes or topics in order to build knowledge or to compare the approaches the authors take.

Range of Reading and Level of Text Complexity:

10. Read and comprehend complex literary and informational texts independently and proficiently.

APPENDIX B

COMMON CORE STATE STANDARDS: ENGLISH LANGUAGE ARTS CODES, GRADES 9-12

Common Core State Standards Codes for ELA 9-12 Expository (RI codes) and Narrative (RL codes) texts. From http://www.corestandards.org/ELA-Literacy/

RI codes:

RI.9-10.1. Cite strong and thorough textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.

RI.9-10.2. Determine a central idea of a text and analyze its development over the course of the text, including how it emerges and is shaped and refined by specific details; provide an objective summary of the text.

RI.9-10.3. Analyze how the author unfolds an analysis or series of ideas or events, including the order in which the points are made, how they are introduced and developed, and the connections that are drawn between them.

RI.9-10.4. Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the cumulative impact of specific word choices on meaning and tone (e.g., how the language of a court opinion differs from that of a newspaper).

RI.9-10.5. Analyze in detail how an author's ideas or claims are developed and refined by particular sentences, paragraphs, or larger portions of a text (e.g., a section or chapter).

RI. 9-10.6. Determine an author's point of view or purpose in a text and analyze how an author uses rhetoric to advance that point of view or purpose.

RI.9-10.7. Analyze various accounts of a subject told in different mediums (e.g., a person's life story in both print and multimedia), determining which details are emphasized in each account.

RI.9-10.8. Delineate and evaluate the argument and specific claims in a text, assessing whether the reasoning is valid and the evidence is relevant and sufficient; identify false statements and fallacious reasoning.

RI.9-10.9. Analyze seminal U.S. documents of historical and literary significance (e.g., Washington's Farewell Address, the Gettysburg Address,

Roosevelt's Four Freedoms speech, King's "Letter from Birmingham Jail"), including how they address related themes and concepts.

RI.9-10.10. By the end of grade 9, read and comprehend literary nonfiction in the grades 9-10 text complexity band proficiently, with scaffolding as needed at the high end of the range.

RI.11-12.1. Cite strong and thorough textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text, including determining where the text leaves matters uncertain.

RI.11-12.2. Determine two or more central ideas of a text and analyze their development over the course of the text, including how they interact and build on one another to provide a complex analysis; provide an objective summary of the text.

RI.11-12.3. Analyze a complex set of ideas or sequence of events and explain how specific individuals, ideas, or events interact and develop over the course of the text.

RI.11-12.4. Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze how an author uses and refines the meaning of a key term or terms over the course of a text (e.g., how Madison defines faction in Federalist No. 10).

RI.11-12.5. Analyze and evaluate the effectiveness of the structure an author uses in his or her exposition or argument, including whether the structure makes points clear, convincing, and engaging.

RI.11-12.6. Determine an author's point of view or purpose in a text in which the rhetoric is particularly effective, analyzing how style and content contribute to the power, persuasiveness or beauty of the text.

RI.11-12.7. Integrate and evaluate multiple sources of information presented in different media or formats (e.g., visually, quantitatively) as well as in words in order to address a question or solve a problem.

RI.11-12.8. Delineate and evaluate the reasoning in seminal U.S. texts, including the application of constitutional principles and use of legal reasoning (e.g., in U.S. Supreme Court majority opinions and dissents) and the premises, purposes, and arguments in works of public advocacy (e.g., *The Federalist*, presidential addresses).

RI.11-12.9. Analyze seventeenth-, eighteenth-, and nineteenth-century foundational U.S. documents of historical and literary significance (including The Declaration of Independence, the Preamble to the Constitution, the Bill of Rights,

and Lincoln's Second Inaugural Address) for their themes, purposes, and rhetorical features.

RI.11-12.10. By the end of grade 11, read and comprehend literary nonfiction in the grades 11-CCR text complexity band proficiently, with scaffolding as needed at the high end of the range.

RL codes:

RL.9-10.1. Cite strong and thorough textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text.

RL.9-10.2. Determine a theme or central idea of a text and analyze in detail its development over the course of the text, including how it emerges and is shaped and refined by specific details; provide an objective summary of the text.

RL.9-10.3. Analyze how complex characters (e.g., those with multiple or conflicting motivations) develop over the course of a text, interact with other characters, and advance the plot or develop the theme.

RL.9-10.4. Determine the meaning of words and phrases as they are used in the text, including figurative and connotative meanings; analyze the cumulative impact of specific word choices on meaning and tone (e.g., how the language evokes a sense of time and place; how it sets a formal or informal tone).

RL.9-10.5. Analyze how an author's choices concerning how to structure a text, order events within it (e.g., parallel plots), and manipulate time (e.g., pacing, flashbacks) create such effects as mystery, tension, or surprise.

RL.9-10.6. Analyze a particular point of view or cultural experience reflected in a work of literature from outside the United States, drawing on a wide reading of world literature.

RL.9-10.7. Analyze the representation of a subject or a key scene in two different artistic mediums, including what is emphasized or absent in each treatment (e.g., Auden's "Musée des Beaux Arts" and Breughel's Landscape with the Fall of Icarus).

(RL.9-10.8 not applicable to literature)

RL.9-10.9. Analyze how an author draws on and transforms source material in a specific work (e.g., how Shakespeare treats a theme or topic from Ovid or the Bible or how a later author draws on a play by Shakespeare).

RL.9-10.10By the end of grade 9, read and comprehend literature, including stories, dramas, and poems, in the grades 9-10 text complexity band proficiently, with scaffolding as needed at the high end of the range.

RL.11-12.1. Cite strong and thorough textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text, including determining where the text leaves matters uncertain.

RL.11-12.2. Determine two or more themes or central ideas of a text and analyze their development over the course of the text, including how they interact and build on one another to produce a complex account; provide an objective summary of the text.

RL.11-12.3. Analyze the impact of the author's choices regarding how to develop and relate elements of a story or drama (e.g., where a story is set, how the action is ordered, how the characters are introduced and developed).

RL.11-12.4. Determine the meaning of words and phrases as they are used in the text, including figurative and connotative meanings; analyze the impact of specific word choices on meaning and tone, including words with multiple meanings or language that is particularly fresh, engaging, or beautiful. (Include Shakespeare as well as other authors.)

RL.11-12.5. Analyze how an author's choices concerning how to structure specific parts of a text (e.g., the choice of where to begin or end a story, the choice to provide a comedic or tragic resolution) contribute to its overall structure and meaning as well as its aesthetic impact.

RL.11-12.6. Analyze a case in which grasping a point of view requires distinguishing what is directly stated in a text from what is really meant (e.g., satire, sarcasm, irony, or understatement).

Analyze multiple interpretations of a story, drama, or poem (e.g., recorded or live production of a play or recorded novel or poetry), evaluating how each version interprets the source text. (Include at least one play by Shakespeare and one play by an American dramatist.)

(RL.11-12.8 not applicable to literature)

RL.11-12.9. Demonstrate knowledge of eighteenth-, nineteenth- and early-twentieth-century foundational works of American literature, including how two or more texts from the same period treat similar themes or topics.

RL.11-12.10. By the end of grade 11, read and comprehend literature, including stories, dramas, and poems, in the grades 11-CCR text complexity band proficiently, with scaffolding as needed at the high end of the range.

APPENDIX C

IRB EXEMPTION DETERMINATION NOTICE

IRB
INSTITUTIONAL REVIEW BOARD
Office of Research Compliance,
010A Sam Ingram Building,
2269 Middle Tennessee Blvd
Murfreesboro, TN 37129

MIDDLE TENNESSEE STATE UNIVERSITY

IRBN007 – EXEMPTION DETERMINATION NOTICE

Friday, February 17, 2017

Investigator(s): Laura Briggs; Jwa Kim
Investigator(s') Email(s): lcc2a@mtmail.mtsu.edu; jwa.kim@mtsu.edu
Department: Literacy Studies

Study Title: A COMPARISON OF NARRATIVE AND EXPOSITORY TEXT COMPREHENSION FOR STUDENTS AT VARYING LEVELS OF SES: A LATENT GROWTH CURVE ANALYSIS
Protocol ID: **17-1158**

Dear Investigator(s),

The above identified research proposal has been reviewed by the MTSU Institutional Review Board (IRB) through the **EXEMPT** review mechanism under 45 CFR 46.101(b)(2) within the research category *(4) Study involving existing data* A summary of the IRB action and other particulars in regard to this protocol application is tabulated as shown below:

IRB Action	EXEMPT from furhter IRB review***	
Date of expiration	NOT APPLICABLE	
Participant Size	Existing Data	
Participant Pool	**Existing data provided by Discovery**	
Mandatory Restrictions	Annonymous, de-identified; randomized data covered under provided permission letter on file with the Research Compliance Office	
Additional Restrictions	None	
Comments	None	
Amendments	**Date**	**Post-Approval Amendments**
		None at this time

***This exemption determination only allows above defined protocol from further IRB review such as continuing review. However, the following post-approval requirements still apply:

- Addition/removal of subject population should not be implemented without IRB approval
- Change in investigators must be notified and approved
- Modifications to procedures must be clearly articulated in an addendum request and the proposed changes must not be incorporated without an approval
- Be advised that the proposed change must comply within the requirements for exemption
- Changes to the research location must be approved – appropriate permission letter(s) from external institutions must accompany the addendum request form
- Changes to funding source must be notified via email (irb_submissions@mtsu.edu)

- The exemption does not expire as long as the protocol is in good standing
- Project completion must be reported via email (irb_submissions@mtsu.edu)
- Research-related injuries to the participants and other events must be reported within 48 hours of such events to compliance@mtsu.edu

The current MTSU IRB policies allow the investigators to make the following types of changes to this protocol without the need to report to the Office of Compliance, as long as the proposed changes do not result in the cancellation of the protocols eligibility for exemption:

- Editorial and minor administrative revisions to the consent form or other study documents
- Increasing/decreasing the participant size

The investigator(s) indicated in this notification should read and abide by all applicable post-approval conditions imposed with this approval. Refer to the post-approval guidelines posted in the MTSU IRB's website. Any unanticipated harms to participants or adverse events must be reported to the Office of Compliance at (615) 494-8918 within 48 hours of the incident.

All of the research-related records, which include signed consent forms, current & past investigator information, training certificates, survey instruments and other documents related to the study, must be retained by the PI or the faculty advisor (if the PI is a student) at the sacure location mentioned in the protocol application. The data storage must be maintained for at least three (3) years after study completion. Subsequently, the researcher may destroy the data in a manner that maintains confidentiality and anonymity. IRB reserves the right to modify, change or cancel the terms of this letter without prior notice. Be advised that IRB also reserves the right to inspect or audit your records if needed.

Sincerely,

Institutional Review Board
Middle Tennessee State University

Quick Links:

Click here for a detailed list of the post-approval responsibilities.
More information on exmpt procedures can be found here.