ABSTRACT

Title of dissertation: THE CONTRIBUTION OF EXECUTIVE FUNCTION TO READING COMPREHENSION FOR LINGUISTICALLY DIVERSE LEARNERS

Anna Genevieve Meyer, Doctor of Philosophy, 2018

Dissertation directed by: Professor Rebecca Silverman
Department of Counseling, Higher Education, and Special Education

Research suggests that EF can aid in the prediction of RC. However, much of the existing research into the relationship between these two variables has relied on statistical correlations and simple linear regression, neither of which fully capture the complexity of their theoretical relationship and other known contributors to RC, such as decoding and linguistic comprehension. Accordingly, this dissertation study investigated the relationship between EF and RC through a synthesis of the literature and two separate empirical studies. The first empirical study investigate whether (1) a latent construct of EF, measured by separate assessments of working memory, shifting, and inhibition, makes unique direct contribution to the prediction of RC and (2) whether EF's latent construct mediates the prediction of RC through decoding and a latent construct of linguistic comprehension. The second empirical study investigated whether (1) a latent construct of EF, measured by separate assessments of working memory, shifting, and inhibition, makes unique direct contribution to the prediction of RC and (2) whether EF's latent construct mediates the prediction of RC through decoding and a latent construct of linguistic comprehension. Both empirical studies examined this relationship in linguistically diverse learners (LDLs) as an understudied population to extend the current research base. Specifically, the sample included three groups of LDL students: (a) English Learners (ELs), or students who speak a language other than or in addition to English in the home and who are receiving school-based English language services because they have not passed an English language proficiency exam, (b) R-ELs, or students who speak a language other than or in addition to English in the home but have passed an English language proficiency exam and have thus been recently exited from EL services, and (c) EL students from the above cohorts who the school identified as having a disability. Language and disability status, respectively, were entered as moderators in the above models to test for significant differences by group. Limitations of the dissertation study and directions for future research are discussed.

THE CONTRIBUTION OF EXECUTIVE FUNCTION TO READING COMPREHENSION FOR LINGUISTICALLY DIVERSE LEARNERS

by

Anna Genevieve Meyer

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:

Professor Rebecca Silverman, Chair
Professor Kelli Cummings
Professor Jeffery Harring
Professor Joan Leiber
Professor Ana Taboda-Barber

Acknowledgements

This work would not have been possible without the support of my advisory committee, family, colleagues, and friends. I am especially grateful to Dr. Rebecca Silverman, Associate Professor in the Department of Counseling, Higher Education, and Special Education, who served as the chair of this dissertation, and as my advisor and mentor at the University of Maryland, College Park. Dr. Silverman's leadership and scholarship in the field of literacy research has served as a model for embarking on my own academic path. I would also like to thank Dr. Harring and Dr. Proctor, for their guidance on the formation of my dissertation study and the opportunity to learn from their expertise in my work on the CLAVES Research Project. In addition, I would like to thank each of the members of my Dissertation Committee for their advice and recommendations in shaping this dissertation study. This process has taught me how important collaboration and critique are to the scientific process of reading research. I would also like to thank my family for their boundless love and encouragement as I pursued my graduate studies. Finally, I would like to thank my husband for his partnership and support of all of my academic goals.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

This dissertation explores the relationship between executive function (EF) and reading comprehension (RC) for linguistically diverse learners (LDLs) in upper elementary school. Given the importance of RC in students' academic success, it is important to understand how cognitive factors such as EF might affect RC. In turn, these results may lead to improved identification methods and interventions for students at risk of developing reading difficulties or those who have failed to respond to specific RC interventions. This dissertation makes a significant contribution to the literature for several reasons. First, the population of LDLs is understudied in the literature, and many studies that do include LDLs do not disaggregate results by language status. Second, the age range represents a critical period in RC development. Third, the analytic methods used for investigating these constructs in this study, including latent variable structural equation modeling that permits testing for direct and indirect, or mediation effects, are a major contribution to the literature to date.

Statement of the Problem

Research has consistently shown that students entering fourth grade confront a fundamental shift in the cognitive demands of reading academic text (Chall, 1983) and, often distressingly, demonstrate a 'fourth grade slump' in academic performance (Chall & Jacobs, 2003). Numerous studies have termed RC difficulties that appear in upper elementary school as 'late emerging' and have also investigated the condition's nature and prevalence (Catts, Compton, Tomblin, & Bridges, 2012; Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Etmanski, Partanen, & Siegel, 2014; Leach, Scarborough, & Rescorla, 2003). Together, these studies support hypotheses that upper elementary school reading deficits may emerge as a result of the increased text complexity associated with later grades. However, they also point to possible errors in sensitivity and specificity at earlier screenings or developmental changes in reading comprehension.

English learners (ELs), defined as students who speak a language other than English in the home and have been identified by their school districts as need support due to limited English language proficiency, represent 9.3 percent of the student population enrolled in United States public schools as of 2013, and are at risk of experiencing difficulty in reading, especially as they progress into upper elementary school and beyond (Kena et al., 2016). However, ELs only represent a portion of the linguistically diverse students in the United States. In the present manuscript, I use the term linguistically diverse learners (LDLs) to indicate a diverse group of students at various levels of English proficiency who speak a language other than English in the home and who could benefit from additional English language development support in order to access academic context in school. Within the breadth of LDLs, I use the term EL to refer to students who are currently receiving identified as needing English for Speakers of Other Languages (ESOL) services, I use the term re-classified English learners (R-EL) to refer to those students who have achieved a level of English proficiency to test out ESOL services, in other words, they have been recently exited from the ESOL program, in the past two years. These students often need additional English language development support to continue to attain the level of English needed to access complex content in school. I use the term English only (EO) to refer to students who report English as the only language spoken in the home, but acknowledge that this group represents students across racial/ethnic divisions. Although this dissertation focuses on identifying reading comprehension difficulties within LDL populations, I align this manuscript with multilingual theory that celebrates and seeks to leverage ELs' diverse cultural and linguistic assets in the acquisition of learning English (e.g., Valdés, Capitelli, & Alvarez, 2011). Recognizing the linguistic diversity within the U.S. population is discussed further in this introductory chapter, and is reflected in the decisions made in my synthesis and in the construction of research questions for the two empirical studies.

Historically, word reading and linguistic comprehension have been accepted as the strongest predictors of RC difficulties in upper elementary school (e.g., Catts, Adlof, & Weismer, 2006; Carver, 1998; Kendeou, Savage, & Broek, 2009; Verhoeven & Van Leeuwe, 2008), a model commonly referred to as the Simple View of Reading (Gough & Tunmer, 1986). Children struggling to decode will likely have difficulty with RC because they are focusing all of their energy on simply reading the words. Children grappling with listening comprehension are also likely to have difficulty with RC because they lack the underlying linguistic framework to process meaning in written text.

However, heterogeneous profiles of students with late-emerging reading difficulties complicates identification and matching to appropriate interventions based on students' areas of need. For instance, Catts et al. (2012) and Leach et al. (2003) estimate between one-third to one-half of students with late-emerging reading disabilities had deficits in RC alone, roughly one-third of students had deficits in word reading alone, and one-tenth to one-third had deficits in both comprehension and word reading. These discrepancies have spurred researchers to re-examine current RC models to ensure that students at-risk for reading failure are identified accurately and as early as possible. One concern that has surfaced is that there may be some students who demonstrate adequate decoding and linguistic comprehension, but still experience difficulties, perhaps because they have an inability to coordinate these two sources of information needed for reading comprehension. Research on cognitive processes that may be needed for reading comprehension, over and above decoding and linguistic comprehension, holds promise for discovering other contributors to the emergence of reading difficulty in upper elementary school.

Executive functioning is a cognitive process in which individuals integrate multiple skills in order to reach a goal. Miyake et al., (2000), in their seminal work, suggested that there are three domains within EF: shifting, inhibition, and updating. Shifting means the ability to transfer between different mental processes flexibly and is also described as cognitive flexibility or switching (Brookshire, Levin, Song, & Zhang, 2004). Inhibition is the ability to resist intrusion of competing ideas or prompts (Friedman et al., 2006). Updating is the ability to refresh and retain information in the working memory (Brookshire et al., 2004). These domains were originally studied in adults, and have only recently been applied to children, especially in the area of reading development (Etmanski et al., 2014; Locascio, Mahone, Eason, & Cutting, 2010; Sesma, Mahone, Levine, Eason, & Cutting, 2009). EF is hypothesized as particularly important for students in upper elementary school because the demands of acquiring meaning from text requires flexible manipulation of multiple cognitive processes, the ability to update and refresh information in the working memory, and the faculty to block the intrusion of unnecessary information through inhibitory control.

Investigating the contribution of EF to RC for LDLs in upper elementary school represents another step towards improving the prediction accuracy of screening batteries used to identify students at-risk of developing reading difficulties. Linking the results of assessment batteries to interventions that target specific skills strengthens the ecological validity of universal and secondary screening procedures. Simultaneously, in order to remediate and provide effective support for students struggling to read, it is important for teachers to understand why a student is having difficulty understanding text. If EF is found as a unique contributor to RC ability, then researchers and educators who screen for EF capabilities alongside traditional measures of reading proficiency will be better positioned to understand students' reading and cognitive profiles. Subsequently, this will allow teachers to make more informed decisions about the components of RC interventions to target those skills, whether they are a combination of or exclusively in EF, decoding, or linguistic comprehension.

Reading Comprehension

The theoretical framework guiding the work reported in this dissertation is the Simple View of Reading (SVR). In this model, RC is the product of skillful decoding and linguistic comprehension (Gough & Tunmer, 1986). Numerous studies have confirmed the accuracy of the SVR, including single time point (e.g., Tunmer & Chapman, 2012) and longitudinal studies of students in elementary school (e.g., Catts, Herrera, Nielson, & Bridges, 2015), studies of secondary students and adults (e.g., Braze et al., 2016), and genetic studies of twins (e.g., Harlaar, et al., 2010). A description of the components of the SVR is provided below.

Decoding. A major aspect of RC is the ability to decode or use phonological and orthographic information to read words. Research has consistently found that early decoding is a predictor of RC. However, decoding must be completed in an automatic or fluent manner so that students can focus on comprehending the meaning of words (Pikulski & Chard, 2005; Pinnell et al., 1995). Without fluency, students are thought to be competing between trying to decode the word and access the meaning of the word (semantics). The importance of fluency (i.e., the ability to decode words fluently) to RC was first postulated by LaBerge and Samuels (1974) in the Automatic Information Processing Theory, which other researchers (e.g., Stanovich, 2000) have studied and expanded over time. As explained by Fuchs, Fuchs, Hosp, and Jenkins (2001), "low-level word recognition frees up capacity for higher level, integrative comprehension processing of text" (p. 242). In fact, studies have found that fluency mediates the relationship between decoding and RC for students in upper elementary school (Silverman, Speece, Harring, & Ritchey, 2012).

Linguistic Comprehension. Linguistic comprehension is itself a multidimensional construct, but essentially refers to students' understanding of language, including words, sentences, and verbal communication (Gough & Tunmer, 1986). Since Gough and Tunmer's original publication (1986), the definition of linguistic comprehension has been expanded to include vocabulary, morphology, and syntax. For instance, Tunmer and Chapman (2012) conducted an exploratory factor analysis that found vocabulary and listening comprehension were components of linguistic comprehension in a sample of elementary school students. In a longitudinal study of students from early elementary school through middle school, Adlof, Catts, and Lee (2010) found that two separate measures of syntax were predictive of reading comprehension, although the type of syntactic measure varied at different points in development. Similarly, Proctor, Silverman, Harring, and Montecillo (2012) found that syntax also made significant contributions to reading comprehension in a study of mid- to upper-elementary students. Despite the evidence for an expanded definition of linguistic comprehension, many studies conducted to date have used individual measures of linguistic comprehension in the prediction of RC. Given the complexity of linguistic comprehension, models that account for its diversity across tasks, such as by forming a latent construct of linguistic comprehension, will position researchers to more accurately reflect the full model of RC and its subcomponents. Latent constructs allow researchers to operationalize and measure an underlying construct that is not directly measurable by a single task (Bollen, 2002).

Linguistic Interdependence. According to the Linguistic Interdependence Hypothesis (Cummins, 1979), students learning English as a second or other language are able to draw on linguistic awareness in their native language. Hoover and Gough (1990) tested the theoretical model of the SVR with an extant longitudinal dataset of English learners in the primary grades, and found decoding and linguistic comprehension contributed to RC outcomes in the same way as it had for English monolingual students. The study also supported the hypothesis that linguistic comprehension was increasingly important to the progress of RC as students advanced through elementary school. Linguistic comprehension has been found to contribute to ELs' RC outcomes in English (Proctor, Carlo, August, & Snow, 2005; Silverman et al., 2015).

This effect has also been replicated to some degree in studies of students' Spanish literacy on English literacy. Proctor, August, Snow, and Barr (2010) present a theoretical path model of the direct and indirect effects of oral language (or linguistic competence) and alphabetic knowledge (or decoding) in English and Spanish on both Spanish and English RC outcomes. A meta-analysis of 86 studies by Prevoo, Malda, Mesman, and van IJzendoorn (2016) found an average positive effect size of 0.08 for associations between first language proficiency on second language proficiency, although the association was not positive for all studies. In a study conducted with struggling readers in sixth grade that included both EL and EO students, Lesaux and Kieffer (2010) found both groups had low, generalized vocabulary scores compared to national norms, but adequate word reading skills. Across these studies is the finding that ELs are a heterogeneous group, and that first language skills are sometimes, but not always, associated with higher performance in a second language. This dissertation seeks to expand the research base on ELs language performance in English and its relation to cognitive skills, namely executive functions.

Executive Function

EF is multifaceted construct. Baggetta and Alexander (2016) provide a thorough review of the many models of the components of EF. Their synthesis found that the most commonly described components are working memory, updating, and shifting. Mikaye et al. (2000) proposed the most frequently cited model, which suggests that EF includes shifting, inhibition, and updating. In a research synthesis I conducted (see Chapter 2 of this manuscript), other components of EF that were implicated in the prediction of reading comprehension include processing speed, planning, and attention rating. Ultimately, however, I chose to focus on the contribution of working memory, shifting, and inhibition to the latent construct of EF because they were most frequently significant in the prediction models across the studies I reviewed. There is also theoretical evidence that the other dimensions may be subsumed under these three domains. Below, I present a brief overview of each component of EF and how each is theoretically and empirically implicated in the process of reading.

Working Memory. Working memory is the ability to store, update, and recall information rapidly in the mind (Diamond, 2013). Working memory is separate from long-term memory, which is responsible for storing and later recalling information over longer periods of time. Working memory is related but separate from short-term memory, which merely holds information in a static sense. Using a computer-based analogy, working memory can be thought of as the “buffering” process that the mind undergoes to hold, manipulate, and make sense of new information in the context of previous knowledge, or long term memory (Baddeley, 2000). Working memory has been shown to be associated with and predict RC for children in elementary school (Cain, Oakhill, & Bryant, 2004; St Clair-Thompson & Gathercole, 2006) and in a meta-analysis of children and adults with RC difficulties (Carretti, Borella, Cornoldi, & De Beni, 2009).

Thinking back to the cognitive demands imposed when decoding and understanding text, it is clear that working memory is deeply connected throughout the process of reading comprehension, potentially in multiple aspects of the process. Studies of children with reading difficulty have suggested that there is theoretical capacity limit of working memory and a potential for individual differences thereof. For instance, Gathercole, Alloway, Willis, and Adams (2006) found working memory predicted RC performance in elementary school aged children with reading difficulties. In the research synthesis I conducted that is presented in Chapter 2 of this manuscript, six out of the seven core studies I reviewed found that working memory contributed uniquely to the prediction of RC for students in upper elementary school, even when controlling for decoding and linguistic comprehension (e.g., Christopher et al., 2012; Gerst, Cirino, Fletcher, & Yoshida, 2015; Jacobson et al., 2016; Nouwens, Groen, & Verhoeven, 2016; Sesma et al., 2009). Although two of the seven studies did not find working memory as predictive of RC, there is a large body of research that has previously linked working memory to RC (Alloway & Alloway, 2010; Cain et al., 2004; Swanson, 1999).

Shifting. Shifting, much like the literal use of the term for an automobile, implies awareness of moving between two or more states or concepts. Shifting has also been referred to in the literature as *switching* and *cognitive flexibility*. For reading, the ability to shift between tasks is hypothesized as important because students need to stop and think about what they have read and whether it made sense in order to monitor their understanding. Additionally, students need to connect flexibly between what they are currently reading and integrate what they know from other contexts, thus shifting their mental focus. In the research synthesis presented in Chapter 2, I found that shifting was investigated as a predictor of RC in five studies and found to be significant in two studies (Kieffer, Vukovic, & Berry, 2013; Nouwens et al., 2016), yet was non-significant in three other studies (Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016).

Inhibition. In a behavioral context, inhibition is the action of restricting a voluntary response to a stimulus. Inhibition is also referred to as *inhibitory control* in some studies (e.g., Kieffer et al., 2013). For reading, inhibition may be involved when students disregard or suppress words and concepts that are associated with what they are reading, but are not the most important elements to consider. As students read increasingly detailed and complex texts, inhibition is even more important as the main idea must be distilled from a host of examples and details. In a study of EF that followed children from preschool children into kindergarten, inhibition measured in preschool was found as significantly and positively correlated with phonemic awareness and letter knowledge measured in kindergarten; however, attention shifting was not significantly related to these skills (Blair & Razza, 2007). When testing the relationship in a regression model, inhibition in preschool was not a significant predictor of either early reading measures in kindergarten, but did show concurrent predictive validity in kindergarten. In Chapter 2 of this manuscript, I present a research synthesis which found that inhibition was investigated as a predictor of RC in five studies and significant in two studies (Cutting, Materek, Cole, Levine, & Mahone, 2009; Kieffer et al., 2013). However, Cutting and colleagues' (2009) inhibition measure was a combination of inhibition and planning, thus confounding the relative impact of each individual skill.

Across EF studies, discrepancies in the definition and operationalization of EF hinder the evaluation of its relationship to reading outcomes. Both Baggetta and Alexander (2016) and Jacob and Parkinson (2015) concluded that the current definition of EF is obscure. Therefore, it is difficult to ascertain the relative contribution of each domain or draw conclusions about which domains might be collapsed into a more parsimonious model. Based on my analysis of the literature, working memory, shifting, and inhibition independently show the most promise of contributing to RC outcomes and, together, represent the full range of concepts implicated in the multidimensional construct of EF. Directions for future research are to include multiple measures for each EF domain in order to form latent constructs and strengthen the construct validity for EF.

Outline of the Dissertation Manuscript

As described in this present chapter (Chapter 1), my dissertation reviews and extends the current literature on the relationship between RC and EF. This manuscript unfolds across five sequential chapters, which are in turn described here. Chapter 1 provides a statement of the problem, theoretical framework, and definition of key terms to be used in subsequent chapters. Chapter 1 also provides an outline of how the subsequent chapters, including the synthesis and two empirical studies, are connected by a common thread and overarching purpose: to systematically examine the quantitative evidence for the contribution of EF to LDL RC outcomes using theory-driven models.

Chapter 2 presents a synthesis of the current literature on the relationship between EF and RC, controlling for decoding and linguistic comprehension. Chapter 2 begins with a description of the theoretical constructs of interest, which include RC, EF, upper elementary school, English learners, and students with disabilities, as well as the studies conducted to date that intersect with one or more of these areas. Chapter 2 also presents an original synthesis I conducted on the correlational research between EF and RC, which was important for several reasons. First, it included studies that controlled for decoding and linguistic comprehension, which are known predictors of RC. Therefore, the studies must show the additive value of EF beyond decoding and linguistic comprehension. Second, the synthesis seeks to explore how the relationship may differ for students at risk of experiencing reading difficulty: English learners and students with disabilities. Third, the synthesis focuses on the upper elementary grades, which are a critical point in academic development.

Findings from the synthesis I conducted suggested that additional work is needed to understand how EF contributes to RC using latent constructs of each term, rather than single measures. Additionally, there is limited research conducted with linguistically diverse students to see whether the contribution of EF was moderated by students' language and disability status. Given the growing number of LDLs in the United States, it is important to uncover explanatory factors in the prediction of RC that may aid in both identification and intervention research as potentially malleable factors.

Chapter 3 describes the first of two empirical studies aimed to address the gaps found in the literature synthesis by simultaneously replicating and extending the current research on the prediction of RC by EF. The following research questions guide the first empirical study: (1) Does EF, as measured by working memory, inhibition, and shifting tasks, make a unique, direct contribution to RC beyond the contributions of linguistic comprehension and decoding, among LDLs? and (2) Does EF, as measured by working memory, inhibition, and shifting tasks, mediate the explanation of RC via linguistic comprehension and/or decoding among LDLs? Each question will also investigate whether language status (R-EL, EL) and/or disability status (IEP vs. typical peers) moderate the unique, indirect contribution of EF to RC via linguistic comprehension and/or decoding. This study is important to the field because very few studies have used latent constructs of EF, RC, and linguistic comprehension while controlling for decoding within a structural equation model (SEM) to test the direct and mediation effects of EF on RC via linguistic comprehension and decoding. For the present study, latent constructs of EF, RC and linguistic comprehension are preferable to observed variables because they approximate the theoretical underlying constructs that cannot be measured in a single test (Bollen, 2002). Additionally, none of the studies to date have used the aforementioned model with an LDL sample and tested whether a multi-tier indicator of students' language status moderated those relationships between the key variables of interest.

Chapter 4 describes a related and second empirical study to investigate whether students' EF moderates the effect of intervention targeting RC. The following research questions guide this second empirical study: (1) Does latent EF, as measured by working memory, inhibition, and shifting, moderate the effect of a supplemental reading intervention on LDL students' latent RC, controlling for linguistic comprehension and decoding? and (2) Does the moderation effect differ for LDL students at different levels of English (EL, R-EL) or by disability status (IEP, non-IEP)? In study two I used the same core dataset and similar methods as described in the first study, but I extended the research by investigating whether EF itself moderates students' reading comprehension in the context of their assignment to a supplemental reading intervention. This study is important to the field because few studies to date have examined whether a latent construct of EF moderates students' change in latent RC over the course of a reading intervention. The study has implications for what elements are incorporated into future reading interventions if students’ relative levels of EF are indeed found to influence their propensity for growth in RC.

Chapter 5 ties all of the findings together from the synthesis and two empirical studies. Chapter 5 also provides overarching conclusions and implications about the relationship between EF and RC for linguistically diverse learners in upper elementary school. This dissertation has sought to evaluate the existing evidence for the contribution of EF to RC outcomes. Further, this dissertation extends the current research base by proposing and conducting two related empirical studies that examine whether latent executive function makes both a direct and indirect contribution to reading comprehension through decoding and linguistic comprehension for upper elementary school LDLs. Finally, this dissertation explores whether EF moderates the effect of a supplemental language-based reading intervention and whether the moderation effect differs for students with lower levels of English proficiency and/or disabilities.

This dissertation makes a significant contribution to the field in several ways. First, the synthesis presents a thorough review of the literature on the relationship between EF and RC controlling for linguistic comprehension and decoding for students in upper elementary school. Previous syntheses have not isolated the results by students with disabilities and language status in upper elementary school. Second, the synthesis identified several gaps in the literature, namely the underrepresentation of students with disabilities and linguistically diverse students, the dearth of empirical studies that include both decoding and linguistic comprehension in models of EF and RC, and the limited number of studies that have represented EF and RC as latent constructs. Third, Study 1 found empirical support for the representation of EF as a latent construct and its direct and mediating contribution to RC through decoding and linguistic contribution. Fourth, Study 2 found empirical support for the moderating effect of EF on students' response to a reading intervention. Fifth, Study 1 and 2 both found differing effects for students with disabilities and limited English language proficiency. In sum, this dissertation provides needed exploration into EF's contribution to RC for linguistically diverse students in upper elementary school. Continuing this line of inquiry with future research may assist in building enhanced methods for identifying and treating linguistically diverse students with reading difficulties that take into account the contribution of executive function.

Chapter 2: Synthesis of Research on Executive Function and Reading Comprehension for Linguistically Diverse Learners in Upper Elementary School

Introduction

Interest in EF has surged in the past decade. As described by Miyake and colleagues (2000), EF is a multicomponent construct consisting of shifting, inhibition, and updating. The compelling work by Miyake et al. (2000) spurred numerous correlational and experimental studies that sought to wed higher levels of EF ability to higher academic achievement and social/emotional well-being. Within this large body of work, EF has been theorized to both predict student outcomes and serve as outcomes in their own right. In this respect, EF has been framed as both the cause of and solution to the problem. For literacy research in particular, EF holds promise because generalized measures of reading achievement, such as RC, are multifaceted constructs that require the successful orchestration of multiple skills. Moreover, students' differential response to reading interventions as a function of their attention has been noted in some studies, but not fully explored as a causal factor in the literature (Al Otaiba & Fuchs, 2006; Greulich et al., 2014; Torgesen et al., 2001). Ratings of student attention have, in some studies, been included as aspects of EF (Cutting et al., 2009; Locascio et al., 2010; Sesma et al., 2009; Toplak, Bucciarelli, Jain, & Tannock, 2008). Further investigating the role of attention, and executive function more broadly, may provide insight to why reading interventions have failed to see gains in standardized reading comprehension outcomes (Greulich et al., 2014; Ritchey et al., 2012). However, the manifold definitions and operationalizations of EF make it difficult to determine exactly how it contributes to reading. Furthermore, EF measures often overlap with one another, and many require basic literacy skills, both of which weakens the statistical inferences to be drawn about relations between the constructs. Friedman and Miyake (2004) have termed this limitation of the research as the “task impurity problem.” With an understanding of the complexity of this topic, this research synthesis seeks to examine the unique contribution of EF measures to RC, while taking into account component skills required for RC. Finally, current syntheses to date (e.g., Jacob & Parkinson, 2015) have excluded linguistically diverse students and students with disabilities, or have not focused on reading outcomes for students in upper elementary school. Therefore, the relationship across studies between EF and RC for these populations has not yet been adequately examined. A description of the theoretical constructs of RC and EF is provided below.

Theoretical Background

Reading Comprehension

According to the Simple View of Reading (SVR; Gough & Tunmer, 1986), reading comprehension involves the successful interaction between decoding and linguistic comprehension, while decoding is defined as the ability to “read isolated words quickly, accurately, and silently” and linguistic comprehension is “the process by which, given lexical information, sentences, and discourses are interpreted” (Gough & Tunmer, 1986, p. 7). The validity of this theoretical model is well supported (e.g., Carver, 1998; Catts et al., 2006; Kendeou et al., 2009; Verhoeven & Van Leeuwe, 2008). In its most basic interpretation, the SVR model suggests children who struggle to decode will likely have difficulty understanding complex texts because they cannot adequately decipher its words. Similarly, children who struggle with listening comprehension are also likely to have difficulty with RC because they lack the underlying linguistic foundation to process the meaning of written text.

Under an expanded definition of RC, such as that presented by Scarborough (2001), other component skills of RC include fluency (e.g., Pikulski & Chard, 2005), vocabulary (e.g., Ouellette & Beers, 2010; Tumner & Chapman, 2012), morphology, and syntax (e.g., Proctor et al., 2012). In the definition of RC for the present study, these factors are subsumed in either decoding or linguistic comprehension. Since quick and accurate decoding results in fluency, fluency is thus specifically subsumed under decoding. Relatedly, since vocabulary, morphology, and syntax are necessary components of linguistic comprehension, these skills are included under the latter term.

Executive Function

Emerging evidence suggests that EF may also play a role in contributing to RC development across age ranges (e.g., Blair & Razza, 2007; Cartwright, 2012; Locascio et al., 2010; Yeniad, Malda, Mesman, van IJzendoorn, & Pieper, 2013), though the delimitation of EF components varies across studies. EF, an umbrella term referring to a constellation of goal-oriented cognitive skills discussed below, does not appear to replace or substitute for any of the SVR equation component skills. However, because RC is a complex task that involves integrating background knowledge with novel information over a sustained period of time, it is likely that relative strengths in EF may contribute to RC outcomes, over and above decoding and linguistic comprehension skills alone.

In Baggetta and Alexander's (2016) recent synthesis of EF as a theoretical construct, they counted a total of 61 references for the definition of EF, but "no single reference cited more than five times" (p. 12). Cartwright (2012) conceptualized EF as "deliberate mental actions" towards a goal and proposed nine subcomponents of EF: "planning, strategic processing, focused attention, inhibition, reflecting on others' perspectives (metacognition), organization, cognitive flexibility, memory, and response to feedback" (p. 25). Locascio et al. (2010) described EF as a "multidimensional construct" that includes three main subcomponents: *response inhibition*, *planning*, and *working memory* (p. 442). Blair and Razza (2007) suggested that EF consists of "attention shifting, working memory, and inhibitory control cognitive processes" that operate under "affectively neutral" conditions (p. 648). Jacob and Parkinson (2015) suggested that EF is "the set of cognitive skills required to direct behavior toward the attainment of a goal" and presented seven aspects: *prioritizing and sequencing*, *inhibiting familiar responses*, *maintaining task-relevant information*, *resisting distractions*, *switching between task goals*, *using information to make decisions*, and *creating abstract rules* (p. 512). As the expanse of this list implies, EF has been theorized to represent a broad array of constructs, but many of these have not been operationalized in terms of measurement and have not been linked directly to RC. See Figure 2.1 for a reference of the commonly cited domains of EF.

Despite the disparate definitions of EF throughout the literature, other researchers of EF most frequently cited the model described by Miyake et al. (2000) according to Baggetta and Alexander's review (2016). In this seminal work, Miyake et al. (2000) proposed that EF is comprised of *updating*, *shifting* and *inhibition*, validated with a sample of 137 college-age students. *Shifting* is defined as both "the ability to engage and disengage appropriate task sets . . . and the ability to perform a new operation in the face of proactive interference or negative priming" (Miyake et al., 2000, p. 56). *Updating* is defined as the ability to "actively manipulate relevant information in working memory, rather than passively store information" (Miyake et al., 2000, p. 57). *Inhibition* is defined as the ability to "deliberately stop… a response that is relatively automatic" (Miyake et al., 2000, p. 58). Across all of the studies reviewed, the three most common components of EF were *working memory*, *updating*, and *shifting* (Baggetta & Alexander, 2016).

Though the discrepancies in the definition of EF within the literature may complicate research on the contribution of EF to RC, there is a need for more research in this direction. Jacob and Parkinson's recent meta-analysis (2015) provided a timely review of EF's contribution across the developmental span for both reading and math achievement, and defined EF across four domains: *response inhibition*, *attention control*, *attention shifting*, and *working memory*. Across constructs, measures, and age, Jacob and Parkinson (2015) found a moderate, unconditional association between EF and achievement. However, they also noted that there is only narrow evidence of a causal relationship between EF and reading due to the preponderance of correlational methods used to study EF. Jacob and Parkinson (2015) did not look specifically at reading, focus on a particular developmental stage, or investigate whether findings were consistent across populations, including students with disabilities and linguistically diverse learners.

Reading is important to study in the present synthesis because it has been shown to predict long-range social and scholastic outcomes (e.g., Kern & Friedman, 2008). Upper elementary is important to study because of the expanding literacy demands that arise during these grades. Specifically, the complexities of texts presented in upper elementary require greater lengths of sustained attention and coordination of skills than in earlier grades. These requisite skills overlap substantially with the definitions of EF presented above. At-risk populations are important to study because students who struggle academically are more likely to have limited opportunities after high school, withdraw from school prematurely, or become incarcerated (Reynolds, Temple, Roberson, & Mann, 2002). Thus, the present synthesis reviews the extant research base on the role of EF in RC for students in upper elementary school with particular attention to at-risk populations, such as ELs and students with disabilities.

Developmental and Linguistic Considerations

Upper Elementary. When studying EF, it is important to consider how the relative contribution of EF may change as students progress through school. In Huizing, Dolan, and van der Molen's (2006) study of students from seven to twenty-one years old, they found that some components of EF developed at different rates. Existing syntheses for contributors to RC in lower elementary school are numerous (e.g., Piasta & Wagner, 2010; Shanahan & Lonigan, 2010) and have found that early reading skills, such as phonological awareness and decoding, are consistent indicators of reading difficulty in lower elementary school. However, as students progress, the expectations for RC progress exponentially. In upper elementary school, there is growing recognition of "late-emerging poor readers," or, "late-emerging poor comprehenders," in both English monolingual (Catts et al., 2012; Compton et al., 2008; Leach et al., 2003) and Spanish-speaking EL (Kieffer, 2010) populations. Across studies, at least some percentage of the sample demonstrated adequate decoding skills, but fell below average RC outcomes and vice-a-versa. One plausible explanation is that these students were missed at an earlier screening, and thus they are not late-emerging, but simply late-identified. However, another explanation may be that these deficits only truly emerged in upper elementary school when students were faced with processing complex texts, a task that requires interplay of skills beyond decoding (Catts et al., 2012; Compton et al., 2008; Lipka, Lesaux, & Seigel, 2006). Further, screening for linguistic comprehension is not typically performed in isolation from reading measures, which may account for some students who are not identified as struggling in earlier screenings. Finally, there may be some students who demonstrate adequate decoding and linguistic comprehension skills, yet have difficulty coordinating the two, which may implicate EF. Therefore, investigating the contribution of EF to RC for students in upper elementary grades is particularly warranted.

Reading Difficulties. Students with reading difficulties face a myriad of risks beyond poor academic performance, such as withdrawing from school or being incarcerated, compared to peers without disabilities (Reynolds et al., 2002). Therefore, studies that identify potentially malleable factors related to reading outcomes are vital for the field. There is also evidence that students with reading difficulties demonstrate weaknesses in the areas of EF, such as working memory, inhibition, and planning (Locascio et al., 2010; Swanson, Howard, & Saez, 2006). An additional facet of the complexity of these intersecting constructs is that students with reading difficulties in upper elementary school are “likely a more heterogeneous group of poor readers compared to younger students” because of the wide variety of domains required for comprehension in upper grades (Speece et al., 2010, p. 259). For example, Locascio et al. (2010) compared EF skills in three subgroups of reading ability: students with word reading deficits, specific comprehension deficits, or their controls without deficits. The authors also found variations in the relative strengths in each EF domain present by subgroup, including: students with word reading difficulties were more likely to have deficits in working memory; students with specific comprehension showed weaknesses in the area of planning; and students without word reading or specific comprehension deficits had higher scores in all EF areas. Thus, it may be particularly important to consider students’ relative strengths and weaknesses in decoding and linguistic comprehension when investigating the contribution of EF to RC for students with reading difficulties. It may also be helpful to employ statistical models that allow for the direct and indirect effects of EF on RC to be explored, such as the direct effect of EF on RC and the potential indirect effects of EF on RC via linguistic comprehension and decoding.

Attention Disorders. The link between EF and RC may be particularly important for students who have attention disorders. This concern is reflected in the breadth of literature conducted to determine the congruence and independence of deficits in EF and the clinical diagnosis of attention deficit hyperactivity disorder (ADHD) (e.g., Biederman et al., 2004; Martel, Nikolas, & Nigg, 2007; Zelazo & Müller, 2002). Additionally, the guidelines for an ADHD diagnosis overlap with some of the domains presented in the EF literature. According to a recent review of 35 studies of achievement for students with ADHD, students with the predominantly inattentive subtype of ADHD actually demonstrated poorer educational achievement than the predominantly hyperactive subtype, although the former group is less often referred for a diagnosis (Van der Kolk, van Agthoven, Buitelaar, & Hakkaart-van Roijen, 2015). Rucklidge and Tannock (2002), found adolescent students with dual diagnoses of ADHD and reading disabilities presented greater deficits in executive function compared to students with ADHD alone. Willcutt, Doyle, Nigg, Faraone, and Pennington's synthesis of studies of EF and ADHD (2005) found that while all students with ADHD demonstrated deficits in EF, the relative severity and EF subcomponent of that deficit varied across studies. Therefore, while EF plays a role in the ADHD profile, it is not the only contributing variable. A review of the literature conducted by Sexton, Gelhorn, Bell and Classi (2012) found prevalence rates for the co-occurrence of ADHD and reading disabilities between 9% to 60%. In natural samples of students where ADHD and reading disabilities often present clinically as comorbid disorders, researchers face greater issues of statistical power in order to distinguish whether the contribution of EF to RC differs for students with ADHD, reading disabilities, or dual diagnoses.

English Learners. Finally, EF skills are likely to be important, and may even function differently, for ELs or students learning English in school while speaking another language at home. Learning a second or additional language requires an ongoing cognitive process of connecting what is known in the first language (L1) with what is known in the second language (L2). The cognitive demands of switching between languages, updating the lexicon in each language as new words are learned, ignoring or inhibiting false cognates, translating when necessary, and maintaining information across languages in working memory all appear on the surface to relate closely to EF's many domains. This distinction between students who speak more than one language, either at home or in school, compared to students who speak only one language has been termed the bilingual advantage and received much attention in the literature as of late, with sometimes conflicting results (Morton & Harper, 2007; Paap & Greenberg, 2013; Prior & MacWhinney, 2010; Qu, Low, Zhang, Li, & Zelazo, 2016; Yang, Hartanto, & Yang, 2017).

A number of studies have evidenced the bilingual advantage in EF measures. For example, Bialystok (1999) found that preschool bilingual students demonstrated greater inhibitory control than their monolingual peers while completing a dimensional card sorting task, which measures set-shifting. The task required students to sort cards based one feature, and then a second feature. The type of errors students make during each stage are recorded and analyzed. Bialystok and Martin (2004) replicated the sorting task of the Bialystok (1999) study and found similar results. In addition, Bialystok and Martin (2004) compared monolingual and bilingual students' inhibition performance on perceptual features versus semantic features of the sorting task and found that bilingual students outperformed monolingual students on only perceptual demand. Bialystok (2011) also demonstrated an advantage for 8-year-old bilingual students compared to monolingual students on a compound task requiring the synchronization of inhibition, working memory, and shifting. Carlson and Meltzoff (2008) compared a sample of native Spanish-English bilingual kindergarten students with English monolingual students to English monolingual students in a Spanish immersion program. Across the three groups, the study found that native bilingual students achieved the highest scores on measures of EF. Similarly, Bialystock, Barac, Blaye, and Poulin-Dubois (2010) found that bilingual students in early childhood outperformed monolingual students on multiple measures of EF, despite the monolingual students' relative strengths in vocabulary. Given the apparent importance of EF for bilingual students, further investigations are needed to connect performance on discrete tasks of EF to the prediction of RC, controlling for linguistic comprehension and decoding.

Present Study

The present research synthesis investigates the contribution of EF to RC, controlling for known reading-related skills such as decoding, fluency, and linguistic comprehension for students in upper elementary school. This synthesis also considers whether EF and RC's relationship differs for at-risk students, including those with reading difficulties, attention disorders, and ELs. This synthesis is needed because other syntheses to date have not controlled for linguistic comprehension and decoding when examining the relationship between EF and decoding. Additionally, no syntheses have been performed that isolate the effects of EF on RC for students in upper elementary school, which is a critical time for reading development and intervention before middle school. Finally, none of the EF syntheses to date have examined whether the effects of EF on RC differ for students with disabilities and ELs while controlling for linguistic comprehension and decoding.

EF is defined as a multi-component construct that includes the following cognitive skills: working memory, processing speed, response inhibition, task updating, planning, organization, cognitive flexibility, attention shifting, and self-regulation (Baggetta & Alexander, 2016; Blair & Razza, 2007; Cartwright, 2012; Jacob & Parkinson, 2015; Locascio et al., 2010; Miyake et al., 2000). In order to reflect EF's manifold definitions in the literature, the operationalization of EF in the present study was allowed to vary as defined by the original studies. Given that all preeminent EF definitions (see list above) include at least two domains, studies must have administered at least two EF assessments to be included in the present synthesis. RC is defined as the skillful use of fluent decoding and linguistic comprehension, as per the definitions set forth by Gough and Tunmer (1986) in the original articulation of the SVR. To be included in the present analysis, the studies must have concurrently administered: (a) at least one standardized, norm-referenced measure of RC, (b) at least one measure of decoding or fluency, and (c) at least one measure of linguistic comprehension. Additional criteria for inclusion in the present study are described in the methods section below.

This synthesis seeks to address the following research questions:

1. What is the unique contribution of EF to RC, controlling for decoding and linguistic comprehension, for students in upper elementary school?
2. Does the contribution of EF to RC vary for at-risk populations of students, including students with disabilities in reading and/or attention and English learners?

Method

The search for peer-reviewed articles to be included in this synthesis followed the process described by Cooper and Hedges (2009) in their chapter *Research Synthesis as Scientific Process.* Their description of the research synthesis' stages was used as a model in the crafting of this present synthesis. Namely, the chapter undertook the following described steps: problem formation, literature search, data evaluation, data analysis, interpretation of the results, and public presentation. A description of the decisions made at each stage in the current synthesis is detailed below.

Stages of Research Synthesis

Problem Formation. Research suggests that students in upper elementary school, and particularly students with reading difficulties, attention disorders, and diverse language backgrounds, may be at risk for difficulty in reading comprehension. Research also suggests that EF performance is correlated with RC outcomes and that EF may explain or contribute to RC performance above and beyond variables known to be associated with RC including decoding and linguistic comprehension. The present synthesis builds on reviews by Baggetta and Alexander (2016) and Jacob and Parkinson (2015), mentioned above, and distinguishes itself by focusing on upper elementary school as well as reading difficulties, attention disorders, and diverse language backgrounds. Additional distinctions are addressed further in the discussion.

Literature Search. The literature search was conducted in three stages. In the first stage, existing syntheses on reading and EF were reviewed to determine common search terms. Of these studies, the most relevant and commonly used key words were *reading comprehension* and *executive function*, which became the primary search terms for the present synthesis. The search was conducted using the WorldCat online database, which automatically searched numerous individual databases relevant to education and psychology research, such as Academic Search Premier, JSTOR, PubMed, ERIC, and ScienceDirect. Additional filter criteria were added to omit articles not yet published in peer-reviewed journals, those not accessible in English, and/or those not published since 2000, which is when Miyake et al. (2000) published their seminal article. The first, and broadest, stage of the literature search method generated approximately 650 articles. In the second stage, the articles were scanned for their applicability for further review, using information provided by title, journal, date, keywords, and abstract. Following the second stage, roughly 200 articles met the basic inclusion qualifications for the present study mentioned above. In the third, and most intensive, stage of the literature search, each of these articles was reviewed in detail and coded for whether they met the final inclusion criteria. These criteria are listed below:

1. Published in a peer-reviewed journal and accessible in English.
2. Participants were assessed in grade 4 or 5, also referred to as "upper elementary".
3. Dependent variables included one or more standardized, norm-referenced measures of RC.
4. Independent variables included two or more standardized measures of EF.
5. Control variables, or covariates, included one or more measures of fluent decoding and linguistic comprehension.
6. The research design followed procedures for determining concurrent validity, which is a type of regression or prediction equation in which two or more measures thought to be related are assessed at the same time. Concurrent validity contrasts from predictive validity, in which two or more measures are assessed at different time

points and the regression equation predicts from one time point to the next.

7. Quantitative analysis was conducted to determine the relation between the predictors of RC and the measures of EF, decoding, or linguistic comprehension for the participants.

After the three stages of increasingly specific criteria were applied as described in the literature review section above, only seven studies remained in the synthesis.

Data Evaluation. The data evaluation process was conducted concurrent to the final literature search stage, and similarly followed a three-step iterative coding process. A starting list of codes was generated from the literature in the first stage; open coding was applied to reflect additional information presented in the studies under review in the second stage; and the codes were collapsed to reflect the themes that emerged from the synthesis in the third stage.

Data Analysis and Interpretation. After the completed data evaluation, the following overarching categories, or level 1 codes, were used to organize the data analysis and interpretation of the synthesis: study features, theoretical constructs and measurement, and analytic findings. Descriptions of the sub-codes, or level 2 codes within each level 1 category are provided below.

Study features. Within the study features category at level 1, studies were coded at level 2 for their publication date, research design, age and grade range, populations studied, and geographical location and affiliation. For publication date, the year the study was first published in a peer reviewed journal was recorded. For study design, the term concurrent validity was used to indicate the study's design to measure the relationship between EF and RC. Studies that did not include controls for decoding, fluency, and linguistic comprehension were omitted from the final sample of studies. For age and grade range, the lowest grade or youngest age to the highest grade or oldest age of all of the participants was recorded. Because the present study is focused on upper elementary school, the number of participants in fourth and/or fifth grade in the study was recorded separately in addition to the total sample size. The studies were coded for student demographics to reflect whether the sample included: (a) typical or on-grade-level readers, (b) students with reading difficulty or below-grade-level readers, (c) students with attention disorders such as ADHD, and (d) ELs. The student demographic codes relied on the definitions presented in the original studies. A definition description for each code is provided in the results below. In order to evaluate whether there were geographical or affiliation biases present in the research, the studies were also coded for their location and institution, when described in the articles. See Table 2.1.

Theoretical constructs and measurement. The level 1 code of theoretical constructs and measurement included the following level 2 codes: dependent measures, independent measures, and covariates or controls. Each of these level 2 codes also received level 3 codes, which were developed using an open and iterative process to reflect the types of data found in the original studies. Refer to Table 2.2 for a complete summary of the original studies' theoretical constructs and measurements.

Reading comprehension. Reading comprehension outcomes served as the dependent measure in the present synthesis. Each study in this synthesis was coded for RC according to whether the measures assessed were multiple-choice format, cloze passage format, or other format. Refer to Table 2.2.

Executive function. EF served as the independent measure in the present synthesis. Each study was coded for EF according to whether the assessments addressed one or more of the following categories: working memory, processing speed, shifting, inhibition, planning, organization, and attention rating scales. Refer to Table 2.2.

Covariates. Covariates, or control variables, that studies used within the present synthesis were coded as well. These codes included measures of: decoding, fluency, linguistic comprehension, and other reading measures. Refer to Table 2.2.

Analytic findings. The procedures introduced by Cooper and Hedges (2009) were used to analyze the results of the studies. For each study, the dependent measures, independent measures, control measures, coefficient values (β) and significance level (p) were analyzed to determine the strength and direction of the relationship between EF and RC measures. Refer to Table 2.3. Additionally, the reliability and validity information for each measure, as reported by each study, were considered when evaluating the studies' methodological strength.

Excluded Studies

There were some studies that met most, but not all, of the criteria. For instance, Altemeier, Abbott, and Berninger's (2008) cohort sequential study with 241 students from first through sixth grade was removed because it did not include a measure of linguistic comprehension as a covariate for RC performance. This means that it is possible that the performance variation attributed to EF measures by the study was partially due to the variation in linguistic comprehension, and thus the study was not suitable for the current synthesis. However, this study is notable because the authors included a subsample of students with reading disabilities (in this case, dyslexia) and conducted additional analyses to determine whether there were significant differences in the relationship between EF and RC for these students compared to their typical peers. Additionally, this study appears to be the only one that has considered the influence of EF on writing. Locascio and colleagues (2010) conducted a study with 86 students between 10–14 years old that contributes to our understanding of the link between EF and RC. This study is important because the authors disaggregated the results by reading ability, including a typical reader group, a word recognition deficit group, and a specific RC deficit group. However, the statistical methods that the authors used did not meet the current study's inclusion criteria. Locascio et al. (2010) conducted ANCOVAs to examine group differences by EF domain, but did not predict RC status based on the EF measures while controlling for decoding, fluency, and linguistic comprehension.

Results

All of the seven studies under review found one or more EF measures as significantly and positively related to RC, and furthermore, one or more EF measures explained a unique portion of the variance in RC performance, controlling for decoding, linguistic comprehension, or both. The results of the synthesis are organized under three main categories: (1) study features, (2) theoretical constructs and measurement, and (3) analytic findings.

Study Features

Design. As dictated by the synthesis parameters, all of the studies investigated EF and RC's relationship concurrently (i.e., at a single assessment time point). All of the measures were administered within a one-month window.

Age/Grade. Three of the studies examined the relationship between EF and RC in fourth and/or fifth grade specifically. Gerst et al. (2015) studied students in fourth and fifth grade; Kieffer et al. (2013) sampled students in fourth grade; and Nouwens et al. (2016)

included students in fifth grade. The remaining four studies included students in their analytic sample outside of the upper elementary age range. These studies were incorporated in the present analysis because they accounted for differences in participants' age by including age as a control variable and/or using age-based standard scores in their analyses.

Sample Size. The sample size in each study ranged from 56 to 483 students. A-priori power analysis through G*Power suggested that the minimum number of participants needed to detect a small effect is 543 participants, assuming a power of .8 with 3 predictors and a probability level of .05 (Faul, Erdfelder, Buchner, & Lang, 2009). Under the same condition, the minimum number of participants to detect a medium effect is 76, and to detect a large effect is 36. Given the number of variables in the analytical models of each study and using the guidelines for multiple regression outlined by Cohen, Cohen, West, and Aiken (2003), it appears that all of the studies were underpowered to detect a small effect; two of the studies were underpowered to detect a medium effect (Cutting et al., 2009; Sesma et al., 2009); and all of the studies were adequately powered to detect a large effect.

Subgroups.

Reading difficulties. All of the studies in the synthesis included typical readers and students with reading difficulties in their samples, but the measurement and categorization of reading difficulty status varied widely across studies and the authors did not always describe them. For example, Nouwens et al. (2016) included students with dyslexia (n=14), ADHD (n=12), Asperger Syndrome (n=1), and comorbid disorders of ADHD, dyslexia, and dyspraxia (n=2), although the authors did not specify the criteria for any of those diagnoses. Cutting et al. (2009) categorized student participants as having either a general reading disability (GRD) or specific-RC disability (S-RCD) based on researchers' administered assessment results. In this study, restrictions in performance on multiple measures defined the criteria for GRD, including average scores below the 25th percentile on separate measures of word identification, decoding, and reading comprehension (word identification, word attack, and passage comprehension, respectively, from the Woodcock Reading Mastery Test - Revised Normative Update [WRMT-R/NU] and/or comprehension from the Gray Oral Reading Test - Fourth Edition [GORT-4]) (Cutting et al., 2009). The S-RCD criteria was a score “at or above the 40th percentile on the basic reading composite despite scoring at or below the 25th percentile on one or more of two measures assessing RC (GORT-4) Comprehension and WRMT-R/NU Passage Comprehension” (Cutting et al., 2009, p. 39). Sesma et al. (2009) also included students with different reading profiles, which they grouped into those with word reading deficits (WRD) and RC deficits (RCD). Students were assigned to the WRD category if they received a score “below the 25th percentile on a single word reading measure (Word Reading from the Wechsler Individual Achievement Test – Second Edition [WIAT-II])” and assigned to the RCD category if they received a score “below the 25th percentile on two of three measures of RC (Reading Comprehension from the WIAT-II, Comprehension from the GORT-4, Passage Comprehension from the WRMT-R)” (Sesma et al., 2009, p. 235).

Gerst et al. (2015) included 5 students currently receiving Special Education services (about 5% of their total sample) and 8 students in the Gifted and Talented Program (about 9% of their total sample) in their sample of 93 students, based on school records of supplemental services enrollment. Christopher et al. (2012) included 128 students (about 27% of the total sample) with a school-reported reading disability, 38 of which were comorbid with an ADHD diagnosis. Jacobson et al. (2016) considered participants “likely to have a reading disability” if they reported a “history of poor reading skills… falling below expected level for age or grade, and/or provision of special services in the area of reading” (p. 3). These students were contrasted to the control group, who were defined as performing “at or above current expectations for grade and performance … [and] above the 40th percentile on standardized school or clinical testing” (Jacobson et al., 2016, p. 3). However, Jacobson did not describe the proportion of their 761-student sample that met the criteria for reading disability or control status, and those categories were not used to disaggregate the results. Gerst et al. (2015) included special education and gifted and talented status as controls in their evaluation, but did not disaggregate the results based on those categories. Nouwens et al. (2016) conducted separate regression analyses using disability category as a between-subjects factor (the results of that analysis are presented in the analytic results section). None of the other six studies compared whether the predictive relationship between EF and RC under the full model differed for students with reading difficulties, although most of the studies reported group means by each measure.

Attention disorders. Five of the studies specified that students with attention disorders were included in their samples, although each study’s definition of an attention disorder differed. For example, Christopher et al. (2012) and Nouwens et al. (2016) considered students in their sample to have ADHD based on school reports of disability status, but the school criteria for that diagnosis were not reported. In the sample from Christopher et al. (2012), 38 out of the 93 students with ADHD had a comorbid reading disability diagnosis. In contrast, Sesma et al. (2009) used parent reports of inattention from the Behavior Assessment System for Children (BASC) Attention Scale (Reynolds & Kampus, 2004) to identify students as having an attention disorder in their final model. Only one study (Christopher et al., 2012) explored separate Structural Equation Models (SEM) with and without students with attention disorders.

English learners. Two studies included significant numbers of ELs in their sample. Gerst et al. (2015) reported that 32% of their sample included ELs defined as “limited English language proficiency” according to school records (p. 7). Kieffer et al. (2013) reported 67% of their sample “spoke a language other than English at home” (p. 337). Nouwens et al. (2016), in their study of Dutch children, included a small number (less than 3%) of language minority learners. Curiously, although Jacobson et al. (2016) purposefully sampled students from minority backgrounds (African American and/or Hispanic/Latino students), they excluded students “who indicated Spanish as their first and/or primary language and/or who were receiving English language support” from their sample (p. 3). Cutting et al. (2009) purposefully excluded “non-native English speakers” (p. 39). None of the studies included EL status as a covariate in their prediction models or disaggregated the results by English learners, although Jacobson et al. (2016) acknowledged that other studies found a “bilingual advantage” for students who speak another language on measures of EF in their discussion (p. 9). The focus of the present synthesis was constrained to studies that measured English language proficiency for students who speak another language than English at home.

Location and Affiliation. Five of the studies were conducted from multiple regions across the United States and one study included a sub-sample in Canada and Puerto Rico; one study was conducted in the Netherlands; and one study did not specify the location. Three of the studies specified their affiliation with independent research centers and/or government funded research projects.

Theoretical Constructs and Measurement

Reading Comprehension. All of the studies included at least one standardized measure of RC and controlled for either decoding or fluency in their predictive model, which were criteria for inclusion in the present synthesis. Across all seven studies, a total of nine different RC measures were used, showing that researchers chose a diversity of RC measures to represent this construct. The three most commonly used types of measures were reading inventories (such as the Qualitative Reading Inventory – Third Edition [QRI-3]); cloze sentence tasks (such as the GORT-3 or GORT-4 and the Scholastic Reading Inventory-Second Edition [SRI-2]); and silent passage reading with multiple choice questions (such as the Gates MacGinitie Reading Test, Fourth Edition: Passage Comprehension subtest [GMRT-4: PC]). Reading inventories require students to read a passage aloud and then answer a mix of factual and inferential questions. Cloze sentence assessments require students to provide a missing word in a sentence using sentence-level inferencing, including both the word's literal meaning in the sentence (i.e., semantics) as well as the missing word's function and form given the sentence structure (i.e., syntax and morphology). Multiple choice questions following passage reading, such as in the GMRT-4: PC subtest require students to activate a high level of RC strategies and inferencing based on the literal and implied information provided in the narrative and expository passages. In this measure, students read silently to themselves for the entire assessment. Two of the studies included more than one RC measure (Christopher et al., 2012; Cutting et al., 2009), but only Cutting et al. (2009) investigated whether the relationship between EF and RC differed depending on the type of RC assessments administered. Christopher and colleagues' (2012) study was the only one to construct a latent variable of RC. As reading comprehension is understood to be a multifaceted construct, using latent RC variables in statistical analysis yields a more robust representation of RC than a single measure. Support for latent constructs of RC in education research is well established (Fletcher, 2006; Mehta, Foorman, Branum-Martin, & Taylor, 2005; Vellutino, Tunmer, Jaccard, & Chen, 2007). Refer to Table 2.2 and Table 2.3.

Decoding and fluency. All of the studies included at least one measure of fluent decoding, which was a criterion for inclusion in the synthesis. Each study's decoding and fluency measures fell into three categories: untimed word reading (n = 6 studies), timed word reading (n = 5 studies), and timed passage reading (n = 3 studies). The most commonly used, untimed word reading measures were the WIAT-II Word Reading subtest, Woodcock Reading Mastery Test -Revised (WRMT-R) Word Attack subtest, the Comprehensive Test of Phonological Processing (CTOPP) Elision subtest, and the Woodcock Johnson - Third Edition (WJ-III) Letter Word Identification (LWID) subtest. The most commonly used timed word reading measures were the Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests of the Test of Word Reading Efficiency (TOWRE). The most commonly used timed passage reading measures were the GORT-3: Rate and the GORT-4: Fluency subtests.

Linguistic Comprehension. As per the criteria for inclusion in the present study, all of the studies incorporated a measure of linguistic comprehension. Measures of vocabulary and oral comprehension were the most commonly used types of linguistic comprehension. For vocabulary, the only measure used was the Peabody Picture Vocabulary Test – Third Edition (PPVT-III; n = 4 studies), although an adapted version was used in the Nouwens et al. (2016) study because the students were Dutch speakers. The PPVT-III measures expressive vocabulary knowledge. For linguistic comprehension, the Woodcock Johnson-

Third Edition Oral Comprehension (WJ-III OC) subtest was the most commonly used measure (n = 2 studies), in which students listen to short passages and then provide the last word. Other measures of linguistic comprehension required students to listen to sentences and combine them (e.g., Test of Language Development-3: Sentence Combining), or evaluate the meaning of sentences presented orally (e.g., Test of Language Competence-Expanded Edition: Ambiguous Sentences and Making Inferences).

Other Measures. In addition to the RC, fluent decoding, and linguistic comprehension measures, three studies also used spelling and generalized intelligence measures. Christopher et al. (2012) included a measure of spelling as a control. Kieffer et al. (2013) used measures of phonological awareness, working memory, and processing speed as control variables. Cutting et al. (2009) and Nouwens et al. (2016) included generalized intelligence quotient (IQ) measures as controls.

Executive Function. Based on the extant literature, EF subcomponents were defined within eight domains: working memory, processing speed, shifting, updating, inhibition, planning, organization, and self-regulation. However, the studies reviewed did not directly assess updating, self-regulation, or organization, so they were omitted from the final analysis. Across all seven studies, the most commonly assessed domains of EF were working memory, shifting, and inhibition. Only one study (Christopher et al., 2012) constructed latent variables of EF. None of the studies assessed all eight domains attributed to EF and domains of EF measured were not consistent across studies, exhibiting the still-existent disagreement in the field over what constitutes core EF skills.

Working memory. The most commonly used measure of working memory was the Digit Span subtest of the WISC-R/WISC-III/or WISC-IV, although the studies varied in their use of either individual subtest scores or composite scores based on multiple subtests in their analysis. The second most common measure of working memory was the Freedom from Distractibility Index (FDI) from the WISC-III/WISC-IV. The FDI is a composite measure of the Arithmetic and Digit Span subtests. In each of these tests, participants are orally given sequences of digits or letters and required to recall and repeat the information either forward or backward.

Processing speed. Processing speed was assessed in three out of the seven studies, but different measures were used in each. These measures included the Colorado Perceptual Speed (CPS) Test 1 and 2, the Identical Pictures subtest of the Educational Testing Service (ETS) measure, the Processing Speed Index of the WISC-IV, and the Visual Matching subtest of the WJ-III. Of these measures, only the Processing Speed Index of the WISC-IV was a significant contributor to RC (Jacobson et al., 2016). The Processing Speed Index is a composite of the Coding and Symbol Search subtests. The Coding subtest requires participants to decipher symbols using a code. The Symbol Search subtest requires participants to identify whether a target symbol occurs in a given set.

Shifting. Shifting was assessed in five studies and, in addition to shifting, was referred to as *naming speed*, *shifting/switching*, *attention shifting*, and *cognitive flexibility*. The most commonly used measure of shifting was the D-KEFS: Trail Making Test (TMT) Number-Letter Switching condition (n = 2 studies). In the TMT Number-Letter Switching subtest, participants are given a scrambled array of letters and numbers from A to L and 1 to 16 and instructed to draw lines to sequence the letters and numbers from 1-A, 2-B, 3-C and so on. Other measures included variations of the Rapid Automatized Naming test and the 64-card version of the Wisconsin Card Sorting Test. All of these measures required the student to shift or switch between recalled information after being given a picture, letter, or number prompt.

Inhibition. Inhibition was measured in five studies and additionally referred to as *inhibitory control*. The most commonly used measure of inhibition was the DKEFS: Color Word Interference Test (CWIT) Inhibition condition, which was used in two studies. In the DKEFS CWIT Inhibition subtest, participants are provided with written color words that are printed in contrasting colors and required to name the ink color not the color word. Other measures included the Gordon Diagnostic System Continuous Performance Test Vigilance and Distractibility subtests (GDS CPT: V & D); the Stop-Signal Reaction Time (SSRT) test; and a modified version of a Number-Quantity Stroop test. In one study (Cutting et al., 2009), inhibition was assessed through a combined shifting and planning measure (i.e., Tower of London), in this case referred to as a *spatial planning, rule learning, and inhibition*. All of these measures required the student to restrain a trained response when presented with interfering information under timed conditions.

Planning. Planning was measured in five studies and additionally referred to as *problem solving*; *planning, organization, and monitoring*; and *spatial planning, rule learning, and inhibition*. The most commonly used measure was the Tower of London or an adapted version of this task (I-3 Tower Task and D-KEFS: Tower). In the Tower subtests, participants are given pegs and discs and required to replicate a tower based on a model using the fewest moves possible. Other measures included the Elithorn Perceptual Maze test and the WISC-IV: Matrix Reasoning test. All of these measures required the student to move a group of objects into a given shape or formation while following a list of rules about the objects' placement under timed conditions.

Attention rating. Two studies (Cutting et al., 2009; Sesma et al., 2009) included rating scales of attention, which parents or teachers completed about students' general behavior in the classroom and at home as measures of students' EF. Both studies used the measures as control variables that were a proxy for ADHD status. Methodological and theoretical issues related to the use of these measures are addressed further in the discussion.

Study Findings

Prediction of Reading Comprehension. Across studies, at least one, but not all EF measures were significant predictors of RC, controlling for decoding, fluency, and linguistic comprehension. These EF domains were working memory, processing speed, shifting, inhibition, planning, and attention rating. The most frequent, significant predictor across studies was working memory, which was significant in five studies (Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016; Nouwens et al., 2016; Sesma et al., 2009), but non-significant in two studies (Cutting et al., 2009; Kieffer et al., 2013). Planning was found as significant in three studies (Cutting et al., 2009; Jacobson et al., 2016; Sesma et al. 2009), but non-significant in two studies (Gerst et al., 2015; Nouwens et al., 2016). Shifting was found as significant in two studies (Kieffer et al., 2013; Nouwens et al., 2016), but non-significant in three studies (Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016). Inhibition, which was measured in five out of seven studies, was significant in two studies (Cutting et al., 2009; Kieffer et al., 2013). Attention rating scales were used in two studies, but attention rating was significant in only one study (Sesma et al, 2009). A summary of the regression analyses is provided in Table 2.3, and individual study findings are described below.

Christopher et al. (2012) found that working memory was a significant RC predictor ($p < .01$) and processing speed was 'marginally' predictive ($p = .08$), controlling for decoding. However, inhibition and shifting were not significant RC predictors. Cutting et al. (2009) found that working memory and a combined planning and inhibition measure were insignificant predictors of RC measured by WRMT-R/NU Passage Comprehension and controlling for decoding, fluency, and oral language (b = -0.07, p = .57; b = -0.14, p = .24). However, using a different RC measure (GORT-4: C) as the outcome and the same control variables, the combined planning and organization measure was significant, (b = -0.28, p = .05) but the working memory measure was not (b = -0.01, p = .93). Gerst et al. (2015) found that two measures of working memory (b = .24, p = .01; b = -0.25, p = .002) were significant RC predictors controlling for age, oral comprehension, decoding, and gifted and talented status. However, measures of shifting, inhibition, and planning were not significant (b =.06, p = .39; b = -0.07, p = .41; b = .07, p = .37). Jacobson et al. (2016) found that processing speed and a composite EF measure, including problem solving, working memory, and switching, were both significant RC contributors, controlling for reading-related language skills ($p < .001$; $p < .001$), and explained an additional 1.5% and 4.5% of the variance, respectively. However, shifting was not a significant predictor on its own within the EF composite (b = .031, p = .196). Kieffer et al. (2013) found that both shifting and inhibition were significant RC predictors, controlling for decoding, language comprehension, working memory, processing speed, and phonological awareness (b = .16, $p < .05$; b = .19, $p < .05$). However, measures of working memory and processing speed were not significant predictors in this model. Nouwens et al. (2016) found that shifting and working memory were significant predictors of performance on a listening span task. However, storage, inhibition, and planning were not significant predictors of RC (b = -.019, p = .834; b = .145, p = .132; b = -.091, p = .255). Sesma et al. (2009) found that attention rating (b = -.13, p

= .001), working memory (*b* = .17, *p* = .032), and planning (*b* = -.25, *p* = .013) were all significant RC predictors, controlling for attention rating, decoding, fluency, and vocabulary.

Subgroup Differences. Although the present synthesis was intended to investigate EF's contribution to RC for at-risk students, the majority of the reviewed studies failed to conduct the necessary statistical analyses to support any broad conclusions about the differential relationship between EF and RC for students with reading and/or attention disorders or ELs.

Reading difficulties. Cutting et al. (2009) and Nouwens et al. (2016) examined group differences for students with reading difficulties and their peers on RC and EF measures. Cutting et al. (2009) found that students with reading comprehension specific difficulties performed significantly worse on a combined planning and inhibition measure (TOL) compared to students with general reading difficulty ($p = 0.04$) and typical readers ($p < 0.001$), although both groups performed similarly to one another, though significantly worse than typical readers ($p = 0.004$), on a different planning measure (Elithorn Perceptual Maze Test). All groups (S-RCD, GRD, control) performed similarly on a measure of working memory (WISC-III/IV DSB). However, Cutting et al. (2009) did not include reading difficulty status in a full hierarchical regression model, so this study does not provide evidence about how reading difficulty status and EF interact with RC when fluent decoding and linguistic comprehension are in the model. Nouwens et al. (2016), comparing group means on individual measures, found that students with dyslexia scored lower than controls on the RC and word recognition measures, but did not differ on vocabulary, working memory, and EF tasks. Furthermore, Nouwens et al. (2016) entered reading disability status as an interaction term in the regression analysis, but found that it was not significant, thus suggesting that these results were not differentiable for students with and without disabilities.

Attention disorders. Four studies included students with attention disorders, but only one investigated whether the relationship between EF and RC differed for this group of students, controlling for decoding, fluency, and linguistic comprehension. Sesma et al. (2009) used the attention rating score as a predictor in their full model and found that it was significant (b = -.13, p = .001). The authors indicated that the measure, which has been validated for the diagnosis of ADHD, was designed to "assess the extent to which a child is easily distracted and has difficulty sustaining concentration" (Sesma et al., 2009, p. 236). However, the rating scale only identified 17 students as having attention difficulties, even though 27 of the students were identified as having ADHD based on a previous diagnosis from a full-scale assessment. This inconsistency calls into question the generalizability of the results to students with ADHD as defined by different measures. Cutting et al. (2009) also used an attention rating, but did not include the measure in the full hierarchical regression analysis. The authors indicated that the attention rating was used to "establish levels of ADHD symptomatology" (Cutting et al., 2009, p. 39), but this proxy ADHD status was not used to disaggregate the results. Nouwens et al. (2016) looked at group differences for students with ADHD compared to their peers and found they scored lower than controls on the RC measure, but did not differ on word reading, vocabulary, working memory, and EF tasks. The study also found that ADHD status was not significant when entered as an interaction term into the full regression model. Again, this null result regarding the interaction between ADHD status and the contribution of EF to RC should not be over-interpreted due to the conducted study's limitations in power. Future research should continue to examine whether EF and RC's relationship differs for students with attention disorders.

English learners. Three of the studies (Gerst et al., 2015; Kieffer et al., 2013; Nouwens et al., 2016) included ELs in their samples, but none conducted additional analyses, either comparing means between EL status and individual measures, or within the regression analysis, to test whether EF and RC's relationship differed for this population.

Discussion

This study found evidence that some, but not all, EF measures are significant RC predictors, controlling for decoding, fluency, and linguistic comprehension. The current analysis provides a needed addition to the literature because it reviewed an extensive body of work conducted since 2000 to investigate EF's contribution to RC for underserved populations, with a focus on upper elementary grades, a critical point in academic development. Findings from this synthesis can be used to inform future directions in research.

Issues of Definition and Measurement

Several trends are notable across studies. EF's definition and operationalization hinders an unequivocal evaluation of the primary research question. Both Baggetta and Alexander (2016) and Jacob and Parkinson (2015) concluded that the current definition of EF is obscure. While all studies in the present analysis suggested that EF is comprised of multiple and related but separable skills, working memory was the only domain consistently assessed across all studies. Therefore, it is difficult to ascertain the relative contribution of each domain or draw conclusions about which ones might be collapsed into a more parsimonious model. Based on the current analysis, there is evidence that processing speed, shifting, inhibition, planning, and attention rating may offer unique contributions to RC outcomes for a variety of learners. Directions for future research include conducting a measurement study to determine how to best define and measure EF, which will likely require multiple measures for different domains of EF that could be used to form a latent construct or a new measure that includes items that assess the varied components under the EF umbrella.

Significance of Working Memory. Working memory was the only measure administered in all studies, and also the most frequently significant predictor of RC across all EF domains. This finding is in agreement with previous studies that have found that measures of working memory are likely to predict RC performance (Carretti et al., 2009; St Clair-Thompson & Gathercole, 2006). Interestingly, Baggetta and Alexander (2016) found that working memory was the second-most commonly cited EF domain after response inhibition. Theoretically, working memory's contribution to RC seems logical because even at the sentential level, the reader must hold information presented in the beginning of the sentence to make inferences about its relation to the end. At the same time, the reader must also add relevant information and ignore irrelevant information, which are both assumed components of working memory.

Significance of Speed. While EF performance may explain some of the underlying cognitive skills that students possess that are also drawn upon in reading tasks, EF is not expected to explain all of the variance in RC. With the exception of the raw score as a planning measure of the number of moves performed in order to reach the goal state or shape, the majority of EF measures included an aspect of speed because raw scores were reported as total time to completion. On one hand, the timed nature of almost all of the tests allow the researcher to compare quickly across measures. On the other hand, fluency or speed of performance is only one aspect of skilled reading. RC requires the successful product of decoding and linguistic comprehension, which requires task-situated critical thinking in order to parse meaning from text. In that sense, discrete, process-oriented EF measures can only explain part of the variance in RC performance, as phonetic knowledge and vocabulary are central to RC. The findings from this synthesis present an argument that EF may aid in the prediction of RC, but would not surpass the contributions of decoding and linguistic comprehension.

Literacy Overlap. Similarly, many of the EF measures required basic literacy and/or numeracy skills, which limits the ability of the EF measures to tap exclusively into EF skills. For example, students must remember increasingly long sequences of letters, digits, words, or even sentences in many of the working memory tasks. Measures of shifting, such as the D-KEFS Number-Letter Switching subtest, requires students to draw upon their knowledge of sequential order for alphabetic and numeric units (i.e., A, B, C, 1, 2, 3, etc.). Due to the potential overlap between the EF measures and foundational literacy and numeracy skills, it is plausible that some of the variance in RC explained by EF is due to confounds in the way EF is measured (e.g., remembering sequences of letters may overlap with decoding ability, which is a known component of RC, or being able to name letters and sounds, which requires language skills may overlap with linguistic comprehension, another known component of RC). This suggests that future EF and RC studies should choose EF tasks that require minimal literacy and numeracy knowledge in order to adequately test the individual contribution of EF to RC, controlling for decoding and linguistic comprehension.

RC was the primary, dependent variable in the present study. However, in the present analysis, only one study (Cutting et al., 2009) investigated whether EF and RC's relationship differed according to the type of RC measure used. There are several advantages of including multiple RC measures. For example, multiple measures could be used to explore latent constructs, reduce error in the analyses, or to examine how the relationship between EF and RC depends on the RC measures under consideration. Similarly, a single measure was used to assess almost all of the EF domains, which again limits the ability to generalize beyond that individual measure, and also increases the amount of error associated with using only one test as a predictor.

Limitations of Statistical Inference

The studies reviewed in the synthesis were methodologically sound, on the whole, but there were some limitations that should be discussed regarding the use of non-standardized measures and limits in statistical power. Several of the studies used researcher-altered EF measures. For instance, Christopher et al. (2012) included an adapted measure of working memory and shifting, and Sesma et al. (2009) included a scale from an extended teacher-rating questionnaire. In both cases, however, the authors did not include reliability data for those adapted measures. For the rating scales in particular, which are not direct assessments, there are inherent issues of reliability and validity. Rating scales have different types of error associated with them than direct observational assessment. Moreover, the questions regarding EF may have been intended to aid in the diagnosis of ADHD, not executive dysfunction.

Additionally, most of the studies did not include a large enough sample size to detect a small to moderate effect if it had been present. Some of the studies also included potentially extraneous controls in the model, which may have reduced the model's power to detect a small effect for the EF variables if they were present. For example, Cutting et al. (2009) included IQ as a covariate, and Gerst et al. (2015) included age and gifted/talented status as covariates. This presents an issue in the primary research question's interpretation, particularly if multiple EF measures were investigated in the same study. For example, Cutting et al. (2009) found a significant effect for a planning and organization measure, but not for a working memory measure. In this case, it is possible that there may be a small effect for working memory, but it could not be detected due to the sample size (n = 56). As described above, across all of the studies, each EF domain was found to be significant in one or more studies, but was also found to be non-significant in one or more studies.

All of the studies reviewed were based on correlational designs, whether they utilized structural equation modeling, hierarchical linear regression, path analysis, or multivariate analysis. In each of these models, the estimates obtained about the relationship between the independent and dependent variables are unidirectional. Christopher et al. (2012) appropriately cautions, "these loadings should not be interpreted as suggesting that a change in the independent variable causes change in the dependent variable" (p. 479). For instance, although this synthesis found some evidence that higher EF performance was associated with higher RC scores, it cannot be concluded that the higher EF performance caused the higher performance in RC. Future research should employ experimental studies to determine whether an EF-targeted intervention is associated with higher RC performance than a non-EF condition or whether EF moderates the effects of intervention on RC. These experimental studies would have actionable results for educational interventions because the design would allow researchers to draw more clear conclusions about the causality of the relationship between EF and RC performance. This is particularly important for students at-risk for reading difficulty, and who may stand to benefit the most from intervention on a potentially malleable factor that directly influences RC performance.

Limitations in Populations Studied

In their thorough review of the research on EF's contribution to reading and math achievement, Jacob and Parkinson (2015) concluded that "more studies that include strong sets of controls for child background characteristics and especially that include measures of the various subcomponents of EF in the same regression are needed so that the relative impact of each can be explored" (p. 542). This limitation still applies in the current synthesis' reviewed studies. Curiously, though, none of Jacob and Parkinson's (2015) reviews of the seven studies overlap with the studies in this synthesis. This may be due to the differences in Jacob and Parkinson's (2015) established search criteria, which included only correlational associations between one (or more) EF measure and either math or reading achievement generally. Whereas in the present synthesis, the inclusion criteria required at least one additional measure of decoding or linguistic comprehension as a covariate. Additionally, Jacob and Parkinson (2015) specifically excluded samples that only included students with disabilities. The inclusion criteria for the present study was more restrictive in the measurement of RC, requiring concurrent administration of control variables known to influence reading outcomes (i.e., fluency, decoding, and linguistic comprehension). The current study was also broader in the definition of EF, requiring two or more assessed domains. Finally, the current study incorporated diverse populations of students, including ELs, students with disabilities, and those with attention disorders. Nevertheless, only three of the studies in the present review included separate analyses to determine if EF and RC's relationship differed for students with reading difficulties or attention difficulties, and none of the studies investigated whether language status was significant. Given the findings from the present synthesis that multiple domains of EF contribute to RC performance for typically developing students, it is important to consider whether and how that relationship varies for learners with a diversity of backgrounds and abilities. Future research should look specifically at ELs, students with reading difficulties, and those with attention difficulties to determine how EF relates to RC.

This synthesis has implications for educational research with regard to how EF may contribute to successful RC. According to the studies reviewed, multiple EF domains, including working memory, processing speed, shifting, inhibition, planning, and attention, were found to be significant RC predictors when controlling for decoding, fluency, and linguistic comprehension. Future exploratory studies should seek to employ latent constructs of EF, RC, and linguistic comprehension to address the research base on EF. Additionally, studies should consider advanced statistical models, such as mediation, in order to test the direct and indirect effects of EF on RC with linguistic comprehension and decoding in the same model. Future research is urgently needed on EF and RC's relationship among diverse student populations at-risk for reading difficulties, including linguistically diverse learners and students with disabilities. Finally, future research should consider whether EF moderates the effects of a reading intervention on students latent RC among students from varying linguistic backgrounds.

Chapter 3: The Contribution of Latent EF to Latent RC via Decoding and Linguistic Comprehension for Linguistically Diverse Learners in Fourth Grade

Introduction

EF is receiving increased attention in educational research. Recently, Baggetta and Alexander (2016) reviewed 106 empirical studies related to EF. While evidence converged in some areas (e.g., that EF is multidimensional), different definitions and EF measures rendered interpretation across studies problematic. Specifically, depending on EF's definition and measures, different studies show varying relationships between EF and RC (e.g., Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016). Using information from multiple EF measures in a latent construct and advanced methods of statistical analysis (e.g. structural equation modeling) may help measure the underlying, rather than observed, construct of EF and its relationship to RC and reduce errors thereof. This manuscript chapter describes an empirical analysis of data from a linguistically-diverse sample of fourth grade elementary school students to examine the direct and mediating effects of latent EF on latent RC, controlling for decoding and latent linguistic comprehension.

The bulk of EF and RC research outcomes has been conducted with English monolingual students (e.g., Christopher et al., 2012, Cutting et al., 2009, Locascio et al., 2010). When studies are conducted with samples of students with diversity in language background and ability, language or disability status are often not used to disaggregate the results (e.g., Gerst et al., 2015). Some studies that have compared monolingual and bilingual children or adults have found evidence of a 'bilingual advantage' in EF (e.g., Bialystok, 2015), but the results have not been replicated with younger, emerging bilingual students. Additional research on whether English proficiency moderates EF's effect on RC is needed.

The objective of this study is to examine whether a latent construct of EF, formed by measures of working memory, shifting and inhibition, contributes a unique, direct effect on RC performance, over and above the contribution of linguistic comprehension and decoding. In addition, this study investigated whether school-defined English language proficiency and disability status moderated the relationship between the aforementioned variables.

Theoretical Background

Reading Comprehension. Linguistic comprehension and fluent decoding are known RC predictors or contributors. Gough and Tunmer (1986) first codified the relationship between decoding and linguistic comprehension in the Simple View Reading (SVR). Students must be able to read, or decode, texts at an adequate rate to allow them to make sense of the passage (LaBerge & Samuels, 1974; Pikulski & Chard, 2005). At the same time, students must be able to make meaning at the word level (semantics, morphology) and at the sentence level (syntax) in order to comprehend academic texts (Nation & Snowling, 2000; Ouellette & Beers, 2010; Proctor et al., 2012; Tumner & Chapman, 2012). The empirical validity of the unique contribution of both decoding and linguistic comprehension to RC has been well documented (Carver, 1998; Catts et al., 2006; Kendeou et al., 2009; Verhoeven & Van Leeuwe, 2008).

The SVR model has informed intervention research that focuses on supporting decoding, linguistic comprehension, or both for students in need of support. Decoding and linguistic comprehension have been found as malleable factors through intervention research (e.g., Rashotte, MacPhee, & Torgesen, 2001; Case et al., 2014; Shaywitz et al., 2004; Torgesen et al., 2001). However, interventions targeting both decoding and linguistic comprehension do not always lead to significant RC improvements (e.g., Greulich et al., 2014; Ritchey et al., 2012), leaving researchers to wonder what other cognitive skills should be addressed in intervention research. Additionally, across intervention research for students with reading difficulties, teacher and/or observer ratings of student attention and behavior have contributed small, but significant variance to students' reading outcomes (Al Otaiba & Fuchs, 2006; Greulich et al., 2014; Torgesen et al., 2001). Given that decoding and linguistic comprehension's relative influence on RC is in flux as students begin to coordinate longer and more complex texts in upper elementary school (Catts et al., 2006; Rasinski et al., 2005), investigating the EF's role at this critical juncture is important.

Executive Function. A constellation of studies found varying EF components associated with RC outcomes for students of varying ages (e.g., Alloway, Banner, & Smith, 2010; Blair & Razza, 2007; Cartwright, 2012; Locascio et al., 2010; Yeniad et al., 2013). In most studies, EF was comprised of at least three components, including working memory, shifting, and inhibition, although research has not coalesced around a precise EF definition. Baggetta and Alexander's (2016) synthesis reported the ten most frequently cited EF components: inhibition/inhibitory control, working memory, shifting, updating, cognitive flexibility, planning, switching, attention, emotional control/regulation, and fluency. In a synthesis conducted on EF's contribution to RC in upper elementary students (see Chapter 2), working memory, inhibition, and shifting were most frequently found as predictive of RC for students in upper elementary school.

Working memory. Working memory is perhaps the most widely studied construct out of all EF dimensions, although the model of working memory itself has been updated several times since the earliest model proposed by Baddeley and Hitch (1974; see Cowan, 1999; Engle, 2001). Variations on digit span or sentence span tasks are typically used to measure working memory (see Seigel & Ryan, 1989; Wechsler, 2003). In these tasks, students are given increasingly long sequences of information to recall. In some versions of the task students are asked to transform the sequences backwards, which requires manipulation of the information in addition to its storage. In a synthesis I conducted (see Chapter 2 of this manuscript), working memory was the most common EF subcomponent to contribute significant unique variance to the prediction of RC (see Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016; Nouwens et al., 2016; Sesma et al., 2009). Reading comprehension is thought to utilize working memory when the reader stores and holds information from the text at the sentence level, and especially the larger passage level. By rapidly storing, recalling and manipulating elements from across the full breadth of language (from phonology to morphology to syntax, etc.), the skilled reader is able to form literal and inferential judgments about the text. This active RC process the reader undertakes is also called "extracting and constructing" knowledge (Snow & Sweet, 2003, p. 1). Reading also requires other cognitive processes beyond working memory, which are described in the sections on inhibition and shifting below.

Inhibition. Inhibition is defined as "the ability to override prepotent or automatic responses" (Jacob & Parkinson, 2015, p. 519). Initially, a variation on the Stroop task (Stroop, 1935) was commonly used to measure inhibition. In some versions of this task, a student is asked to read a series of color words (e.g., red, orange, blue) written in opposing colors (e.g. the word orange written in blue ink). Other inhibition measures require the student to name shapes according to a conflicting pattern, such as calling a square a circle and vice versa (see NEPSY-II Inhibition subtest). Tasks that measure inhibition are intended to assess students' ability to supersede an ordinary response when processing and responding to information in the environment. In a synthesis of executive function and reading comprehension for upper elementary school students (see Chapter 2), inhibition was found to contribute significant unique variance to the prediction of RC in two studies (see Cutting et al., 2009; Kieffer et al., 2013). Inhibition is thought to relate to reading because reading for understanding requires sustained attention to both the code and meaning of a text for comprehension.

***Shifting*.** Shifting is "the ability to intentionally move backward and forward between tasks, mental sets, or goals" (Baggetta & Alexander, 2016, p. 15). Shifting is also referred to as *cognitive flexibility* (Diamond, 2013) or *switching* (Miyake et al., 2000). Dimensional card sorting tasks are often used to measure shifting, such as variations on the Wisconsin Card Sorting Task (see Kongs, Thompson, Iverson, & Heaton, 2000). In a synthesis I conducted (see Chapter 2 of this manuscript), shifting contributed significant, unique variance to the prediction of RC in two studies (see Kieffer et al., 2013; Nouwens et al., 2016). Shifting is thought to relate to reading because processing texts requires the flexible use of implicit 'rules' regarding text structure and organization. Good readers are encouraged to 'stop and think' while they read and to make connections outside the text, which are both strategies that require shifting of mental images and thoughts. To borrow from Kendeou, Broek, Helder and Karlsson's language (2014) in their argument for a cognitive view of RC, it is important for researchers to dissect the *process* of reading as much as the *product* in order to understand why it succeeds or fails.

Diverse Learners

Linguistically Diverse Learners. The relationship between EF and RC may differ depending on students' language background. I use the term linguistically diverse learners

(LDLs) purposefully to indicate the range of language ability students may possess in their native language (L1) and their second language (L2), which is English for most studies conducted in the United States. Students who are fully bilingual are those who have proficiency in both languages. Emerging bilingual students have proficiency in one language and are learning another. When schools identify students as ELs, they often do so because those students speak a language other than English at home and they have not passed an English proficiency test, but their proficiency in their home language may not have been measured. In reviewing other studies for this manuscript, I replicated the language status terms as they originally appeared in each study, but will use the terms which the participating schools prescribed to classify and describe my own sample of students.

In a synthesis of the research (see Chapter 2) on EF for students in upper elementary school, three of the seven studies reviewed included ELs, broadly defined (Gerst et al., 2015; Kieffer et al., 2013; Nouwens et al., 2016). However, none of the three studies conducted any additional analyses to test whether EF and RC's relationship differed for this population (i.e., comparing group means on individual measures or as an interaction term in the regression analysis). More explicitly, several of the other studies I reviewed actively excluded students who the school district or parent survey identified as ELs or as speaking a language other than English at home, respectively.

Aside from my synthesis, other studies of early childhood, elementary school, and adult native bilingual and second-language learner students have found EF performance differences between groups. However, the 'bilingual advantage' is often found in individuals who are native bilinguals, not emerging bilinguals (e.g., Bialystock, Barac, Blaye, & Poulin-Dubois, 2010; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Costa, Hernández &

Sebastián-Gallés, 2008). Some studies have found that emerging bilingual students (i.e. those not fully proficient in two languages) do not perform as well as fully bilingual students on measures of cognitive control (Bialystok, 1988). Furthermore, ELs show greater RC deficits compared to their monolingual peers as they move into upper elementary school, plausibly because of the increasing demands of comprehending academic texts. Therefore, EF studies are urgently needed to observe how the relationship may differ for linguistically diverse students who face additional challenges in RC development in upper elementary school.

Students with Disabilities. There is mounting evidence that students with disabilities in reading also show weaknesses in EF measures. For instance, Locascio and colleagues (2010) conducted a study with three groups of students ages 10–14: students with word reading deficits, students with specific RC deficits, and typical readers. They found that students with specific RC deficits performed below their typically developing peers on an EF planning factor, despite controlling for decoding. Although the students in the word reading deficit group performed below the typical reader group on an EF inhibition factor, this difference was mitigated when controlling for decoding. Borella, Carretti, and Pelgrina (2010) conducted a study of 10–11-year-old “good” and “poor comprehenders” and found that poor comprehenders showed deficits in working memory and inhibition compared to their peers (p. 541). Sesma et al. (2009) and Cutting et al. (2009) both investigated students in upper elementary school with RC difficulties and found that they performed lower than their peers without RC deficits on EF measures, despite controlling for decoding and linguistic comprehension. Across studies, prevalence rates of specific RC difficulties range from 3–15% (e.g., Leach et al. 2003; Nation & Snowling, 1997; Stothard & Hulmne, 1995). These studies suggest that students with reading difficulties may demonstrate additional EF deficits compared to their peers, which has implications for both the identification and intervention of such students.

English Learners with Disabilities. English learners with disabilities may have even more difficulty with RC because they are dealing with language and disability issues simultaneously (Ortiz & Artiles, 2010; Hauerwas, Brown, & Scott 2013). Yet, limited research has been done on this population of students. Distressingly, a recent review of multi-state practices regarding ELs with disabilities opened with the bleak statement: “No single method has proven effective in differentiating between English learner students who have difficulty acquiring language skills and those who have learning disabilities” (Burr, Hass, & Ferriere, 2015, p. i). The authors do recommend a structured list of steps that local educational agencies (LEAs) can undertake to reduce the current rates of error in the identification of ELs with disabilities, which paradoxically includes patterns of both over-identification and under-identification (Rueda & Windmueller, 2006; Sullivan, 2011). An in-depth look at referral practices within one state found that EL students were underrepresented in special education in elementary school, but overrepresented in secondary school (Artiles, Rueda, Salazar, & Higareda, 2005). This study points again to the critical importance of disentangling the multiple factors that ELs with disabilities face. As I do not know of any research conducted to date, a study that examined the contribution of EF to RC with even a small sample of EL students with disabilities would be of vital significance to advancing the field of second language and disability education.

Purpose

The purpose of this study was to extend the current literature base on EF’s contribution to RC. The following research questions guided this study:

1. Is there a relationship between latent EF and latent RC, controlling for LC and DC? Is the relationship direct or mediating?
2. Does language status moderate the effect of latent EF on latent EC?
3. Does disability status moderate the effect of latent EF on latent RC?

Hypotheses

Findings from previous research suggest that EF will make a significant, unique, and direct contribution to RC (Kieffer et al., 2013; Christopher et al., 2012). Additionally, EF is hypothesized to mediate the contribution to RC via linguistic comprehension and decoding. For the subgroups of interest in the study, it is hypothesized that language and disability status will modify EF and RC's relationship. A hypothesized path model for the direct and mediating contributions of EF to RC is presented in Figure 3.1. This model draws on the previous work of Kieffer et al. (2013) and Christopher et al. (2012) about the hypothesized relations between the variables of interest, while also employing progressive, theory-driven statistical models to extend the current body of work on EF.

Method

Design

The study used data collected from a federally funded quasi-experimental study conducted in the Mid-Atlantic area, known as the CLAVES Project, which stands for *Comprehension, Linguistic Awareness, and Vocabulary for ELs* (Proctor, Silverman, & Harring, 2013). The purpose of the CLAVES project was to design and evaluate the feasibility and efficacy of a supplemental reading intervention for fourth and fifth grade emerging bilingual students. The CLAVES project was conducted in two sites. Data in this manuscript is from one site in the Mid-Atlantic region. At this site, the sample included 128 fourth grade students nested within 10 classrooms who were identified by their schools as either English learners (EL) or reclassified English learners (R-ELs) at the start of the school year.

Sample

Data used in this study was collected in four public elementary schools that serve a predominantly Hispanic/Latino population of students and their families. Student background data was obtained from school reports about participating students' race/ethnicity, gender, and eligibility status for supplemental services related to language, disability, and socioeconomic status. Home language data was requested from parents via a supplemental survey and was used for descriptive purposes only. Across these four schools, about 90% of students were eligible for free or reduced lunch, which is an indicator of lower socioeconomic status. According to public records, approximately 57% of the students across all four schools were identified as ELs and were receiving supplemental English as a second or other language (ESOL) services at the time of the study. About 7% of the students were identified as receiving special education services through an Individualized Education Plan (IEP). See Table 3.1. All students who agreed to participate in the CLAVES study were eligible for inclusion in the current study (N = 128). However, only a portion of the CLAVES assessment and student background data were used in the present study.

Procedures

Language Status. Language status was obtained from school records. All students in the participating district undergo an initial assessment of English language proficiency if their parents indicate that a language other than English is spoken in the home when they enroll in school. Students whose parents do not report speaking a language other than English are considered "English Only" (EO) by the district. Students who speak another language are assessed using the World Class Instructional Design and Assessment (WIDA) Assessing Comprehension and Communication in English State-to-State (ACCESS) for English Language Learners test (WIDA Consortium, 2010). Based on students' performance on a variety of tasks, students receive four individual proficiency ACCESS scores for the domains of listening, reading, speaking, and writing (ranging in value from 1.0 "Entering" to 6.0 "Reaching"), as well as an overall composite score that contains the weighted average across the four domains. In the school district participating in the current study, students whose average score is below 5.0 are considered "English Learners" (ELs) and receive supplemental English language instruction. Students whose score is 5.0 or above are considered to have passed the proficiency test and are termed "re-classified English learners" (R-ELs) by the district. For the current study, the district's classification of students' language background and proficiency were used as the categorical measure of language status, which include R-EL and EL. Students were eligible to participate in the present study if they had been reclassified in the previous two academic years by the district.

Disability Status. Student records of enrollment in an Individualized Education Plan (IEP) were obtained for participating students from school records. Additionally, a dummy variable for students' disability status was created for the presence of a disability (coded as 0 = no, 1 = yes).

Socioeconomic Status. School records of eligibility for free and reduced meal status (FRP) were obtained for participating schools. This status is collected by the National School Lunch Program (NSLP) and is used by Local Educational Agencies (LEAs) to provide meals to students at reduced prices based on economic need. Although FRP status interpretation only provides an approximation of household income, it has been widely used as a crude indicator of socioeconomic status more broadly in education and public health literature. School FRP status was used as a descriptive variable only to contextualize the sample. Since the majority of the students in the present sample are eligible for FRP (above 90%), it was not used as a control variable in the regression equation.

Data Collection. Data used for the present study were collected in spring of one academic year. Most assessments were group-administered unless otherwise noted below. Administrators of all measures were trained on how to give the tests with reliability prior to entering the field. Additionally, research assistants were trained to score all of the measures according to the standardized test procedures. All data were single and double scored and single and double entered to ensure reliability of the scores. When needed, supplemental rubrics were developed to determine the appropriate coding of student responses.

Measures

Assessments were group administered and raw scores were used in analyses unless otherwise noted below. All assessments were administered at one time point in the spring of the academic year.

Latent Reading Comprehension. Three group administered norm-referenced measures of RC were used to form a latent construct. There is a robust line of research that finds latent constructs represent the underlying theoretical conceptualization of RC (Fletcher, 2006; Mehta, Foorman, Branum-Martin, & Taylor, 2005; Vellutino, Tunmer, Jaccard, & Chen, 2007). More specifically to this study, previous research supports the validity of using these three measures to form a latent construct of RC (Silverman et al., 2015; Proctor et al., 2012). The advantages of using a latent variable as the dependent variable of interest are discussed further in the analytic plan.

MAZE. The Achievement Improvement Monitoring System (AIMSweb) MAZE assessment is a measure of silent reading efficiency and comprehension at the sentence and passage level (Pearson, 2014). Students are given three minutes to read a passage and select the correct missing words. After the first sentence, every seventh word is replaced with three words in brackets, which include two distractors and one correct word. Students are instructed to circle the word that makes sense in the sentence. Students earn one point for each correct word circled, with a maximum of 46 points. The test makers report split-half reliability of .95 and alternate form reliability of .95 for fourth grade students.

Test of Silent Reading Efficiency and Comprehension. The Test of Silent Reading Efficiency and Comprehension (TOSREC) is a measure of silent reading fluency and comprehension (Wagner, Torgeson, Rashotte, & Pearson, 2010). Students are given three minutes to read a series of statements and choose whether each statement is true or false, by marking 'yes' or 'no.' Students earn one point for each correct answer, yet lose one point for each incorrect answer. The raw score is calculated by subtracting the total incorrect answers from the total correct answers. Test makers report alternate form reliability of .86 for fourth grade.

Gates MacGinitie Reading Comprehension. The Gates MacGinitie Reading Test (GMRT) Reading Comprehension subtest is an extended measure of passage reading comprehension that can be administered in a group setting (MacGinitie, MacGinitie, Maria, & Dreyer, 2002). In this measure, students are given 35 minutes to read multiple passages of narrative and informational text and are expected to answer 4-6 inferential and literal questions about each passage, for a total of 48 questions. Students receive one point for every correct answer, with the total number of points forming the raw score (maximum 48). This measure was selected because it has been used frequently in previous studies as a measure of RC outcomes (Cutting & Scarborough, 2006; Locascio et al., 2010; Ritchey, Silverman, Schatschneider, & Speece, 2013; Tong, Deacon, & Cain, 2013) The GMRT reading comprehension subtest demonstrated test-retest reliability of .81 for fourth grade according to the test maker (MacGinitie et al., 2002).

Latent Executive Function. Three norm-referenced measures representing established components of EF, working memory, shifting, and switching, were used to form a latent construct of EF. Although many studies have examined these domains as separate constructs (e.g., Cutting et al., 2009, Gerst et al., 2015, Nouwens et al., 2016), there is at least theoretical support for the fundamental unity of the construct as well (Miyake et al., 2000; Diamond, 2013). The advantages of using a latent construct are discussed further in the analytic plan below. Trained research assistants individually administered all of the EF measures individually.

Working memory. The Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) Digit Span Forwards (DSF) and Digit Span Backwards (DSB) are two individually administered conditions of the working memory subtest (Wechsler, 2003). In the DSF condition, students listen while an examiner reads a list of numbers and then students repeat the numbers aloud. In the DSB condition, students listen to a list of numbers and then repeat the sequence of numbers backwards. The DSF task provides a measure of students' working memory storage. The DSB task provides a measure of how many numbers a student can hold and manipulate. In the DSF condition a sequence of up to 9 numbers are presented and in the DSB condition a sequence of up to 8 numbers are presented. Students receive one point for each number recalled correctly. The total percentage correct on the DSB condition was used in analyses. Test makers report split-half Spearman Brown reliability of .87.

Inhibition. The Neuropsychological Developmental Assessment – Second Edition (NEPSY-II) Inhibition subtest is a timed measure of inhibition (Korkman, Kirk, & Kemp, 2007). The subtest contains four successive conditions in which students are given rules for rapidly naming shapes and directionality of arrows and then prompted to reverse or inhibit the conventional naming procedures. For example, students are prompted to verbally identify opposing shapes, such as referring to squares as circles, or directions, such as saying arrows that are pointing up are pointing down. In this way, the combined subtests provide a measure of students' ability to inhibit a reflexive response for a novel response. For each condition, the examiner records the total time to complete the task, self-corrections, and errors. The total number correct was used in analyses. Test makers report reliability of .73-.90.

Shifting. The Delis-Kaplan Executive Function System (D-KEFS) Trail Making Test (TMT) consists of five successive conditions, which provide individual measures of scanning, sequencing, shifting, and motor speed, respectively. In the first three conditions, the student scans a sheet of paper and connects sequences of numbers and letters. In the fourth condition, *Letter-Number Switching,* the student connects an alternating sequence of numbers and letters: A, 1, B, 2, C, 3, and so on. Students are given a maximum of 4 minutes to complete the task. The total number of correct number/letter sequences within the time limit is calculated by the administrator after the test. The total time to complete the task used in analyses. Test makers report test-retest reliability of .89.

Linguistic Comprehension. Linguistic comprehension was represented by performance on the Core Academic Language Skills assessment (CALS; Uccelli et al., 2015).

The CALS assessment measures linguistic awareness across content area tasks presented in eight subtests, which include connecting ideas, tracking themes, organizing texts, breaking words, comprehending sentences, identifying definitions, epistemic stance, and metalanguage. Test makers report split-half reliability of .90 and coefficient α of .93.

Decoding. The Test of Sight Word Reading Fluency-Second Edition (TOSWRF-2) was used as a measure of reading accuracy and fluency at the word level (Mather, Hammill, Allen & Roberts, 2014). Students are given a list of 220 words and prompted to identify the beginning and end of each word by placing a slash at the end of the word. For example, students would be expected to four draw lines to delineate the beginning and ending of words in the sequence UPSEEWHYLONG. The words in the assessment are organized from easiest to most difficult. Students are given 3 minutes to complete the task. Test makers report test-retest reliability of .93.

Data Analysis

Analytic Plan. The analytic plan for the study followed the conventional set of procedures for conducting quantitative analyses. First, threats to validity were examined and mitigated to the extent possible. Next, missing data were identified and determined not to disproportionately affect the analysis. Then, descriptive statistics were obtained and reviewed. Then, model assumptions were tested. Finally, structural equation modeling, including confirmatory factor analysis and path analysis, were used to answer the research questions of interest. SPSS AMOS Version 25 was used to conduct structural equation modeling (Kline, 2015).

Missing Data. Prior to running the models below, the datasets were checked to determine the scope of missing data and its potential for bias on the analysis. Out of 128 participants, 2 were missing assessment data due to absences within the assessment window at the school sites. The missing data represents less than 2% of the total participant data collected and was missing at random.

Statistical Analysis.

Descriptive statistics. Preliminary data analysis included examining the descriptive statistics for the variables of interest. The means and standard deviations of each measure were calculated. Additionally, the variables of interest were examined visually through histograms and P-P and Q-Q plots to look for issues of skew, kurtosis, univariate and multivariate outliers, and normality of residuals. See Table 3.2 for means and standard deviations. The range in skew was from -0.021 and 1.796 which is within the guidelines for normal distribution. The range in kurtosis was from 0.097 and 5.597. The kurtosis value of 5.597 for the MAZERS is outside of the acceptable range.

Multicollinearity. Structural equation modeling, like ordinary least squares regression, assumes that there is not multicollinearity between the variables of interest. Multicollinearity was examined in the current study by reviewing the Variance Inflation Factors (VIF) for each of the predictor variables. VIF values were between 1.069 and 1.335. These values are considered acceptable. VIF values above 10, indicate multicollinearity must be addressed in the model (Menard, 1995). Additionally, a tolerance statistic (1/VIF) was obtained for the predictor variables. The tolerance statistics ranged from 0.749 to 0.935, which are considered acceptable. Tolerance values below 0.2, indicate multicollinearity of predictors may be an issue (Menard, 1995).

Structural Equation Modeling. To answer the primary research question, a path analytic model was used to examine the direct and mediating effects of working memory, inhibition, and shifting on a latent variable of RC, controlling for latent linguistic comprehension and decoding. Path analysis conducted within structural equation modeling seeks to understand and explain patterns of covariance among a set of variables (SEM; Bowen & Guo, 2011; Hancock & Mueller, 2013; Kline, 2015). Path analysis was conducted in SPSS AMOS.

Given the current research questions, structural equation modeling (SEM) offered several advantages over regression analyses. First, SEM is flexible enough to allow for the construction of one multi-component model that reflects the theoretical relationships among multiple dependent and independent variables. In the study, the modeling process included mapping the latent variable of EF onto the latent variable of RC, while accounting for other potential contributors to RC (i.e., decoding and linguistic comprehension). Second, SEM allows for the use of both single measures and latent constructs within the same model. In this case, decoding and linguistic comprehension were obtained via single variables or manifest variables, while EF and RC represent latent constructs formed by three individual measures each. Third, SEM allows for the specification of measurement error, which is especially important in a model with a relatively small sample size.

Model testing. The following path analysis model was tested using multilevel structural equation modeling and confirmatory factor analysis through SPSS AMOS (CFA; Browne & Cudeck, 1993; Kline, 2015). See Figure 3.1. Circles in the model represent latent variables and rectangles represent measured variables. The circle labeled “Latent Executive Function” represents the latent variable of EF, formed by three measured variables, which are represented by the boxes “Working Memory,” “Inhibition,” and “Shifting.” The circle “Latent Reading Comprehension” represents the latent variable of RC, formed by three measured variables, which are represented by the boxes “MAZE,” “GMRT”, and “TOSREC.” The rectangle “Linguistic Comprehension” is an observed variable of linguistic comprehension formed by a single assessment that encompasses items related to semantics, morphology, and syntax. The rectangle “Decoding” is a measured variable formed by a single assessment. Single-headed arrows indicate regression while double-headed arrows indicate covariances among the predictors (“latent EF”) and the residual covariance between the mediators (“decoding” and “linguistic comprehension”). The statistical significance of the direct and indirect paths was obtained by evaluating the path weights. Additionally, model fit was evaluated by examining several fit indices: chi-square test (x^2), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), goodness of fit index (GFI), and Akaike information criterion (AIC). A GFI above 0.90 suggests that the proposed model fits the data well (Kline, 2015). Guidelines for acceptable fit for each index are chi-square tests above .05, CFI above 0.90, TLI above 0.90, and RMSEAs below 0.10 (Hu & Bentler, 1999; Kline, 2015).

Results

The purpose of the study was to explore whether EF contributes to RC. The primary research questions investigate whether latent EF, as measured by working memory, inhibition, and shifting tasks, made a unique, direct contribution to latent RC beyond the contributions of language comprehension and decoding, among linguistically diverse learners. The secondary research question considered whether latent EF mediated the explanation of RC via language comprehension and/or decoding. For each research question, I also explored whether disability status and/or language status moderated the effects of EF to RC.

Means and standard deviations for the observed measurement variables are provided in Table 3.2. Significant differences between subgroups based on language status and disability status were present based on one-way ANOVAs, conducted as secondary analyses. EL students performed significantly below their peers on two measures of executive function: working memory ($F[1,124] = 12.03$, $p = .001$) and inhibition ($F[1,124] = 13.97$, $p = .001$), decoding ($F[1,124] = 5.59$, $p = .020$), linguistic comprehension ($F[1,124] = 12.03$, $p = .001$), and two measures of reading comprehension: GMRT ($F[1,124] = 6.38$, $p = .013$) and TOSREC ($F[1,124] = 3.35$, $p = .070$). Students with IEPs performed significantly below their peers on two measures of executive function: working memory ($F[1,124] = 6.42$, $p = .013$) and inhibition ($F[1,124] = 5.43$, $p = .021$), decoding ($F[1,124] = 7.436$, $p = .007$), and one measures of reading comprehension: MAZE ($F[1,124] = 4.31$, $p = .040$).

Correlations among observed assessment variables are provided in Table 3.3. Among the executive function measures, working memory and shifting showed a small, but significant correlation ($r = .286$, $p = .001$). Working memory showed a small but significant correlation with decoding ($r = .238$, $p = .007$) and GMRT (.197, $p = .027$). Shifting showed a small but significant correlation with decoding ($r = -.219$, $p = .014$). Inhibition showed a small but significant correlation with decoding ($r = .241$, $p = .006$), linguistic comprehension ($r = .229$, $p = .010$), and two measures of reading comprehension (GMRT [$r = .265$, $p = .003$], MAZE [$r = .194$, $p = .029$]). Decoding showed a moderate correlation with linguistic comprehension ($r = .425$, $p = .001$) and with all reading comprehension measures (GMRT [$r = .386$, $p = .001$], MAZE [$r = .606$, $p = .001$], TOSREC [$r = .408$, $p = .001$]). Linguistic comprehension showed a moderate correlation with all reading comprehension measures (GMRT [$r = .622$, $p = .001$], MAZE [$r = .523$, $p = .001$], TOSREC [$r = .378$, $p = .001$]).

Among the reading comprehension measures, all showed moderate correlations with each other GMRT, MAZE, and TOSREC showed moderate correlations (GMRT to MAZE [r = .505, p = .001]; GMRT to TOSREC [r = .427, p = .001]; MAZE to TOSREC [r = .529, p = .001). Examining patterns among these associations, it appears different measures of executive function have differing associations with decoding, linguistic comprehension, and reading comprehension measures. While inhibition is significantly correlated with linguistic comprehension, decoding, and all measures of reading comprehension, working memory is significantly correlated with linguistic comprehension and one measure of reading comprehension, and shifting is only significantly correlated with decoding.

Research Question 1

The first research question examined whether there was a relationship between latent EF and latent RC controlling for linguistic comprehension and decoding and, if so, whether that relationship was direct or indirect. Results showed that latent EF has direct relationship with latent RC when decoding and linguistic comprehension are not included in the model and indirect relationship with latent RC when decoding and linguistic comprehension are included in the model. Model 1, which represents the direct effect relationship between latent EF and latent RC without decoding and linguistic comprehension, had the following indices: x^2 (8, N=126) = 8.601, p=0.377, CFI=.994, TLI = .988, RMSEA = .025, and AIC = 34.601. The relationship between latent executive function and latent reading comprehension was significant (p=.037). The standardized estimate for the path weight from latent EF to latent RC was .40. Model 2, which represents the direct relationship between latent EF and latent RC including decoding and linguistic comprehension had the following indices, which indicate poor fit: x^2 (19, N=126) = 68.676, p = 0.001, CFI = .784, TLI = .682, RMSEA

= .145, and AIC = 102.676. The relationship between latent RC and decoding was significant (p=.001), the relationship between latent RC and linguistic comprehension was significant (p=.001), but the relationship between latent executive function and latent RC was not significant (p=.997). Model 3, which represents the mediational relationship between latent EF and latent RC through decoding and linguistic comprehension, showed acceptable fit with the following fit indices: x^2 (18, N=126) = 35.957, p = 0.007, CFI = .922, TLI = .879, RMSEA = .089 and AIC = 71.957. The relationship between latent executive function and decoding was significant (p=.013), and the relationship between latent executive function and linguistic comprehension was significant (p=.014), the relationship between latent RC and decoding was significant (p=.001), and the relationship between latent RC and linguistic comprehension was significant (p=.001). Model 3 presented the best fit indices and was theoretically consistent with the hypothesized relationship between the variables of interest, therefore all subsequent subgroup analyses were conducted using Model 3. The results of the primary research question suggest a mediating effect of latent EF to latent RC through linguistic comprehension and decoding. Figures 1, 2 and 3 depict the structural models for Models 1, 2, and 3, respectively, along with the corresponding standardized regression estimates for each path. Table 3.4, Table 3.5 and Table 3.6 depicts the measurement information for each model. Table 3.7 depicts the comparative fit indices for each model. Principal component analysis coefficients for latent RC and latent EF are displayed in Table 3.8 and Table 3.9.

Research Question 2

For research question 2, model 3 was used to test for differences between EL and REL groups using AMOS multi-group analysis. The model itself was considered acceptable having the following model fit indices: x^2 (36, N=126), 51.387, p = .046, CFI = .920, TLI = .876 and RMSEA = .059. A chi-squared difference test of the nested model comparisons found significant differences between the EL and REL groups for the overall model (x^2 (8, N=126), 25.941, p = .001). However, no significant differences between EL and REL groups were found for the paths from LEF to LC (p = .702) and LEF to DC (p = .361). These results suggest the relationship between latent EF and decoding and latent EF and linguistic comprehension did not differ significantly based on language status for students in the sample.

Research Question 3

For research question 3, model 3 was used to test for differences between IEP and non-IEP groups using AMOS multi-group analysis. The model itself was considered acceptable having the following model fit indices: x^2 (36, N=126), 69.109, p = .001, CFI = .858, TLI = .779, and RMSEA = .086. A chi-squared difference test of the nested model comparisons found significant differences between the groups for the overall model (x^2 (8, N=126), 18.850, p = .016). However, no significant differences between IEP and non-IEP groups were found for the paths from LEF to LC (p = .329) and LEF to DC (p = .719). These results suggest the relationship between latent EF and decoding and latent EF and linguistic comprehension did not differ significantly based on disability status for students in the sample.

Discussion

Ultimately, this study advances the understanding of the relationship between EF and RC in fourth grade, and by doing so, takes another step towards more appropriately targeting interventions for students with deficits in RC. This study found latent EF adds to the prediction accuracy of latent RC, which provides support for future research to investigate whether interventions to improve EF skills are linked with better RC outcomes, especially for linguistically diverse learners and students with disabilities. The present study has implications for diagnostic assessment batteries for identifying students who struggle with reading because it offers insight into the domains affected that may contribute to reading difficulty.

It was hypothesized that executive function would make both a direct and indirect contribution to the prediction of reading comprehension. In isolation, latent executive function presented a small to moderate sized standardized regression coefficient (.40) to latent RC. However, when decoding and linguistic comprehension were added to the model, executive function no longer directly contributed to reading comprehension. Therefore, the model that accounts for the established models of reading comprehension, which includes decoding and linguistic comprehension, is most theoretically sound. When viewing the mediating contribution of executive function to reading comprehension through decoding and linguistic comprehension, latent executive function had a direct effect of .61 and .66 on decoding and linguistic comprehension, respectively, as well as a mediating effect on latent RC of .29 and .35, respectively.

Considering these results in the context of previous research, the findings seem consistent with studies that employed similar models of executive function and reading comprehension. Sesma et al. (2009) found working memory ($b = .17$, $p = .032$), planning ($b = -.25$, $p = .013$), and attention ($b = -.13$, $p = .001$) were significant predictors of RC (b = .17, b = - .25, while controlling for decoding, reading fluency, and vocabulary breadth. Christopher et al. (2012) found that working memory ($p < .01$) and processing speed (p

= .08), were significant and marginally significant predictors of RC, respectively, when controlling for decoding. Cutting et al. (2009) found planning (*b* = -0.28, *p* = .05) was a significant predictor of RC, when controlling for decoding, reading fluency, and oral language. Gerst et al. (2015) found working memory measures (*b* = .24, *p* = .01; *b* = -0.25, *p* = .002) were significant RC predictors controlling for age, oral comprehension, decoding, and gifted and talented status. Jacobson et al. (2016) found processing speed ($\Delta R2 = .015$, $p < .001$) and a problem-solving, working memory, switching composite measure ($\Delta R2 = .045$, $p < .001$), were significant predictors of RC, controlling for language skills. Kieffer et al. (2013) found shifting (*b* = .16, $p < .05$) and inhibition (*b* = .19, $p < .05$) were significant predictors of RC, controlling for decoding, language comprehension, working memory, and processing speed. The size of the regression weights from latent EF to decoding, latent EF to linguistic comprehension, and latent EF indirectly to RC through decoding and linguistic comprehension in the present study are larger than the beta coefficients found in previous studies. This finding lends support to the methodological decision to form a latent variable for EF in order to discern the combined impact of multiple domains of EF on the RC, decoding, and linguistic comprehension variables of interest.

Differences by Language Proficiency

The results of the present study found no significant differences in the path from latent executive function to decoding and to linguistic comprehension between EL and REL students. However, in a secondary analysis, one-way ANOVAs revealed EL students performed significantly below their REL peers on measures of working memory, inhibition, decoding, linguistic comprehension, and all three measures of reading comprehension. These analyses were conducted to inform future work with EL and REL populations. A synthesis of research on executive function and reading comprehension found empirical studies have often excluded students with limited English proficiency (see Chapter 2). When English learners are included in samples, studies have not conducted any subgroup analyses to examine group differences either through regression or interaction terms (Gerst et al., 2015; Kieffer et al., 2013; Nouwens et al., 2016). As such, the present study extends the research base by systematically investigating group mean differences and moderation by language status for students with limited English proficiency.

Differences by Disability Status

The present study found revealed students with disabilities performed significantly below their peers on measures of working memory, inhibition, decoding, and one measure of reading comprehension when examining group differences through one-way ANOVAs, conducted as a secondary analysis. However, these differences between students with disabilities and students without disabilities were not significant when chi-square difference tests were performed for the regression path weights from latent EF to decoding and latent EF to linguistic comprehension. Among other studies that have examined the link between executive function and reading in students with disabilities, Borella et al. (2010) found students with RC difficulties performed below their typical reader peers on measure of working memory and inhibition. Nouwens et al. (2016) found that students with disabilities performed significantly below peers without disabilities on measures of decoding and RC, but performed similarly on measures of vocabulary, working memory, inhibition, storage, planning, and cognitive flexibility based on ANOVA with disability status as between-subjects factor. Nouwens et al. (2016) also added interaction terms for disability status and performance on each measure in the full regression model predicting reading comprehension, but found those terms were not significant. Cutting et al. (2009) found students with reading comprehension difficulties performed significantly worse on measures of inhibition and planning compared to proficient readers. However, students with reading comprehension difficulties performed similarly to their peers on a measure of working memory (Cutting et al., 2009). Locascio et al. (2010) found that students with RC difficulties performed worse than their peers on planning, when controlling for decoding, and students with decoding difficulties performed comparatively worse on inhibition, although this difference disappeared when decoding was added to the model. The majority of previous studies that included students with disabilities only conducted independent ANOVAs to test for mean differences among subgroups, rather than including disability status in the full hierarchical regression models. The present study extends the existing base of literature on students with disabilities by examining whether those differences impact the relationship between executive function and linguistic comprehension and decoding.

Limitations

Several limitations hampered the ability of the present study to fully investigate the research questions posed. First, the relatively small sample size given the number of parameters to be estimated reduced the degrees of freedom and ability to detect small effects, had they been present. Additionally, due to the relatively low prevalence of students with disabilities in the typical population, it was not possible to have a large enough sample of students with disabilities to meaningfully interpret differences in their results. For instance, the percentage of students with disabilities in the participating schools ranges from just 5% to 9%. Similarly, the present study included only 10 students with disabilities, or 8% of the total sample. Further, although the latent construct of EF has advantages in terms of reducing pooled error, the present study used only one measure for each factor of inhibition, shifting, and working memory. This limited the practical context for each executive function construct to one type of task. Finally, the present sample included students who were currently classified as English learners or who had recently exited from the English learner program in the past two years. The inability to detect differences between EL and REL students may similarly be due to the small sample size in the present study. It may also be possible that the purposeful selection of students at higher levels of English language proficiency for admission in the study (WIDA levels 3.0 and above) limited the ability of the present study to adequately explore differences between the full breadth of EL students compared to REL students. Furthermore, the present study was conducted in schools with students of predominantly Spanish-speaking backgrounds, yet the primary language of instruction in school was English. Additional Spanish language proficiency assessments were not administered by the researchers to determine each student's degree of bilingual literacy. Therefore, the study cannot be compared directly to previous research with fully bilingual adults. The sample also did not include students who were native English speakers, so direct comparisons cannot be made to research conducted with English-only speakers. The current study represents assessments administered at one time point and offers only correlational, not casual, data on the relationships between executive function and reading comprehension.

Future Directions

The present study makes an important contribution to the field, but also presents questions and implications for future research. The debate over the unity of the construct of EF continues to present differences in the field. Future studies should seek to explore the measurement of executive function and employ large sample sizes as well as multiple measures for each domain of EF. In order to understand how executive function may differ among students with disabilities, future research should purposefully over-sample students with disabilities in order to adequately obtain a sample size in which half of the students have disabilities.

While the present study examines correlational relationships between executive function and reading comprehension for linguistically diverse students in upper elementary school, future research should consider whether and how executive function interacts with students' growth in reading throughout the year. In addition, executive function may interact with students' response to reading interventions. Future research should examine EF moderates the effects of small group reading intervention on RC. Such studies would have strong implications for reading interventions because the experimental design would yield stronger conclusions about the direction of relationship between EF and RC. In particular, it would be important to include students with disabilities and students with limited English proficiency in these samples to consider how EF functions for linguistically diverse learners.

Chapter 4: The Contribution of Latent EF to the Effects of a Reading Intervention on Students' Reading and Linguistic Comprehension

Introduction

Executive function is a term that encompasses the interrelated cognitive processes used to perform tasks. A host of individual studies have connected EF to reading comprehension (RC) outcomes, and findings have been summarized in several comprehensive meta-analyses. For instance, Baggetta and Alexander (2016) provide a review of definitional agreement in the field with regard to theoretical models of EF, summarize the assessments associated with each component of EF, tabulate populations and disciplines within which EF has been conducted, and summarize outcomes with which EF is associated. Although the review provides thorough analysis of the qualitative trends across studies, it does not include a lengthy discussion of the analytic models used to date to examine the relative contribution of EF to each outcome of interest.

Jacob and Parkinson's (2015) review of EF across studies also provide guidance for the current study. Specifically, the authors used correlational techniques in conducting of a meta-analysis to examine the average meta-effect sizes for reading and math outcomes. Additionally, Jacob and Parkinson (2015) disaggregated the unconditional meta-analytic correlations across four EF domains of (working memory, response inhibition, attention control, and attention shifting), three age groups (3–5 years, 6–11 years, and 12–18 years), two study designs (concurrent and predictive validity), and two settings (naturalistic or laboratory-based). Related to the present study, Jacob and Parkinson's (2015) meta-analysis provides a basis for understanding the magnitude and direction of EF's relationship with reading outcomes in typically developing students in elementary school: the relationship is positive and the average effect size is moderate (r = 0.36). However, their meta-analysis excluded any study conducted solely with students with disabilities or ELs. Given the growing diversity of U.S. schools, conducting research on the role of EF with these populations is of importance to the field (see Chapter 2 and Chapter 3).

While correlational research is helpful in revealing relationships between EF and its component variables and reading comprehension, intervention research is needed to determine whether or how EF interacts with instruction. Jacob and Parkinson (2015) provide a meta-analysis of interventions designed to directly improve EF and, by influencing EF, indirectly improve reading outcomes. However, these researchers did not find fully convincing evidence that EF intervention studies led to gains in achievement when the necessary covariates were added into the models. Furthermore, these students only focus on studies focused on changing EF. These studies did not investigate whether EF moderates the effects of intervention targeting other reading-related skills.

While research on the role of EF in instruction targeting reading comprehension is fairly nascent, research on intervention focused on skills much more closely associated with reading comprehension such as decoding, strategy, and/or language instruction has a long history and provides substantial evidence of potential effects on RC outcomes. However, research is needed on whether students' EF moderates their response to literacy intervention. In this scenario, EF could be the unmeasured third variable not within the model that would explain students' differential response patterns to the same reading intervention. In other words, if some students have simultaneous underlying EF, decoding, and/or linguistic comprehension deficits but are assigned reading interventions that do not target EF deficits directly, then those students may not realize the same gains as students without EF deficits.

The preponderance of the research connecting EF to reading outcomes has been conducted with English monolingual students in early childhood and early elementary school (see Jacob & Parkinson, 2015 for a review). Concurrently, the majority of research connecting EF to reading outcomes with linguistically diverse populations has been conducted with fully bilingual children and adults (see Bialystok, 2015). Therefore, there are several populations that have been not been largely ignored in the literature to date. These populations include (a) English Learners (ELs), or students who speak a language other than or in addition to English in the home and who are receiving school-based English language services because they have not passed an English language proficiency exam, (b) Recently-exited ELs (R-ELs), or students who speak a language other than or in addition to English in the home but have recently (within the past 2 years) passed an English language proficiency exam and have thus been recently exited from EL services, and (c) EL or REL students who have been identified as having a disability.

This study used data from a larger quasi-experimental study of linguistically diverse learners (LDLs) in fourth grade participating in a language-based reading intervention in order to test whether EF levels moderated gains in RC, ultimately addressing a need in the research base. Using path analysis in SEM, both EF and RC were measured by latent constructs and decoding and linguistic awareness were accounted for in the model. The preponderance of studies to date on EF and RC have included only one or two measures of EF and RC, but not latent constructs. The inclusion of latent constructs in this study increases the generalizability of the findings because the multiple domains of EF and RC, respectively, are captured through the multiple measures used to assess them. The present study further expands on previous research by accounting for students' assignment to the intervention condition and tested whether EF levels moderated the changes in students' demonstration of RC from the pre-program to post-program assessment to expand the model. The purpose of the present study is to understand whether EF moderates the effect of a language-based reading intervention program for LDL learners in fourth grade.

Theoretical Background

Reading Comprehension Interventions. Skillful RC is accepted as the product of linguistic comprehension and decoding according to the Simple View of Reading (SVR; Gough & Tunmer, 1986). The SVR model has served as the basis for a variety of reading interventions. For instance, Rashotte et al., (2001), Case et al. (2014), Shaywitz et al. (2004), and Torgesen et al. (2001) describe interventions with struggling readers across elementary school that have been successful in improving decoding, linguistic comprehension, and/or reading outcomes. However, students have not universally responded to such interventions (e.g., Greulich et al., 2014; Ritchey et al., 2012), and those responses (either by demonstrating gains in reading or not) are sometimes linked to ratings of attention (Al Otaiba & Fuchs, 2006; Greulich et al., 2014; Torgesen et al., 2001). Although they represent slightly different concepts attention ratings have often overlapped with operational definitions and assessments of EF.

Much recent reading intervention research with English monolingual students has been conducted under the principles of tiered instruction. These models of assignment to instruction are built on the principle that students are moved through successively intensive tiers of literacy interventions if they do not respond (i.e., demonstrate gains), given the present level of literacy instruction. The Response-to-Intervention (RTI) model's validity and reliability are well documented for students with reading difficulties (Justice, 2006), and the majority of the United States has developed explicit guidelines for the use of RTI to prevent and remediate reading difficulties (Hauerwas, Brown, & Scott, 2013). More broadly, teachers make decisions about students' reading group assignments and time spent within specialized intervention based on their performance throughout the school year. If students are identified as failing to keep up with their peers, data-driven instructional principles would expect students to be provided with increasingly intensive levels of literacy instruction. Students are regularly assessed at the beginning, middle, and end of the year on their reading skills to determine whether they are making progress towards benchmarks. Given that EF has been shown to predict RC outcomes, when students do not show equivalent progress compared to their peers, it makes sense to consider whether cognitive factors such as EF might be involved in reading difficulty.

Executive Function. EF is an umbrella term for a multifaceted cognitive construct used in goal-directed behavior (Miyake et al., 2000). As evidenced in several recent syntheses of the literature (see Chapter 2; Jacob & Parkinson, 2015) as well as empirical studies (see Chapter 3), EF measures account for significant and unique variance in the prediction of RC. The components of EF most frequently found as significant in the prediction of RC are *working memory*, *shifting*, and *inhibition*. Working memory is both the ability to store and manipulate information in the mind (Baddeley & Hitch, 1974; Cowan, 1999; Engle, 2001). Syntheses of the literature (see Chapter 2) and empirical studies (see Chapter 3) suggest that working memory plays an important role in EF (Christopher et al., 2012; Gerst et al., 2015; Jacobson et al., 2016; Nouwens et al., 2016; Sesma et al., 2009). Inhibition is the ability of the mind to override a reflexive response (Miyake et al., 2000). According to a synthesis of the literature on the concurrent validity of EF as a predictor of

RC for students in upper elementary school (see Chapter 2), inhibition also plays a significant role in the prediction of RC in some studies (Cutting et al., 2009; Kieffer et al., 2013). Finally, shifting is the ability of the mind to switch or change between tasks or mental sets (Baggetta & Alexander, 2016) and was also found as crucial to the prediction of RC in two out of seven studies (Kieffer et al., 2013; Nouwens et al., 2016) reviewed in a recent synthesis (see Chapter 2).

Purpose

The present study extends the current literature base on EF's known relationship with RC. The following research questions guided this study:

1. Does latent EF, as measured by working memory, inhibition, and shifting, moderate the effect of a supplemental reading intervention on LDL students' latent RC, controlling for linguistic comprehension and decoding?
2. Does the moderation effect differ for LDL students at different levels of English (EL, R-EL)?
3. Does the moderation effect differ for LDL students by disability status (IEP, non-IEP)?

Hypotheses

Students' latent EF is hypothesized to moderate the effect of the CLAVES reading intervention. Furthermore, it is hypothesized that the moderation effect will differ for students at different levels of English proficiency and disability status. The theoretical path model tested under the current study is presented in Figure 4.1. This model is an extension of the theoretical relationship between EF and RC presented in the empirical study conducted in Chapter 3. In the present study, the indirect contribution of latent EF through decoding and linguistic comprehension is augmented to include the assignment to the CLAVES intervention as a contextual variable.

Method

Design

The study used data collected by the CLAVES Project, which stands for *Comprehension, Linguistic Awareness, and Vocabulary for ELs* (Proctor, Silverman, & Harring, 2013). The CLAVES Project evaluated the feasibility and efficacy of a supplemental reading intervention for upper elementary school students. Data for the present study were collected at one of the two sites for the project, the one located in the Mid-Atlantic region. At this site, 128 fourth-grade students grouped within 10 classrooms across 4 schools participated in the research project. Half of the students were assigned to an intervention condition and half to a control or "business as usual" condition. Although the assignment was not fully randomized within each classroom due to the limited sample size and existing groupings of the students within each school, the two groups (intervention and control) were balanced by reading level and World Class Instructional Design and Assessment (WIDA) Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS) level prior to implementation (WIDA Consortium, 2010). The sample includes both recently reclassified English Learners (R-ELs) and English Learners (ELs). Assessment data was collected in the fall (pre-test) and late spring (post-test). Additional data sources, such as student background demographics, were obtained through school reports. Fidelity of implementation was assessed according to a standardized protocol developed by the research team (Proctor, Silverman, Harring, Jones, & Hartranft, in review).

Data were analyzed using a structural equation model approach (Bowen & Guo, 2011; Hancock & Mueller, 2013; Kline, 2015).

Sample

The CLAVES study was conducted in four public elementary schools that serve a predominantly Hispanic/Latino population of students and their families. Across these four schools, about 90% of students were eligible for free or reduced lunch, which is an indicator of lower socioeconomic status. According to public records, approximately 57% of the students across all four schools were identified as ELs and were currently receiving supplemental English as a second or other language (ESOL) services. About 7% of the students were identified as receiving special education services through an Individualized Education Plan (IEP). See Table 4.1 for a breakdown of participants by condition, English learner designated services, and special education services. The sample size for the present study includes R-ELs (n =67), ELs (n = 53), and students with disabilities (n = 9).

Demographics and Background Information. Student background data was obtained from school records, such as race/ethnicity, gender, and EL status, IEP status, and free and reduced meal status. The research team distributed a supplemental home language survey in order to understand students' language use in the home and data were used for contextual purposes.

Language status. School records of students' home language status was compiled based on school records. In this district, all students are subject to an assessment of English language proficiency at the time of their enrollment if parents indicate a language other than English is spoken in the home. Students are assessed by the district the World Class Instructional Design and Assessment (WIDA) Assessing Comprehension and

Communication in English State-to-State for English Language Learners ACCESS (ACCESS; WIDA Consortium, 2010). Students are scored on four domains: listening, reading, speaking, and writing, with individual scores for each domain ranging from 1.0 to 6.0. Scores of 1.0 are termed "Entering" and represent the lowest level of performance. Scores of 6.0 are termed "Reaching" and represent the highest level of performance. In the district for the present study, students were exited from the EL program, and considered "re-classified English learners" (R-ELs) when they reached an average score of 5.0 or greater across all four domains. Recruitment for the present study was targeted at students from WIDA levels 3.0 or greater or students who had recently exited within the past two years. Targeting this population of students was guided by research that has shown ELs and RELs often need continued support for academic language after exiting EL services. Language status was represented in the dataset as 1 for currently receiving EL services and 0 for recently exited students not receiving EL services.

Disability status. School records of students' disability status were obtained for the project. For use in the present study, categorical variables were created to indicate whether students have an Individualized Education Plan (IEP; 1) or not (0).

Socioeconomic status. Student records of eligibility for free and reduced meal status (FRP) were obtained for the project as well. FRP provides a rough estimation of household income and is used as a proxy for socioeconomic status in education literature. This variable was only used to contextualize the sample and was not used in the quantitative analysis since the majority of the students (over 90%) qualify for FRP.

Procedures

Lesson Implementation. Teachers and specialists (n = 10) who volunteered to participate in the study attended a one-day intensive training in the fall of the academic year and agreed to implement the CLAVES program that year. The full program consisted of 36 lessons designed to be taught in small-groups approximately 3 days per week for a duration of approximately 30 minutes each. The lessons were divided into three thematically-based instructional units that are aligned with the regularly scheduled English Language Arts curriculum in the school district. Due to time constraints, the majority of the teachers taught only two out of the three units during that academic year, with an average of approximately 30 lessons, including introduction and closing. Each unit contained guided reading of two texts, supporting informational videos, language instruction in semantics, morphology, and syntax, and discussions about central issues in the text or dialogic reasoning. Semantics instruction consisted of discussion of the meaning of target vocabulary words, related words, multiple meanings, cognates, and contextual use in the text and at home. Morphology instruction consisted of practice and application of affixes using target vocabulary from the semantics instruction. Syntax instruction consisted of analysis and discussion of sentence structure within the text and in related activities. Dialogic reasoning consisted of discussion about one or more central ideas from the text with support from textual and personal evidence. Reading comprehension instruction consisted of guided application of reading strategies, such as summarizing, clarifying, and making inferences based on the texts and supporting materials.

The nature and type of language instruction for students in the *business as usual* condition was determined through teacher interviews, surveys, and curriculum review to contextualize the regular reading instruction. In general, teachers reported the CLAVES program provided more language-based instruction while the business as usual program provided more fluency and decoding based instruction.

Fidelity of Implementation. Observations of teachers and fidelity of implementation data collection took place throughout the implementation of the project. Each teacher was observed at least two times during the year. Lessons were video and audio-taped for review and coding by the research team. The fidelity rating system was built around the lesson plans provided to the teachers. For each step in the lesson plan, teachers received a score of 2, 1, or 0. Teachers earned a 2 if they implemented 2 or more of the items listed in the lesson plan for that particular step, a 1 if they implemented 1 of the items listed in the lesson plan for that particular step, and a 0 if they did not implement any of the items listed in the lesson plan for that particular step. Fidelity of implementation was rated by two trained research assistants with interrater reliability above .90 (Cohen's Kappa) for 10 percent of the total observations. Fidelity scores ranged from 75-100% implementation of target lesson components across participating teachers at the Mid-Atlantic site. The average percent completed across the sample was 87%.

Data Collection. Assessment data collection took place in the fall (time 1) and spring (time 2) of one academic year. Assessments were administered in a group setting unless otherwise described below. Training was provided to test administrators to ensure fidelity and reliability of implementation. Scoring was done by trained research assistants and all measures were single and double scored for accuracy. Research assistants met regularly and followed rubrics to ensure reliability in scoring.

Measures

A combination of latent and observed variables were used in the present study. All assessment measures have been normed on fourth grade students and LDL students are represented in most of the normative samples. Latent constructs were preferable for analysis, compared to observed variables, because they reduce the error associated with each individual. This allows for statistical inferences to be made about the broader construct, in this case latent RC, latent EF, and latent linguistic comprehension, rather than focus on differences between each measure (Bollen, 2002).

Latent Reading Comprehension. Three reading comprehension measures were used to form a latent construct of RC. The validity of these measures as a latent construct of RC has been established by previous correlational and longitudinal research (Silverman et al., 2015; Proctor et al., 2012). More broadly, there a multitude of studies have found latent constructs of RC adequately represent the underlying theoretical conceptualization of RC (Fletcher, 2006; Mehta, Foorman, Branum-Martin, & Taylor, 2005; Vellutino, Tunmer, Jaccard, & Chen, 2007). In the present study, all three measures were used to form a latent construct of RC. Latent RC assessments were administered in the fall (time 1) and spring (time 2) of the year, coinciding with the start and end of the quasi-experimental trial.

MAZE. The Achievement Improvement Monitoring System (AIMSweb) MAZE assessment is an established measure of silent reading fluency and comprehension (Pearson, 2014). Students read a passage silently for three minutes and identify the missing words in the sentences from a field of three. Students receive one point for each correct word identified, with a maximum of 46 points. Test makers report split-half reliability of .95 and alternate form reliability of .95 for fourth grade students (Pearson, 2014).

***Test of Silent Reading Efficiency and Comprehension*.** The Test of Silent Reading Efficiency and Comprehension (TOSREC) was also used a measure of silent reading fluency and comprehension (Wagner et al., 2010). Students silently read short sentences containing either true or false information for three minutes and indicate whether each sentence is true or false. Students receive one point for each sentence correctly marked as true or false, yet lose one point for each sentence incorrectly marked. Raw scores are calculated by summing the total correct answers minus the incorrect answers. Test makers report alternate form reliability of .86 for fourth grade (Wagner et al., 2010).

***Gates MacGinitie Reading Comprehension*.** The Gates MacGinitie Reading Test (GMRT) Reading Comprehension subtest was used as an extended measure of passage RC (MacGinitie et al., 2002). Students read narrative and informational passages silently and answer literal and inferential comprehension questions about each passage. There are a total of 48 questions with four to six questions per passage. Students receive one point for each question answered correctly for a maximum of 48 points. Test makers report test-retest reliability of .81 for fourth grade (MacGinitie et al., 2002).

Linguistic Comprehension. Measures of linguistic comprehension were administered in the fall and spring of the project. In the fall, three measures were administered and a latent construct of linguistic comprehension was formed by those measures: an expressive vocabulary measure, a derivational morphology measure, and a grammaticality judgment measure. In the spring, one measure was administered that contained eight sub-sections including vocabulary, morphology, syntax, as well as themes, organization, and opinion within texts. Each linguistic comprehension measure is described below.

Expressive Vocabulary. The Woodcock Muñoz Language Survey –Revised (WMLS-R) Picture Vocabulary subtest is an individually administered measure of expressive vocabulary (Woodcock, 2004). In this assessment students are shown pictures of objects and asked to orally name them. The early items are highly frequent household objects, but the items become increasingly difficult with each page. Administration discontinues after six consecutive items on a page are incorrect. Students earn one point for each correct answer. Raw scores were used in analyses. Test makers report test-retest reliability of .88-.92. This measure was administered one-on-one and in the fall only.

Morphology. The Extract the Base (ETB) measure was used as a measure of morphology and administered in a group setting (Goodwin et al., 2012). In this assessment students are given cloze sentences and a target word. Students are prompted to 'extract the base' or find the correct morphological base from the target word in order to complete the sentence. The test is untimed and consists of 28 items, each of which are scored on a 0 – 2 scale (0 = incorrect, 1 = partially correct, 2 = fully correct). Raw scores were used in analyses. Test makers report internal reliability of 0.86. This measure was administered in the fall only.

Syntax. The Comprehensive Assessment of Spoken Language (CASL) Grammaticality Judgment subtest was used as a measure of syntax (Carrow-Woolfolk, 1999). The measure is administered individually. For this measure, the examiner reads a series of increasingly complex sentences and prompts the student to identify: (a) whether the sentence is correct and (b) if it is incorrect, how to change the sentence to fit the standards of English grammar. Some of the sentences are syntactically correct while others contain only one error in grammar. Students receive one point for correctly identifying whether the sentence is syntactically correct or not and an additional point for changing one word in the sentence to make it grammatically correct. The test is untimed and raw scores were used in analyses. Test makers report internal reliability of 0.88. This measure was administered one-on-one and in the fall only.

Core Academic Language Skills. The Core Academic Language Skills (CALS; Uccelli et al., 2015) measure was given in the spring timepoint only. This assessment contains multiple questions across eight subtests related to academic language. These subtests are titled connecting ideas, tracking themes, organizing texts, breaking words, comprehending sentences, identifying definition, epistemic stance, and metalanguage. Test makers report split-half reliability of .90 and a coefficient a of .93 (Uccelli et al., 2015).

Decoding. The Test of Silent Word Reading Fluency - Second Edition (TOSWRF-2) was used as a measure of decoding in the current study (Mather et al., 2014). This assessment provides an indication of students' word-level reading accuracy and fluency. Students silently read a list of 220 words and mark the beginning and end of each word. Earlier words are shorter and more high frequency words, while later words are longer and less familiar. Test makers report test-retest reliability of .93 (Mather et al., 2014). This assessment was given in the spring only.

Latent Executive Function. A latent construct of EF was formed from three measures designed to tap working memory, shifting, and inhibition. The use of a latent construct for EF is supported by previous literature (Miyake et al., 2000; Diamond, 2013), although other studies have not combined individual EF tests into latent measures (e.g., Cutting et al., 2009, Gerst et al., 2015, Nouwens et al., 2016). Academic evidence for the unity of EF to form a construct are reinforced by the results of the current study, as discussed below in the analytic plan and results. All of the EF measures were administered individually in the spring of the academic year.

Working memory. The Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) consists of two subtests: Digit Span Forwards (DSF) and Digit Span Backwards (DSB; Wechsler, 2003). First, students listen while an examiner reads a list of numbers. Then students are prompted to repeat the numbers aloud. This is known as the DSF condition and tests rote recall. Up to 9 numbers are presented in a row. Next, students listen again while an examiner reads a list of numbers. Then, students are prompted to repeat the sequence of numbers backwards. This is known as the DSB task and tests how many items a student can hold and manipulate mentally. Up to 8 numbers are presented in a row. Students receive one point for each number recalled correctly. The total percentage correct on the DSB condition was used in analyses. Test makers report split-half Spearman Brown reliability of .87.

Inhibition. The Neuropsychological Developmental Assessment – Second Edition (NEPSY-II) Inhibition consists of four conditions that provide an overall measure of inhibition (Korkman, Kirk, & Kemp, 2007). In the first and third conditions, students are provided with arrays of shapes and told to name the shapes. In the second and fourth conditions, students are provided with the same arrays and told to name opposing shapes. The test is intended to measure of students' ability to stop an automatic response and provide an artificial response based on a rule. The total number of correct items was used in analyses. Test makers report reliability of .73-.90.

Shifting. The Delis-Kaplan Executive Function System (D-KEFS) Trail Making Test (TMT) consists of five successive conditions, which in sum are intended to measure mental set-shifting. In the earlier conditions, students connect typical sequences of numbers (1, 2, 3 and so on) and letters (A, B, C, and so on). In the fourth condition, *Letter-Number Switching,* students are instructed to draw lines to connect an alternating sequence of numbers and letters: A, 1, B, 2, C, 3, and so on. Students have a maximum of 4 minutes to complete the task. The total time to complete the task used in analyses. Test makers report test-retest reliability of .89.

Data Analysis

Analytic Plan. The analytic plan for the current study adhered to the established procedures for conducting quantitative analyses. Prior to administration, the study design was examined for threats to validity. During and after the study, missing data were identified, and reliability of administration was established. Descriptive statistics were evaluated to ensure normality of distribution. Structural equation modeling was conducted in successive order, with model testing, confirmatory factor analysis, path analysis and model fit indices. Additionally, measurement invariance was tested for measures administered at two time points. SPSS AMOS Version 25 was used to conduct SEM (Bowen & Guo, 2011; Hancock & Mueller, 2013; Kline, 2015).

Missing Data. The dataset was evaluated for missing data in order to determine if bias was present. Out of 128 participants in the larger study, 8 were missing one or more assessments from the full battery. Removing those participants was deemed the most appropriate solution, as SEM requires complete datasets on all cases. In total, 6% of the intended participants were missing data, however the data appeared to be missing at random.

Statistical Analysis.

Descriptive statistics. Descriptive statistics for the variables of interest were calculated, including means, standard deviations, and measures of normality. See Table 4.2

for means and standard deviations. Examinations of normality were conducted to explore issues of skew, kurtosis, univariate and multivariate outliers, and normality of residuals to determine whether the data met the assumptions of normality. For the present sample, the range in skew was from -0.82 and 1.86 which is within the guidelines for normal distribution. The range in kurtosis was from -1.10 and 5.85. The kurtosis value of 5.597 for the MAZE post-test raw score was outside of the acceptable range. Further investigation revealed two students had scores outside of the predicted range. However, the maximum Cook's distance statistic was 0.41, which is within the acceptable range. Therefore, the data was not altered from its raw form.

Multicollinearity. Multicollinearity means there are high levels of correlation between two or more predictor variables. Multicollinearity statistics should be examined in all studies in which linear regression or SEM are used. For the present study, multicollinearity was evaluated by reviewing the Variance Inflation Factors (VIF) for each of the predictor variables. VIF values above 10 and tolerance values below 0.2 indicate multicollinearity must be addressed in the model (Menard, 1995). In the current dataset, VIF values were between 1.073 and 1.884 and tolerance statistics were from 0.542 to 0.932. These values are considered acceptable.

Structural Equation Modeling. Structural equation modeling (SEM) was used to answer the research questions because it offers several advantages compared to multiple regression analysis. The components of the SEM model were built using the structural model described in Chapter 3. Namely, a path analytic model was used to examine the indirect effects of latent executive function, formed by working memory, inhibition, and shifting, on latent variable RC, controlling for linguistic comprehension and decoding. Chi-square difference tests were run to determine differences by subgroups on the paths of interest. The present study added pre-test variables to account for students' initial performance in RC and linguistic comprehension at the beginning of the year, as well as a categorical variable representing assignment to treatment. Figure 4.1 provides the structural model under investigation in the present study.

Model testing. Several path analytic models were tested using structural equation modeling and confirmatory factor analysis within SPSS AMOS (CFA; Browne & Cudeck, 1993; Kline, 2015). See Figure 4.1. Within SEM nomenclature, circles represent latent variables and rectangles represent measured variables. "Latent Executive Function" represents the latent variable of EF, formed by three measured variables, which are represented by the boxes "Working Memory," "Inhibition," and "Shifting." "Latent Reading Comprehension Pretest" and "Latent Reading Comprehension Posttest" represent the latent variables of RC collected in the spring and fall. Each Latent Reading Comprehension variable is formed by three measured variables, which are represented by the boxes "MAZE," "GMRT", and "TOSREC." "Latent Linguistic Comprehension Pretest" represents the latent variable of Linguistic Comprehension collected in fall and formed by three measured variables, which are represented by the boxes "Syntax," "Vocabulary", and "Morphology." "Linguistic Comprehension Posttest" represents an observed variable of linguistic comprehension collected in the spring and formed by a single assessment that encompasses items related to semantics, morphology, and syntax. The rectangle "Decoding" is a measured variable formed by a single assessment. Single-headed arrows indicate regression. Excluded from this structural model are the double-headed arrows indicating covariances among the predictors and error terms associated with each observed variable.

The statistical significance of the direct and indirect paths was obtained by evaluating the path weights. Additionally, model fit was evaluated by examining several fit indices: chi-square test (x^2), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), goodness of fit index (GFI), and Akaike information criterion (AIC). A GFI above 0.90 suggests that the proposed model fits the data well (Kline, 2015). Guidelines for acceptable fit for each index are chi-square tests above .05, CFI above 0.90, TLI above 0.90, and RMSEAs below 0.10 (Hu & Bentler, 1999; Kline, 2015).

Confirmatory factor analysis was used to evaluate longitudinal measurement invariance for latent RC, which was administered in the fall and spring. In longitudinal designs, measurement invariance is important to consider because it shows whether the relationships between the latent constructs are stable across time points. See Table 4.10 for tests of measurement invariance by time for latent RC at pretest and posttest.

Results

The purpose of the current study was to explore whether EF moderates the effect of the assignment to a supplemental reading intervention for linguistically diverse learners. The study also investigated whether the moderation effect differed according to students' language and/or disability status.

Means and standard deviations for the observed measurement variables, disaggregated by condition, language status, and disability status are provided in Table 4.2. Significant differences among subgroups based on condition, language status and disability status were present based on one-way ANOVAs. No significant differences between assessments were found for students in the intervention versus the control condition for the fall (pre-test) variables. For EL students in the intervention condition, significant differences were found compared to REL students in the intervention condition on two measures of reading comprehension at pretest: GMRT ($F[1,55] = 3.89$, $p = .054$) and TOSREC ($F[1,55] = 4.37$, $p = .041$), one measure of linguistic comprehension at pretest: morphology ($F[1,55] = 11.74$, $p = .001$), one measure of reading comprehension at posttest: MAZE ($F[1,55] = 4.03$, $p = .050$), and two measures of executive function: working memory ($F[1,55] = 7.44$, $p = .009$) and inhibition ($F[1,55] = 8.30$, $p = .006$). For EL students in the control condition, significant differences were found compared to REL students in the control condition on one measure of linguistic comprehension at pretest: morphology ($F[1,61] = 5.96$, $p = .018$), all three measures of reading comprehension at posttest: GMRT ($F[1,61] = 4.44$, $p = .039$), MAZE ($F[1,61] = 14.46$, $p = .001$), TOSREC ($F[1,61] = 4.36$, $p = .041$), and one measure of linguistic comprehension at posttest ($F[1,61] = 10.63$, $p = .002$), and one measure of executive function: inhibition ($F[1,61] = 3.95$, $p = .051$).

Students with IEPs performed significantly below their peers on one measure of reading comprehension at pretest: GMRT ($F[1,118] = 10.99$, $p = .001$), two measures of linguistic comprehension at pretest: Syntax ($F[1,118] = 4.49$, $p = .036$), vocabulary ($F[1,118] = 5.13$, $p = .025$), decoding ($F[1,118] = 9.14$, $p = .003$), and two measures of executive function: working memory ($F[1,118] = 3.64$, $p = .059$) and inhibition ($F[1,118] = 4.602$, $p = .034$).

Correlations among observed assessment variables can be found in Table 4.3. All of the pre and post-test reading comprehension measures showed moderate correlations with each other ($r = .408$ to $.695$, $p = .01$). All of the pre and post-test linguistic comprehension measures showed small to moderate correlations with each other ($r = .277$ to $.444$, $p = .01$).

Two of the executive function measures showed moderate correlations with each other: working memory and inhibition ($r = .276$, $p = .01$). The reading comprehension and linguistic comprehension variables all showed small to moderate correlations with each other across pre and post-test ($r = .255$ to $.620$, $p = .01$). Decoding showed small to moderate correlations with most reading comprehension, linguistic comprehension, and executive function variables ($r = .251$ to $.641$, $p = .01$) with the exceptions of pre-test vocabulary to decoding and working memory to decoding, which were non-significant, and shifting to decoding, which was significant at the .05 level ($r = -.189$). The 3 executive function variables showed small correlations with 13 out of the 33 total assessment pairs. Working memory and inhibition showed small correlation with reading comprehension at pretest ($r =.183$ to $.234$, $p = .05$), and slightly stronger, but still small correlations at posttest ($r = .202$ to $.273$, $p = .05$ to $.01$). Working memory also showed moderate correlations with morphology at pretest ($r = .260$, $p =.01$) and linguistic comprehension at posttest ($r = 2.41$, $p = .01$). Inhibition showed small correlations with syntax at pretest ($r = .229$, $p =.05$) and linguistic comprehension at posttest ($r = .251$, $p = .01$). Shifting only significantly correlated with one other measure, which was with decoding ($r = -.189$, $p = .05$).

Examining patterns among these associations, it appears all measures of reading comprehension showed moderate correlations with each other, but not perfect correlations. This supports the use of a latent construct for these measures. Linguistic comprehension similarly showed moderate correlations with each other, supporting the use of a latent construct for the pretest measures. Executive functions showed one instance of a small, but significant correlation out of three possible correlations. Although higher correlations between shifting and the other two constructs were expected based on correlations reported in other similar studies, this may be due to the choice of the shifting measure rather than differences in the theoretical domains.

The following research questions guided this study: (1) Does latent EF, as measured by working memory, inhibition, and shifting, moderate the effect of a supplemental reading intervention on LDL students' latent RC, controlling for linguistic comprehension and decoding? (2) Does the moderation effect differ for LDL students at different levels of English (EL, R-EL)? (3) Does the moderation effect differ for LDL students by disability status (IEP, non-IEP)?

Research Question 1

In order to properly conduct this analysis, the first step was to verify the underlying relationships between the variables in a measurement model using the multi-group analysis function in AMOS to evaluate differences by condition. The baseline model included all assessment variables and the condition variable, except for the EF variables and language status variable. See Figure 4.1, Figure 4.2 and Table 4.4. This model demonstrated poor model fit indices: x^2 (85, N = 120) = 240.476, CFI = .710, TLI = .625, RMSEA = .125. All path regression weights were significant at the .01 level. Additionally, the chi-square difference test for multiple groups showed significant differences for the control condition compared to the intervention condition (x^2 [10, N = 120) = 19.996, p = .029). Principal component analysis coefficients for latent RC is displayed in Table 4.8. Tests of measurement invariance by time for latent RC at pretest and posttest are displayed in Table 4.10. Despite the poor model fit indices, this model represents the theoretical relations among the variables (i.e., pre-test reading comprehension would predict post-test reading comprehension, and pre-test linguistic comprehension would predict post-test linguistic comprehension, which would in turn predict post-test reading comprehension). Given the relatively small sample size for the number of parameters to be estimated, it is hypothesized that a larger sample would yield improved fit indices. Therefore, this model was used as the baseline model for all subsequent analyses.

The first research question investigated whether latent EF, as measured by working memory, inhibition, and shifting, moderated the effect of a supplemental reading intervention on LDL students' latent RC, controlling for linguistic comprehension and decoding. The model itself was not considered ideal, having the following poor fit indices x^2 (151, N=120), 482.465, p = .001, CFI = .403, TLI = .280, and RMSEA = .136. See Figure 4.3, Figure 4.4, and Table 4.5. Principal component analysis coefficients for latent EF is displayed in Table 4.9. A chi-squared difference test of the nested model comparisons found significant differences between the control and intervention groups for the overall model (x^2 (14, N=120), 26.366, p = .023), which included latent EF. Examining the path regression weights, differences were seen among the executive function variables across the control and intervention conditions. For the control group, the paths from latent EF to decoding and post-test linguistic comprehension were significant (.808, p = .02; .410, p = .05), meaning that increase in latent EF was associated with increase in decoding and linguistic comprehension. However, for the intervention group, the paths from latent EF to decoding and linguistic comprehension were not significant (.386, p = .071; .720, p = .066). However, with a larger sample size, these observed directional relations may have been statistically significant. This means for students in the intervention, latent EF did not significantly predict decoding performance nor linguistic comprehension. The remainder of the pathways for both conditions were all significant.

Research Question 2

For research question 2, model 1 was used to test for differences between EL and REL groups across control and intervention conditions using AMOS multi-group analysis. See Table 4.6 and Figures 4.5, 4.6, 4.7, and 4.8. The model itself was considered inadequate having the following poor model fit indices: x^2 (307, N=120), 538.509, p = .001, CFI = .592, TLI = .516, and RMSEA = .081. A chi-squared difference test of the nested model comparisons found significant differences between the EL and REL groups across control and intervention conditions for the overall model (x^2 (36, N=120), 56.844, p = .015). This means there were significant differences in the path regression weights across the four groups (EL*Intervention, REL*Intervention, EL*Control, REL*Control). Of particular interest for the current study were the paths from latent EF to decoding and linguistic comprehension by each subgroup. For students in the control group, the EF pathways to decoding and post linguistic comprehension were not significant for either EL students (p = .39; p = .20) or REL students (p = .16; p = .09). For students in the intervention group, in contrast, the pathway from latent EF to decoding was significant for EL students (.96, p = .04), but not significant for REL students (.36, p = .09). These contrasting patterns of significance are interpreted further in the discussion.

Research Question 3

For research question 3, model 2 was intended to be used to test for differences between EL and REL groups across control and intervention conditions using AMOS multi-group analysis. However, given the small number of students with disabilities in each condition subgroup (4 and 5), SEM was determined not to be appropriate. Instead, a multivariate analysis of between-subjects effects was conducted using condition and disability status as fixed factors, decoding and linguistic comprehension as dependent variables, and the observed executive function variables as independent variables. See Table 4.7. This analysis showed a main effect of disability status on decoding (f = 5.051, p = .027), but non-significant effects on linguistic comprehension. No main effect for condition and no interaction effects were detected (p = .078). Examining the means and standard deviations by condition and disability status reveal similar findings. Students with disabilities scored significantly below their peers, irrespective of condition, on decoding (f = 9.142, p = .003). Students with disabilities also scored significantly below their peers on inhibition (f = 4.602, p = .034), and marginally on working memory (f = 3.641, p = .059). These results suggest the relationship between latent EF and decoding and latent EF and linguistic comprehension may differ significantly based on disability status for students in the sample, although drawing statistical inferences must be done with caution given the small sample size of this subgroup.

Discussion

The primary question for this study was to investigate whether latent EF moderates the effect of a supplemental reading intervention on LDL students' latent RC, controlling for linguistic comprehension and decoding. This study found that EF does moderates the effect of the assignment to a supplemental reading intervention for linguistically diverse learners, but only through the path from latent EF to decoding and linguistic comprehension. Additional analyses revealed direct effects of moderation on latent RC were not detected. The ancillary research questions for this study were to explore whether the moderation effect differed for students at different levels of English and for students with disabilities. The study found limited evidence the relationship between latent EF and decoding and linguistic comprehension differed by students' assignment to treatment condition, language status, and disability status.

The present study's findings that decoding and linguistic comprehension are significant and strong predictors of reading comprehension achievement is consistent with literature to date. The current study also extended this research base by adding measures of executive function as direct contributors to linguistic comprehension and decoding, as well as indirect contributions, or mediators, to reading comprehension. Adding latent executive function within the context of the treatment condition improved some aspects of the model fit, as evidenced by the reduction in the RMSEA. This means the expanded model accounted for more of the variance in the outcome variable, namely, latent RC, by adding significant explanatory variables, namely, latent EF.

The current study also found some evidence that the role of EF as a predictor of decoding and linguistic comprehension, and indirectly, reading comprehension, varied by students' assignment to treatment condition, language status, and disability status. Specifically, students in the control group demonstrated a significant effect of EF on decoding and linguistic comprehension, whereas students in the intervention group did not demonstrate a significant effect of EF on decoding and linguistic comprehension. However, these results do not indicate the categorical variables are not necessarily causing the particular relationship between EF and decoding and linguistic comprehension. SEM methodology often evokes causal language, when in this case the relationships between the variables of interest are merely correlational.

Returning to the purpose of the larger reading intervention study, these effects are not surprising. The intervention study did not emphasize of decoding in the lessons, as the focus was intentionally on linguistic awareness and comprehension. In fact, the curriculum contained supports for decoding so that the academic tasks were concentrated squarely on linguistic comprehension. In this way, it may not have mattered whether students had high or low decoding or higher or low EF in the study, because the study intentionally circumvented these pre-requisite skills through scaffolding and instructional modifications. Perhaps if the intervention was more decoding focused or required more independent reading, EF would have moderated effects. More research is needed on the role of EF in moderating the effects of different kinds of interventions with different emphases. Nevertheless, students in the intervention group outperformed the control group on measures of semantics and linguistic awareness (Silverman, Proctor, Harring, Jones, & Hartranft, under review), so the intended concentration of the language-based intervention was met.

Among EL students and students with disabilities, some differences were observed compared to their peers. For EL students in the intervention group, the pathway from EF to decoding was significant. This effect was not found for REL students in the intervention group. This difference may be related to the typical progression of reading development and the varying cognitive loads required as readers progress. When students are still emerging readers, there is a heavier focus on deciphering the code, followed by making meaning from the text. As students enter upper elementary school, their foundational reading skills are expected to be intact and they are taught to process and evaluate what the meaning of the text. For students who are learning English as a second language, that cross-over point may be occurring later than with typically developing English monolingual students. Thus, it is not surprising that EL students in the intervention were relying more heavily on EF and decoding skills compared to REL students, who had higher levels of English language proficiency.

Implications

The present empirical study is poised to contribute to the current literature because it investigates the potential for EF to moderate the effect of a reading intervention, which is a relationship that has not been studied through SEM models with latent constructs in LDL populations. Furthermore, the study tested whether the moderation effect differs depending on students' language status and/or disability status, both of which are populations that are understudied in the literature but at-risk for reading difficulty. Understanding more about how students' cognitive ability may contribute to their reading comprehension holds tremendous promise for the field.

Limitations

The present study was impeded by the restricted sample size, both overall given the complex models intended to be tested, as well as in particular for the questions regarding subgroup analysis. Reduced power in any sort of regression analysis leads to greater likelihoods of failing to detect an effect even if one is present. In the present study, the limitations of reduced power were extenuated by the relatively small size of the hypothesized effect of executive function on reading comprehension. Due to the small number of students with disabilities in the present study, it was not possible to conduct a full SEM regression with all of the variables of interest. Moreover, given the low prevalence of students with disabilities it difficult to generalize the results beyond the current study. The present study also relied on single measures of working, shifting, and inhibition. Decoding and executive function were assumed to represent stable constructs that would be unchanged throughout the course of an academic year, and therefore these assessments were only given once. However, it is unknown if a fully replicated fall and spring battery of assessments would have yielded change in the decoding and executive function scores.

Future Direction

The present study provides support for future research to investigate whether students' EF levels impact their growth in decoding, linguistic comprehension, and reading across the year. Large scale measurement studies of EF are sorely needed in order to understand how the domains of EF contribute to a unity construct and how EF may change over the course of an academic year and across multiple years. Future reading intervention studies should consider assessing students EF levels prior to their assignment to a reading intervention to ensure the program or programs chosen fully addresses the breadth of reading comprehension. Additional research is also needed to study how the profiles of students with from EL and REL backgrounds or with disabilities intertwine with their EF competencies.

Chapter 5: Conclusion

The central goal of this dissertation was to systematically identify and examine the measurable evidence for the contribution of EF to RC outcomes for LDLs in upper elementary school. The dissertation accomplished this aim over three separate, but related studies. A synthesis of the literature (Chapter 2) was conducted on the predictive validity for EF on RC outcomes for students in upper elementary school. An empirical study (Chapter 3) was conducted to extend the current literature on EF by testing whether a latent construct of EF, including working memory, shifting, and inhibition, made direct or indirect contributions to RC. In addition, the study explored whether that relationship differed for students from underrepresented populations, including students currently enrolled in, or recently exited from, English language learner programs, as well as students with disabilities. A second empirical study (Chapter 4) was conducted to explore whether students' latent EF moderated the effect of an intervention targeting RC and linguistic awareness. As a follow-up to that question, the study also explored whether the moderation effect varied by students' language status and disability status. The present chapter (Chapter 5) offers a discussion of the main findings across all three studies and suggests implications gleaned from this study about the relationship between EF and RC for students in upper elementary school.

Executive functioning is the multidimensional process by which children and adults integrate multiple skills in order to reach a goal. Working memory, shifting, and inhibition are the domains of EF are consistently referenced in the literature (Baggetta & Alexander, 2016; Jacob and Peterson, 2015; Miyake et al., 2000). *Working memory* is the ability to refresh, retain, and manipulate information in the working memory (Baddeley, 2000; Brookshire et al., 2004; Diamond, 2013). *Shifting* means the skill of transferring among mental processes with flexibility (Brookshire et al., 2004; Kieffer et al., 2013). *Inhibition* means the skill of preventing competing thoughts or impulses (Friedman et al., 2006). Among these three skills, working memory was most consistently found to be predictive of RC outcomes, according to my synthesis of the literature. Baggetta and Alexander (2016) found that inhibition was the most commonly described domain of EF. Finally, shifting has been implicated as predictive of RC in many studies of RC (Kieffer et al., 2013; Nouwens et al., 2016). Based on the breadth of these findings, the next step in the inquiry around EF was to explore whether these domains exerted the same degree of predictive validity when measured as a latent construct. Indeed, findings from my synthesis of the literature suggested that additional research is needed to understand how EF contributes to RC using latent constructs, instead of single measures. The operationalization of EF as a latent construct in the empirical studies I conducted offers evidence for the unity of the construct and is discussed further below.

Key Findings and Implications

Contribution of Latent EF. This dissertation found consistent evidence that EF makes a small, but significant, contribution to RC through the pathways of decoding and linguistic comprehension. In my empirical studies, I found some evidence that latent EF adds to the prediction accuracy of latent RC. When entered as a single predictor, latent executive function presented a small to moderate sized standardized regression coefficient (.40) on latent RC. I hypothesized that executive function would make both a direct and indirect contribution to the prediction of reading comprehension. However, when decoding and linguistic comprehension were entered into the models in both of my empirical studies, the contribution of EF to RC was indirect through decoding and linguistic comprehension. In some ways, this is consistent with Simple View of Reading (SVR; Gough & Tunmer, 1986) literature, which states that reading is the successful product of decoding and linguistic comprehension (Catts et al., 2006; Carver, 1998; Kendeou et al., 2009a; Verhoeven & Van Leeuwe, 2008). All of the studies reviewed in the synthesis and both the empirical studies found that either decoding and/or linguistic comprehension were significant predictors of RC achievement, and often both were implicated when included in the model (Christopher et al., 2012; Cutting et al., 2009; Gerst et al., 2015; Jacobson et al., 2016; Kieffer et al., 2013; Sesma et al., 2009). Given that the regression weights observed in my empirical studies were comparable to those found in studies from my synthesis, this dissertation offers support for the formation of a latent EF. The advantages of using a latent construct are that it is possible to discern the combined impact of multiple domains of EF on the RC, decoding, and linguistic comprehension variables of interest simultaneously.

Importance of Upper Elementary School. Upper elementary school is critical point in literacy and academic development. Throughout the studies presented in this dissertation, studies on EF were chosen for samples that specifically focused on upper elementary school, and fourth grade in particular. The fourth grade reading curriculum presents a major shift in its cognitive demands for students (Chall, 1983; Chall & Jacobs, 2003). Perhaps not coincidentally, many studies have focused on understanding how and why late emerging reading difficulties that appear in upper elementary school, despite earlier screenings (Catts et al., 2012; Compton et al., 2008; Etmanski et al., 2014; Leach et al., 2003). One competing hypothesis is that deficits may emerge as a result of the increased text complexity required in upper elementary school, though alternative hypotheses are also compelling. Given this critical juncture in development, the importance of EF may, too, re-emerge as students are coordinating and integrating new skills.

Intersections with Language Proficiency. English learners, and linguistically diverse learners more broadly, are often characterized as students who speak a language other than English in the home and may benefit from additional English language support to succeed in school. ELs represent the subset of LDLs who are identified by schools as requiring specialized English language supports. Re-classified English learners (RELs) are students who have recently advanced out of their schools' English language programs upon completion of English language proficiency exams. The purpose of including LDL populations in this manuscript is informed by multilingual research and the recognition of the assets multilingual students bring to school and their learning.

The present cohort of three studies found some evidence that latent EF functions differently for EL students compared to REL students. For instance, in the second empirical study, the path from EF to decoding was divergent for EL students compared to their REL peers within the context of the treatment assignment. Specifically, EF played a significant role in the performance on measures of decoding for EL students in the intervention condition, while the effect was absent for REL students in the intervention condition. However, in the first empirical study, there were no significant differences in the SEM paths from latent EF to decoding and linguistic comprehension for EL versus REL students, despite EL students performing significantly below their peers on observed working memory and inhibition in one-way ANOVA models. These mixed findings suggest EF skills may be important in some conditions as students are learning the code of a language, and also offers suggestions for where to begin interventions for students who are learning English. LDL populations have been understudied in the EF literature on students in upper elementary school (Gerst et al., 2015; Kieffer et al., 2013; Nouwens et al., 2016). Both of the empirical studies contribute to the current literature on LDL populations by investigating the potential for latent EF to influence the relationships between decoding, linguistic comprehension and reading comprehension.

Intersections with Disability Status. In my synthesis of the literature, I found limited research conducted with students with disabilities that investigated whether EF contributed to RC outcomes. Although struggling readers are often classified by researchers in EF research, students who are identified by their schools as having a disability. In the empirical studies I conducted, a secondary analysis of one-way ANOVAs revealed students with disabilities performed significantly below their peers on measures of working memory, inhibition, decoding, and reading comprehension. These findings echo previous studies of observed EF components and RC outcomes of students with reading difficulties (Cutting et al., 2009; Locascio et al., 2010). Even though these differences were not significant in the SEM models in the present empirical studies, the ability to detect such differences had they been present was limited by the small sample size in the studies. Other research on EF that includes students with disabilities similarly found this group performed lower on measures of decoding and RC, but the interaction term representing disability status was not a significant predictor of RC (Nouwens et al., 2016).

Intersections with Reading Interventions. This dissertation examined the impact of latent EF on RC at a single time point, as well as its role as a moderator for change in reading comprehension. The second empirical study investigated whether EF moderated the effect of a language-focused reading comprehension intervention. Although the fit indices for the second empirical study were not ideal, this study found effects for the role of latent EF as a moderator of latent RC through decoding and linguistic comprehension for students in the control condition. The pathway between latent EF and decoding differed for EL students in a language-based intervention, compared to REL students in the intervention and EL and REL students in the business as usual, or control condition. This dissertation makes important contributions to the field because there is a scarcity of research that examines how latent EF may add to the model of RC change throughout an intervention study.

Limitations and Future Directions

The synthesis of the literature and two empirical studies offered additional evidence on the link between latent EF and RC through decoding and linguistic comprehension. However, limitations present in each study are noted below. Several of these drawbacks could be feasibly addressed in future studies, and are presented alongside suggestions for inquiry moving forward.

Sample Size and Subgroups. Across studies, the small sample size for the number of parameters to be estimated reduced the analytic and inferential power of the models assessed. Additional research is needed that contains ample students with disabilities and/or students with limited English proficiency to understand the composition and effects of EF in understudied populations. Balancing the need for larger samples sizes with the resources required to conduct large-scale studies is difficult. In part, though, this could be addressed in recruitment stages of future projects, so that sufficient numbers of students in understudied populations are included in the study from the beginning. Additional solutions may be provided by linking concurrent investigations of RC and EF with understudied populations intentionally so that datasets can be shared and support multiple lines of research when possible.

Measurement of EF. Research into the unity of the construct of EF is ongoing, yet further research is needed to fully explore the measurement of EF and its constituent factors. This includes the need to understand how multiple measures of each EF domain are influenced by task-specific and individual student differences compared to elements that are central to the construct of EF. Directions for future research includes the careful review of measures selected, concurrent use of established measures alongside researcher-developed measures to gauge construct validity, and clear delineations in reports of how each construct is operationalized. Additionally, researchers should consider forming latent constructs when permissible and reasonable given the measures used, and then reporting those factor analytic findings for dissemination and critique from the field.

Longitudinal and Instructional Influences. The studies presented in this dissertation used correlational methods to detect relationships between latent EF and RC for LDLs in upper elementary school. Further research is needed to understand how EF may change across the year and over multiple years. Measuring the development of EF in students directly alongside their instructional contexts would allow for multiple strands of complementary research to be conducted simultaneously and could yield exponentially stronger results. Variations in instructional design that are sensitive to EF development may be associated with stronger gains in both EF and RC outcomes. This dissertation offers both novel and confirmatory findings for the field and suggests several directions for the continued exploration of EF.

References

Adlof, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43(4), 332-345.*

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities*, *39*(5), 414-431.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*, 20-29.

Alloway, T. P., Banner, G. E., & Smith, P. (2010). Working memory and cognitive styles in adolescents' attainment. *British Journal of Educational Psychology*, *80*(4), 567-581.

Altemeier, L. E., Abbott, R. D., & Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology, 30*(5), 588-606.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, *71*(3), 283-300.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind, Brain, and Education*, *10*, 10-33.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327-336.

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, *9*(2), 117-121.

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, *65*(4), 229.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, *70*(3), 636-644.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, *24*(4), 560–567.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, *11*(4), 485-508.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental science*, *7*(3), 325-339.

Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., & Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*(5), 757-766.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. C*hild Development, 78*(2), 647–663.

Bollen, K.A. (2002). Latent Variables in Psychology and the Social Sciences. Annual Review of Psychology, 53, 605-634.

Borella, E., Carretti, B., & Pelgrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, *43*(6)*,* 541–552.

Bowen, N. K., & Guo, S. (2011). *Structural equation modeling*. Oxford University Press.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., ... & Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of reading. *Reading and writing*, *29*(3), 435-451.

Brookshire, B., Levin, H. S., Song, J., & Zhang, L. (2004). Components of executive function in typically developing and head-injured children. *Developmental neuropsychology*, *25*(1-2), 61-83.

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). London: Sage Ltd.

Burr, E., Haas, E., & Ferriere, K. (2015). Identifying and Supporting English Learner Students with Learning Disabilities: Key Issues in the Literature and State Practice. REL 2015-086. *Regional Educational Laboratory West*.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental science*, *11*(2), 282-298.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19*(2), 246–251.

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Circle Pines, MN: AGS.

Cartwright, K. B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. *Early Education & Development*, *23*(1), 24-36.

Carver, R. P. (1998). Predicting reading level in grades 1 to 6 from listening level and decoding level: Testing theory relevant to the simple view of reading. *Reading and Writing*, *10*(2), 121-154.

Case, L., Speece, D., Silverman, R., Schatschneider, C., Montanaro, E., & Ritchey, K. (2014). Immediate and long-term effects of tier 2 reading instruction for first-grade students with a high probability of reading failure. *Journal of Research on Educational Effectiveness*, *7*(1), 28-53.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. Journal of Speech, Language, and Hearing Research, 49(2), 278-293.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, *104*(1), 166-181.

Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing: An Interdisciplinary Journal*, *28*(9), 1407-1425.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278-293.

Chall, J. S. (1983). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth-grade slump. *American educator*, *27*(1), 14-17.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., ... & Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: a latent variable analysis. *Journal of Experimental Psychology: General, 141*(3), 470-488.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences (3rd edition).* Mahwah, NJ: Lawrence Earlbaum Associates.

Cooper, H., & Hedges, L. V. (2009). Research synthesis as scientific process in H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis*. (pp 3-15). New York, NY: Russell Sage Foundation.

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, *18*(3), 329-337.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59-86.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control*, *20*, 506.

Cummins, J. (1979). Linguistic interdependence and the educational development of Bilingual children. *Review of Educational Research, 49*(2), 222-251.

Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of dyslexia*, *59*(1), 34-54.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277– 299.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-kaplan executive function scale.* San Antonio, TX: The Psychological Corporation.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Etmanskie, J. M., Partanen, M., & Siegel, L. S. (2016). A longitudinal examination of the persistence of late emerging reading disabilities. *Journal of learning disabilities*, *49*(1), 21-35.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

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

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of experimental psychology: General*, *133*(1), 101.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological science*, *17*(2), 172-179.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239-56.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of experimental child psychology*, *93*(3), 265-281.

Gerst, E. H., Cirino, P. T., Fletcher, J. M., & Yoshida, H. (2015). Cognitive and behavioral rating measures of executive function as predictors of academic outcomes in children. *Child Neuropsychology,* 1-27.

Goodwin, A. P., Huggins, A. C., Carlo, M., Malabonga, V., Kenyon, D., Louguit, M., & August, D. (2012). Development and validation of extract the base: an English derivational morphology test for third through fifth grade monolingual students and Spanish-speaking English language learners. *Language Testing*, *29*(2), 265-289.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. DOI: 10.1177/074193258600700104

Greulich, L., Al Otaiba, S., Schatschneider, C., Wanzek, J., Ortiz, M., & Wagner, R. K. (2014). Understanding inadequate response to first-grade multi-tier intervention: Nomothetic and ideographic perspectives. *Learning Disability Quarterly*, *37*(4), 204 217.

Hancock, G. R., & Mueller, R. O. (Eds.). (2013). *Structural equation modeling: A second course*. Iap.

Harlaar, N., Cutting, L., Deater-Deckard, K., DeThorne, L. S., Justice, L. M., Schatschneider, C., ... & Petrill, S. A. (2010). Predicting individual differences in reading comprehension: A twin study. *Annals of Dyslexia*, *60*(2), 265-288.

Hauerwas, L. B., Brown, R., & Scott, A. N. (2013). Specific learning disability and response to intervention: State-level guidance. *Exceptional Children, 80*, 101–120.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*(2), 127–160.

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

Huizinga, M., Dolan, C.V., & van der Molen, M.W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017–2036. DOI:10.1016/j.neuropsychologia.2006.01.010

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, *85*(4), 512-552.

Jacobson, L. A., Koriakin, T., Lipkin, P., Boada, R., Frijters, J. C., Lovett, M. W., & Bosson-Heenan, J. (2016). Executive functions contribute uniquely to reading competence in minority youth. *Journal of Learning Disabilities*, DOI: 0022219415618501

Justice, L. M. (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools*, 37(4), 284–297.

Kendeou, P., Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*(1), 10-16.

Kendeou, P., Savage, R., & Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology*, *79*(2), 353-370.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*(4), 765-778.

Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., … & Dunlop Velez, E. (2016). *The Condition of Education 2016 (NCES 2016-144)*. U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch.

Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology*, *27*(5), 505.

Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Researcher*, *39*(6), 484-486.

Kieffer, M. J., Vukovic, R. K., & Berry, D. (2013). Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Reading Research Quarterly*, *48*(4), 333-348.

Kline, R. B. (2015). *Methodology in the social sciences. Principles and practice of structural equation modeling (4th ed.).* New York, NY, US: Guilford Press.

Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (2000). *Wisconsin Card Sorting Test–64 card version.* Lutz, FL: PAR.

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY–Second Edition (NEPSY-II). *San Antonio: The Psychological Corporation*.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of educational psychology*, *95*(2), 211-244.

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*(3), 596-632.

Lipka, O., Lesaux, N. K., & Siegel, L. S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities risk status and profiles over 5 years. *Journal of Learning Disabilities*, *39*(4), 364-378.

Locascio, G., Mahone, E. M., Eason, S., & Cutting, L. (2010). Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities, 43*(5), 441-454.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. G. (2002). *Gates–MacGinitie Reading Test (4th ed.).* Itasca, IL: Riverside.

Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., ... & Denckla, M. B. (2002). Effects of IQ on executive function measures in children with ADHD. *Child Neuropsychology*, *8*(1), 52-65.

Martel, M., Nikolas, M., & Nigg, J. T. (2007). Executive function in adolescents with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(11), 1437-1444.

Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2014). *TOSWRF-2: Test of Silent Word Reading Fluency*. Pro-ed.

Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading*, *9*(2), 85-116.

Menard, S. (1995). Applied logistic regression analysis. Thousand Oaks, CA: Sage.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental science*, *10*(6), 719-726.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied psycholinguistics*, *21*(02), 229-241.

Nation, K., & Snowling, M. J. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359–370.

Nouwens, S., Groen, M. A., & Verhoeven, L. (2016). How storage and executive functions contribute to children's reading comprehension. *Learning and Individual Differences*, *47*, 96-102.

Ortiz, A. A., & Artiles, A. J. (2010). Meeting the needs of ELLs with disabilities: A linguistically and culturally responsive model. In G. Li & P.A. Edwards (Eds.), *Best practices in ELL instruction* (pp. 24-272). New York, NY: Guilford.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*(2), 189-208.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive psychology*, *66*(2), 232-258.

Pearson (2014). *Achievement Improvement Monitoring System (Aimsweb) Technical Manual.* Bloomington, MN: NCS Pearson, Inc.

Piasta, S. B., & Wagner, R. K. (2010). Developing early literacy skills: A meta-analysis of alphabet learning and instruction. *Reading Research Quarterly*, *45*(1), 8-38.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*, 510-519.

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S.

(1995). *Listening to children read aloud.* U.S. Department of Education, National Center for Education Statistics, Washington, DC.

Prevoo, M. J. L., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*, *86*(1), 237-276.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition*, *13*(2), 253-262.

Proctor, C. P., August, D, Snow, C. E., & Barr, C. (2010). The interdependence continuum: A perspective on the nature of Spanish-English bilingual reading comprehension. *Bilingual Research Journal*, *33*(1), 5-20.

Proctor, C. P., Carlo, M. S., August, D., & Snow, C. E. (2005). Native Spanish-Speaking Children Reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246 – 256.

Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish-English bilingual children in elementary school. *Reading and Writing: An Interdisciplinary Journal*, *25*(7), 1635-1664.

Proctor, C. P, Silverman, R. D., & Harring, J. R. (2013). The CLAVES Intervention Project: Developing a supplemental intervention for Comprehension, Linguistic Awareness, and Vocabulary in English for Spanish speakers. Funded by the Institute for Education Sciences.

Qu, L., Low, J. J. W., Zhang, T., Li, H., & Zelazo, P. D. (2016). Bilingual advantage in executive control when task demands are considered. *Bilingualism: Language and Cognition*, *19*(2), 277-293.

Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly*, *24*(2), 119-134.

Rasinski, T. V., Padak, N., McKeon, C., Krug-Wilfong, L., Friedauer, J., & Heim, P. (2005). Is reading fluency a key for successful high school reading?. *Journal of Adolescent and Adult Literacy, 49*, 22–27.

Rueda, R., & Windmueller, M. P. (2006). English language learners, LD, and overrepresentation: A multiple-level analysis. *Journal of Learning Disabilities*, *39*(2), 99-107.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children-second edition (BASC-2).* Circle Pines, MN: AGS.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 cost-benefit analysis of the Title I Chicago child-parent centers. *Educational Evaluation and Policy Analysis, 24*(4), 267-303.

Ritchey, K. D., Silverman, R. D., Montanaro, E. A., Speece, D. L., & Schatschneider, C. (2012). Effects of a tier 2 supplemental reading intervention for at-risk fourth-grade students. *Exceptional Children*, 78, 318–334.

Ritchey, K. D., Silverman, R. D., Schatschneider, C., & Speece, D. L. (2013). Prediction and stability of reading problems in middle childhood. *Journal of Learning Disabilities*, *48*(3), 298-309.

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents withADHD: Effects of reading difficulties and gender. *Journal of child psychology and psychiatry*, *43*(8), 988-1003.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York, NY: Guilford.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child development*, 973-980.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232-246.

Sexton, C. C., Gelhorn, H. L., Bell, J. A., & Classi, P. M. (2012). The co-occurrence of reading disorder and ADHD: epidemiology, treatment, psychosocial impact, and economic burden. *Journal of Learning Disabilities*, *45*(6), 538-564.

Shanahan, T., & Lonigan, C. J. (2010). The National Early Literacy Panel A Summary of the Process and the Report. *Educational Researcher*, *39*(4), 279-285.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., & Fletcher, J. M. (2004). Development of Left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological psychiatry*, *55*(9), 926-933.

Silverman, R. D., Speece, D. L., Harring, J. R., & Ritchey, K. D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading*, *17*(2), 108-133.

Snow, C., & Sweet, A. (2003). Reading for comprehension. In A. Sweet & C. Snow (Eds.), *Rethinking Reading Comprehension.* New York, NY: Guilford.

Speece, D. L., Ritchey, K. D., Silverman, R., Schatschneider, C., Walker, C. Y., & Andrusik, K. N. (2010). Identifying children in middle childhood who are at risk for reading problems. *School psychology review*, *39*(2), 258-276.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*(4), 745–759.

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

Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology and Psychiatry*, *36*(3), 399-408.

Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children*, *77*(3), 317-334.

Swanson, H. L. (1999). Reading comprehension and working memory in learning-disabled readers: Is the phonological loop more important than the executive system?. *Journal of Experimental Child Psychology*, *72*(1), 1-31.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities?. *Journal of learning disabilities*, *39*(3), 252-269.

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: Another piece of the puzzle. *Journal of Learning Disabilities*, *47*(1), 22-33.

Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2008). Executive functions: performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53-72.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of learning disabilities*, *34*(1), 33-58.

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux vocabulary knowledge and the independent components hypothesis. *Journal of learning disabilities*, *45*(5), 453-466.

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sanchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics*, *36*(5), 1077-1109.

Valdés, G., Capitelli, S., & Alvarez, L. (2011). *Latino children learning English: Steps in the journey*. New York, NY: Teachers College Press.

Van der Kolk, A., Van Agthoven, M., Buitelaar, J. K., & Hakkaart-van Roijen, L. (2015). A systematic review of literatures on factors associated with educational and academic performance in attention deficit hyperactivity disorder. *Creative Education*, *6*(2), 164-180.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific studies of reading*, *11*(1), 3-32.

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, *22(3)*, 407-423.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of silent reading efficiency and comprehension*. Pro-Ed.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—4th Edition (WISC-IV).* San Antonio, TX: Harcourt Assessment.

WIDA Consortium. (2010). *Assessing comprehension and communication in English state to state for English language learners: ACCESS for ELLs: Interpretive guide for score reports.*

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*, *57*(11), 1336-1346.

Woodcock, R. W. (2004). Woodcock-Muñoz Language Survey-Revised comprehensive manual. *Riverside. Chicago, IL.*

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *The Woodcock-Johnson III tests of achievement.* Rolling Meadows, IL: Riverside Publishing.

Yang, H., Hartanto, A., & Yang, S. (2017). Bilingualism confers advantages in task switching: Evidence from the dimensional change card sort task. *Bilingualism: Language and Cognition*, 1-19.

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1-9.

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In *Blackwell Handbook Of Childhood Cognitive Development*, U. Goswami (Eds), (pp 445-469). Malden MA: Blackwell Publishing.

Appendices

Tables

Table 2.1. *Study Features.*

Study (Year)	**Total Sample, Age/Grade**	**Upper elementary** (n if specified)		**Subgroups** (n if specified)				**Location**
		4th Grade	5th Grade	Typical Readers	Reading Difficulties	Attention Disorders	English Learners	
Christopher et al. (2012)	N = 483 (8 – 16 years old)	Yes	Yes	Yes	Yes (n = 128)	Yes (n = 93)	No	NW, US
Cutting et al. (2009)	N = 56 (9 – 14 years old)	Yes	Yes	Yes (n = 21)	Yes (n = 25)	Yes	No	NE, US
Gerst et al. (2015)	N = 93	Yes (n = 54)	Yes (n = 40)	Yes	Yes	NS	Yes	SW, US
Jacobson et al. (2016)	N = 761 (8 – 16 years old)	Yes	Yes	Yes	Yes	NS	No	NE, NW, MA, SW, US; CA; & PR
Kieffer et al. (2013)	N = 120	Yes (N = 120)	None	Yes	Yes	NS	Yes	NE, US
Nouwens et al. (2016)	N = 123	None	Yes (N = 123)	Yes	Yes (n = 14)	Yes (n = 12)	Yes	Netherlands
Sesma et al. (2009)	N = 60 (9 – 15 years old)	Yes	Yes	Yes (n = 31)	Yes (n = 26)	Yes (n = 14)	No	NS

Table 2.2. *Dependent variables, independent variables, and covariates across studies.*

Study (Year)	Reading Comprehension			Decoding and Fluency			Linguistic Comprehension		Other
	Reading Inventory	Cloze	Multiple Choice	Untimed Word	Timed Word	Timed Passage	Vocab.	Oral Comp.	
Christopher et al. (2012)	GORT-3 QRI-3: RC	WJ-III: PC	PIAT: C	PIAT: WR	TLORSW			WJ-III: OC QRI-3: LC† Barnes K-I†	PIAT: S
Cutting et al. (2009)	GORT-4: C	WRMT-R/NU: PC		WRMT-R/NU: WA, WI	TOWRE: SWE, PDE	GORT-4: F	PPVT-III	TOLD-3: GC, SC TLC-E: AS, MI	WISC-III: IQ
Gerst et al. (2015)			GMRT-4: PC	CTOPP Elision				WJ-III: OC	
Jacobson et al. (2016)		SRI-2		WJ-III: LWID, WA CTOPP: Elision	TOWRE: SWE, PDE		PPVT-III		
Kieffer et al. (2013)			GMRT-4: PC	WJ-III: LWID CTOPP: Elision				WJ-III: OC	
Nouwens et al. (2016)			Diatekst		EMT		PPVT-III-NL		RCPM
Sesma et al. (2009)	WIAT-II: RC			WIAT-II: WR WRMT-R: WA		GORT-4: F	PPVT-III		

Note. † = adapted measure. WJ-III Oral = Woodcock-Johnson Oral Comprehension (Woodcock, McGrew, & Mather, 2001); QRI-3: RC = Qualitative Reading Inventory-3 Reading Comprehension (Leslie & Caldwell, 2001); Barnes KNOW-IT (Barnes & Dennis, 1996; Barnes et al., 1996); GORT-3 = Gray Oral Reading Test-3 (Wiederholt & Bryant, 1992); QRI-3: LC = Qualitative Reading Inventory Listening Comprehension; WJ-III: PC = Woodcock-Johnson Passage Comprehension; PIAT-C = Peabody Individual Achievement Test Comprehension (Markwardt, 1970); PIAT-WR = PIAT Word Recognition; PIAT-S = PIAT Spelling; TLORSW= Time-limited oral reading of single words (Olson et al., 1994).

Table 2.2 *Continued*

Study	Executive Function Measures					
	Working Memory	**Processing Speed**	**Shifting**	**Inhibition**	**Planning**	**Attention**
Christopher et al. (2012)	WISC-R/WISC-III: DS Sentence Span† Counting Span†	CPS: Test 1-2 ETS Identical Pictures	*(Naming Speed)* RAN Colors and Objects†	GDS CPT: V & D SSRT		
Cutting et al. (2009)	WISC-III/WISC-IV: DS Backwards			*(Spatial Plan., Rule Learning, Inhib.)* TOL	*(Planning, Organization, Monitoring)* Elithorn PMT	DuPaul
Gerst et al. (2015)	BRIEF-T WM WMTB-C: LR		DKEFS: TMT Number-Letter Switching	DKEFS: CWIT Inhibition	I-3 Tower Task	
Jacobson et al. (2016)	WISC-IV: DS Forwards	WISC-IV: PSI	*(Shifting/ Switching)* RAN/RAS Composite		*(Problem Solving)* WISC-IV: Matrix Reasoning	
Kieffer et al. (2013)	SCPT: Visual Matrix	WJ-III: Visual Matching	*(Attention Shifting)* WCST	*(Inhibitory Control)* Number-Quantity Stroop†		
Nouwens et al. (2016)	Competing Language Processing Task *(Storage)* WISC-III-NL: DS Forwards		*(Cognitive Flexibility)* D-KEFS: TMT Number-Letter Switching	DKEFS: CWIT Inhibition	D-KEFS: Tower	
Sesma et al. (2009)	WISC-III: FDI				TOL	BASC

Table 2.3. *Contribution of EF to RC by study.*

Study	EF Predictors	RC Outcomes	Covariates	Full Model R^2	*β*	*p (β)*
Christopher et al. (2012)	WISC-R/WISC-III: DS, SS† & CS† CPS: 1-2 & ETS IP RAN C & O† CPT: VD/SSRT	GORT-3; QRI-3: RC; WJ-III: PC; PIAT: C	WJ-III: OC; Barnes K-I; PIAT: WR; TLORSW; WJ-III: OC; QRI-3: LC†; B K-I†; PIAT: S	NS	.64 .13 -.03 -.10	.01 .08 .72 .52
Cutting et al. (2009)	TOL Ellithorn Mazes	GORT-4	TOWRE; Oral language	.52	-.28 -.01	.05 .93
	TOL Ellithorn Mazes	WRMT/NU	TOWRE; Oral language	.68	-.14 -.07	.24 .57
Gerst et al. (2015)	BRIEF-T WM WMTB-C TOL DKEFS CWIT DKEFS TWT	GMRT-4: RC	Age; WJ-III: OC; CTOPP Ellison; SE; G/T	.68	-.25 .24 .06 -.07 .07	.01 .01 .39 .41 .37
Jacobson et al. (2016)	WISC-IV: PSI EF Composite WISC-IV: DSF RAN/RAS WISC-IV: MR	SRI-2	Demographics; WJ-III: LWID, WA; CTOPP: Elision; TOWRE: SWE, PDE; PPVT-III	.60	 .06 .34 .03	<.01 <.01 .02 <.01 .20
Kieffer et al. (2013)	WCST N-Q Stroop† SCPT: VM WJ-III: VM	GMRT-4: PC	WJ-III: LWID; CTOPP: Elision; WJ-III: OC	NS	.16 .19 -.05 .13	2.04 *z* 2.57 *z* -0.77 *z* 1.50 *z*
Nouwens et al. (2016)	WISC-III-NL: DSF DKEFS: CWIT I D-KEFS: TMT D-KEFS: T CLPT	Diatekst	RCPM; EMT; PPVT-III-NL	.32	-.02 .15 -.19 -.09 .27	.83 .13 .04 .25 .01
Sesma et al. (2009)	BASC WISC-III FSI TOL	WIAT-II: RC	WRMT-R: WA; GORT-4: F; PPVT-III	.63	-.13 .17 -.25	.001 .032 .013

Table 3.1 *Sample size disaggregated by language status and also by disability.*

	N (%)
Overall	126
Language Status	
EL	55 (44)
REL	71 (56)
Disability	
IEP	10 (8)
No IEP	116 (92)

Table 3.2 *Means (SD) for Executive Function, Covariates, and Reading Comprehension, disaggregated by language status and also by disability.*

	Executive Function			Covariates		Reading Comprehension		
	Working Memory	Shifting	Inhibition	Decoding	Linguistic Comp.	GMRT	MAZE	TOSREC
Overall	38.44	186.99	152.67	78.96	475.69	21.10	10.95	22.89
	(9.55)	(47.67)	(4.12)	(18.17)	(19.44)	(7.64)	(6.04)	(8.02)
Language Status								
EL	35.23***	190.93	151.18***	74.69*	468.91***	19.18**	8.73***	21.42
	(9.57)	(49.74)	(4.24)	(16.00)	(17.56)	(6.19)	(4.01)	(7.80)
REL	40.93	183.94	153.82	82.27	480.94	22.58	12.68	24.03
	(8.83)	(46.13)	(3.67)	(19.15)	(19.30)	(8.35)	(6.77)	(8.05)
Disability								
IEP	31.25**	189.00	149.80*	64.30**	466.90	18.30	7.20*	19.70
	(9.77)	(55.48)	(5.53)	(24.47)	(20.41)	(3.74)	(3.19)	(8.12)
No IEP	39.06	186.82	152.91	80.22	476.45	21.34	11.28	23.16
	(9.32)	(47.21)	(3.91)	(17.08)	(19.25)	(7.85)	(6.12)	(7.99)

Note. *($p < 0.05$), **($p < 0.01$), ***($p < 0.001$), ~($p < .10$).

Table 3.3 *Pearson Correlation Coefficients for All Measures*

	1	2	3	4	5	6	7	8
Working Memory	-							
Shifting	-0.096	-						
Inhibition	.286**	-0.004	-					
Decoding	0.068	-.219*	.241**	-				
Linguistic Comp.	.238**	-0.044	.229**	.425**	-			
GMRT	.197*	-0.038	.265**	.386**	.622**	-		
MAZE	0.086	-0.128	.194*	.606**	.523**	.505**	-	
TOSREC	0.031	-0.028	0.114	.408**	.378**	.427**	.529**	-

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Table 3.4 *Regression Weights for Research Question 1, Model 1*

			Estimate	S.E.	Standardized Estimate	P
LRC	<---	LEF	0.492	0.237	0.396	0.037
GMRT	<---	LRC	1		0.663	
MAZE	<---	LRC	0.928	0.169	0.779	0.001
TOSREC	<---	LRC	1.041	0.19	0.658	0.001
W. Memory	<---	LEF	1		0.426	
Shifting	<---	LEF	-0.935	1.46	-0.08	0.522
Inhibition	<---	LEF	0.672	1.46	0.663	0.101

Table 3.5 *Regression Weights for Research Question 1, Model 2*

			Estimate	S.E.	Standardized Estimate	P
LRC	<---	Decoding	0.135	0.025	0.522	0.001
LRC	<---	Ling. Comp.	0.143	0.024	0.593	0.001
LRC	<---	LEF	-0.001	0.033	-0.005	0.987
GMRT	<---	LRC	1		0.646	
MAZE	<---	LRC	0.909	0.144	0.757	0.001
TOSREC	<---	LRC	0.923	0.18	0.561	0.001
W. Memory	<---	LEF	1		4.302	
Shifting	<---	LEF	-0.025	1.504	-0.022	0.987
Inhibition	<---	LEF	0.007	0.393	0.067	0.986

Table 3.6 *Regression Weights for Research Question 1, Model 3*

			Estimate	S.E.	Standardized Estimate	P
Decoding	<---	LEF	3.720	1.491	0.612	0.013
Ling. Comp.	<---	LEF	4.302	1.743	0.661	0.014
LRC	<---	Decoding	0.136	0.025	0.470	0.001
LRC	<---	Ling. Comp.	0.142	0.025	0.527	0.001
GMRT	<---	LRC	1		0.688	
MAZE	<---	LRC	0.907	0.123	0.790	0.001
TOSREC	<---	LRC	0.921	0.156	0.603	0.001
W. Memory	<---	LEF	1		0.313	
Shifting	<---	LEF	-2.868	2.046	-0.180	0.161
Inhibition	<---	LEF	0.551	0.245	0.399	0.024

Table 3.7 *Comparative fit for path models (N = 126)*

Model	x^2	*df*	*p*	CFI	RMSEA	AIC
1. EF to RC direct relationship without DC LC	8.601	8	0.377	0.994	0.025	34.601
2. EF to RC direct relationship with DC LC	68.676	19	0.001	7.840	0.145	102.676
3. EF to RC indirect relationship through DC LC	35.957	18	0.007	0.922	0.089	71.957

Table 3.8 *Principal Component Analysis for Latent RC*

	Component Score Coefficient
GMRT	0.421
TOSREC	0.400
MAZE	0.412

Table 3.9 *Principal Component Analysis for Latent EF*

	Component Score Coefficient
Working Memory	0.618
Shifting	-0.204
Inhibition	0.587

Table 4.1 *Sample size disaggregated by condition, language status and disability.*

	N (%)	Intervention	Control
Overall	120	57	63
Language Status			
EL	53 (44)	27	27
REL	67 (56)	31	36
Disability			
IEP	9 (7)	4	5
No IEP	111 (93)	53	58

Table 4.2 *Means (SD) for Executive Function, Covariates, and Reading Comprehension, disaggregated by language status and also by disability.*

		Fall Assessment						
		GMRT	MAZE	TOSR	SYN	VOC	MOR	GMRT
Condition								
	Cont.	18.38	7.06	14.48	33.92	30.63	20.16	20.84
		(7.94)	(3.79)	(6.28)	8.44	3.78	3.61	7.78
	Inter.	18.47	7.32	13.84	32.19	29.63	20.21	21.63
		(7.52)	(3.38)	(4.05)	(8.08)	(4.43)	(5.21)	(7.68)
Language Status								
Control	EL	19.50	7.83	15.58	35.44	31.14	21.08	22.58
		(8.72)	(3.80)	(7.17)	(9.11)	(3.57)	(3.49)	(8.52)
	REL	16.89	6.04	13.00	31.89	29.96	18.93	18.52
		(6.63)	(3.59)	(4.60)	(7.12)	(4.00)	(3.44)	(6.07)
Inter.	EL	20.23	8.00	14.84	33.97	30.65	22.19	23.26
		(8.23)	(3.10)	(3.43)	(9.81)	(3.99)	(4.59)	(8.37)
	REL	16.38	6.50	12.65	30.08	28.42	17.85	19.69
		(6.07)	(3.57)	(4.46)	(4.71)	(4.70)	(4.99)	(6.39)
Disability Status								
Control	No IEP	19.19	7.19	14.66	34.41	30.98	20.16	21.14
		(7.69)	(3.74)	(6.46)	(8.46)	(3.57)	(3.75)	(7.97)
	IEP	9.00	5.60	12.40	28.20	26.60	20.20	17.40
		(3.81)	(4.45)	(3.36)	(6.22)	(4.16)	(1.30)	(4.16)
Inter.	No IEP	18.92	7.45	14.08	32.60	29.75	20.43	21.75
		(7.44)	(3.46)	(4.07)	(8.23)	(4.47)	(5.27)	(7.91)
	IEP	12.50	5.50	10.75	26.75	28.00	17.25	20.00
		(6.56)	(1.00)	(2.22)	(0.96)	(4.08)	(3.59)	(3.46)

Table 4.2 *Continued*

		Spring Assessment						
		MAZE	TOSR	CALS	DC	WM	SHIFT	INHIB
Condition								
	Cont.	10.83	23.59	473.46	76.49	39.78	189.48	152.79
		5.19	8.12	19.68	18.18	9.14	49.63	4.12
	Inter.	11.47	22.14	479.05	82.07	36.73	184.65	152.40
		(6.84)	(7.82)	19.53	(16.79)	(9.53)	(44.31)	(4.27)
Language Status								
Control	EL	12.78	25.39	479.97	79.86	41.49	183.11	153.67
		(5.27)	(8.67)	18.80	(19.83)	(9.35)	(49.59)	(3.85)
	REL	8.22	21.19	464.78	72.00	37.50	197.96	151.63
		(3.83)	(6.75)	17.63	(14.92)	(8.49)	(49.33)	(4.24)
Inter.	EL	13.10	22.48	483.97	84.35	39.72	184.26	153.81
		(8.20)	(7.16)	20.12	(19.21)	(8.16)	(41.61)	(3.64)
	REL	9.54	21.73	473.19	79.35	33.17	185.12	150.73
		(4.13)	(8.67)	17.39	(13.22)	(9.96)	(48.16)	(4.42)
Disability Status								
Control	No IEP	11.09	23.90	474.52	78.00	40.19	188.47	152.71
		(5.22)	(8.16)	19.25	(16.24)	(9.30)	(49.67)	(4.16)
	IEP	7.80	20.00	461.20	59.00	35.00	201.20	153.80
		(4.09)	(7.58)	22.75	(30.82)	(5.59)	(53.36)	(3.90)
Inter.	No IEP	11.77	22.17	479.36	83.21	37.26	183.85	152.98
		(7.00)	(7.81)	19.63	(16.20)	(9.17)	(44.17)	(3.76)
	IEP	7.50	21.75	475.00	67.00	29.69	195.25	144.75
		(1.29)	(9.18)	20.31	(19.78)	(12.88)	(51.69)	(3.30)

Table 4.3 *Pearson Correlation Coefficients for All Measures*

Test	1	2	3	4	5	6	7	8	9	10	11	12	13
1GMR	-												
1MAZ	.527**	-											
1TOS	.479**	.446**	-										
1SYN	.373**	.231*	.419**	-									
1VOC	.399**	.324**	.406**	.429**	-								
1MOR	.352**	.347**	.255**	.400**	.277**	-							
2GMR	.695**	.520**	.408**	.354**	.343**	.353**	-						
2MAZ	.522**	.534**	.421**	.417**	.341**	.397**	.513**	-					
2TOS	.440**	.510**	.489**	.373**	.411**	.336**	.436**	.527**	-				
2CAL	.644**	.467**	.412**	.425**	.444**	.407**	.620**	.514**	.381**	-			
2DC	.498**	.460**	.341**	.258**	0.178	.264**	.450**	.641**	.470**	.458**	-		
2WM	.183*	.199*	0.134	0.105	0.116	.260**	.202*	0.055	-0.027	.241**	0.079	-	
2SHI	-0.101	-0.156	-0.1	-0.093	-0.007	-0.043	-0.073	-0.152	-0.075	-0.058	-.189*	-0.151	-
2INH	.233*	0.142	.234*	.229*	0.13	0.14	.273**	.206*	0.11	.240**	.251**	.276**	-0.029

**. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed).

Table 4.4 *Regression Weights for Research Question 1, Model 1*

			Control				Intervention			
			Est.	S.E.	Std. Est.	P	Est.	S.E.	Std. Est.	P
PostCALS	<---	PreLC	3.811	0.778	0.935	0.001	3.175	0.999	0.718	0.001
PostLRC	<---	PreLRC	0.612	0.155	0.790	0.001	0.606	0.166	0.878	0.001
PostLRC	<---	Decoding	0.094	0.028	0.393	0.001	0.133	0.034	0.610	0.001
PostLRC	<---	PostCALS	0.056	0.023	0.251	0.017	0.072	0.026	0.380	0.006
PostGMRT	<---	PostLRC	1.000		0.623		1.000		0.551	
PostMAZE	<---	PostLRC	0.820	0.170	0.816	0.001	0.872	0.245	0.535	0.001
PostTOSREC	<---	PostLRC	1.140	0.261	0.697	0.001	0.805	0.278	0.411	0.004
PreSyntax	<---	PreLC	1.000		0.570		1.000		0.536	
PreVocab	<---	PreLC	0.372	0.118	0.473	0.002	0.715	0.225	0.713	0.001
PreMorph	<---	PreLC	0.357	0.113	0.477	0.002	0.663	0.232	0.563	0.004
PreGMRT	<---	PreLRC	1.000		0.709		1.000		0.708	
PreMAZE	<---	PreLRC	0.456	0.098	0.677	0.001	0.487	0.110	0.768	0.001
PreTOSREC	<---	PreLRC	0.879	0.169	0.787	0.001	0.351	0.118	0.461	0.003

Table 4.5 *Regression Weights for Research Question 1, Model 2*

			Control				Intervention			
			Est.	S.E.	Std. Est.	P	Est.	S.E.	Std. Est.	P
PostCALS	<---	PreLC	1.123	0.351	0.465	0.001	2.678	0.868	0.629	0.002
PostCALS	<---	LEF	7.511	3.182	0.808	0.018	2.022	1.121	0.386	0.071
Decoding	<---	LEF	3.716	1.930	0.410	0.054	3.396	1.846	0.720	0.066
PostLRC	<---	PreLRC	0.612	0.153	0.768	0.001	0.751	0.179	0.938	0.001
PostLRC	<---	Decoding	0.094	0.029	0.383	0.001	0.071	0.037	0.263	0.055
PostLRC	<---	PostCALS	0.056	0.026	0.232	0.028	0.082	0.034	0.336	0.015
PostGMRT	<---	PostLRC	1.000		0.633		1.000		0.666	
PostMAZE	<---	PostLRC	0.820	0.164	0.823	0.001	0.406	1.034	0.051	0.694
PostTOS	<---	PostLRC	1.140	0.252	0.706	0.001	0.757	0.193	0.550	0.001
PreSyntax	<---	PreLC	1.000		0.915		1.000		0.544	
PreVocab	<---	PreLC	0.234	0.077	0.479	0.002	0.744	0.237	0.736	0.002
PreMorph	<---	PreLC	0.260	0.077	0.557	0.001	0.658	0.233	0.554	0.005
PreGMRT	<---	PreLRC	1.000		0.709		1.000		0.757	
PreMAZE	<---	PreLRC	0.456	0.098	0.677	0.001	0.423	0.095	0.713	0.001
PreTOS	<---	PreLRC	0.879	0.169	0.787	0.001	0.330	0.108	0.464	0.002
W.Memory	<---	LEF	1.000		0.224		1.000		0.374	
Shifting	<---	LEF	0.406	1.034	0.022	0.694	-2.611	2.329	-0.210	0.262
Inhibition	<---	LEF	0.757	0.193	0.278	0.001	0.554	0.304	0.462	0.068

Table 4.6 *Regression Weights for Research Question 2*

			Control*EL				Control*REL			
			Est.	S.E.	Std. Est.	P	Est.	S.E.	Std. Est.	P
Decoding	<---	LEF	1.017	1.178	0.265	0.388	2.941	2.109	0.392	0.163
PostCALS	<---	PreLC	1.387	0.658	0.462	0.035	0.828	0.362	0.415	0.022
PostCALS	<---	LEF	2.069	1.621	0.455	0.202	5.907	3.504	0.895	0.092
PostLRC	<---	PreLRC	0.356	0.148	1.248	0.016	0.862	0.228	0.936	0.001
PostLRC	<---	Decoding	0.061	0.037	0.659	0.101	0.094	0.035	0.330	0.008
PostLRC	<---	PostCALS	0.049	0.031	0.631	0.116	0.044	0.038	0.137	0.239
PostGMRT	<---	PostLRC	1.000		0.229		1.000		0.744	
PostMAZE	<---	PostLRC	0.988	0.192	0.322	0.001	0.660	0.140	0.809	0.001
PostTOS	<---	PostLRC	0.731	0.216	0.147	0.001	1.038	0.234	0.763	0.001
PreSyntax	<---	PreLC	1.000		0.824		1.000		0.959	
PreVocab	<---	PreLC	0.272	0.155	0.399	0.078	0.211	0.085	0.516	0.013
PreMorph	<---	PreLC	0.266	0.133	0.454	0.046	0.222	0.086	0.554	0.010
W.Memory	<---	LEF	1.000		0.466		1.000		0.283	
Shifting	<---	LEF	-3.043	1.874	-0.245	0.105	-0.437	3.269	-0.023	0.894
Inhibition	<---	LEF	0.366	0.350	0.336	0.296	0.350	0.321	0.240	0.275
PreGMRT	<---	PreLRC	1.000		0.732		1.000		0.702	
PreMAZE	<---	PreLRC	0.461	0.159	0.624	0.004	0.432	0.115	0.696	0.001
PreTOS	<---	PreLRC	0.827	0.242	0.873	0.001	0.852	0.217	0.727	0.001

Table 4.6 *Regression Weights for Research Question 2 Continued*

			Intervention*EL				Intervention*REL			
			Est.	S.E.	Std. Est.	P	Est.	S.E.	Std. Est.	P
Decoding	<---	LEF	3.274	1.619	0.955	0.043	2.081	1.352	0.418	0.124
PostCALS	<---	PreLC	3.339	1.788	0.503	0.062	0.847	0.413	0.378	0.040
PostCALS	<---	LEF	1.584	0.956	0.355	0.098	1.432	1.221	0.279	0.241
PostLRC	<---	PreLRC	0.886	0.221	1.593	0.001	0.155	0.176	0.132	0.378
PostLRC	<---	Decoding	0.030	0.030	0.157	0.314	0.205	0.045	0.735	0.001
PostLRC	<---	PostCALS	0.012	0.023	0.085	0.583	0.127	0.038	0.470	0.001
PostGMRT	<---	PostLRC	1.000		0.433		1.000		0.689	
PostMAZE	<---	PostLRC	0.988	0.192	0.527	0.001	0.988	0.192	0.798	0.001
PostTOS	<---	PostLRC	0.731	0.216	0.261	0.001	0.731	0.216	0.500	0.001
PreSyntax	<---	PreLC	1.000		0.542		1.000		0.911	
PreVocab	<---	PreLC	1.448	0.757	0.805	0.056	0.209	0.083	0.466	0.012
PreMorph	<---	PreLC	0.849	0.516	0.445	0.100	0.204	0.097	0.394	0.036
W.Memory	<---	LEF	1.000		0.389		1.000		0.466	
Shifting	<---	LEF	-3.043	1.874	-0.242	0.105	-3.043	1.874	-0.280	0.105
Inhibition	<---	LEF	0.463	0.272	0.405	0.089	0.633	0.359	0.669	0.078
PreGMRT	<---	PreLRC	1.000		0.750		1.000		0.557	
PreMAZE	<---	PreLRC	0.605	0.142	0.771	0.001	0.580	0.346	0.859	0.094
PreTOS	<---	PreLRC	0.279	0.188	0.285	0.138	0.345	0.178	0.461	0.053

Table 4.7 *Principal Component Analysis for Latent RC*

	Pretest Component Score Coefficient	Posttest Component Score Coefficient
GMRT	0.421	0.399
TOSREC	0.400	0.406
MAZE	0.412	0.427

Table 4.8 *Principal Component Analysis for Latent EF*

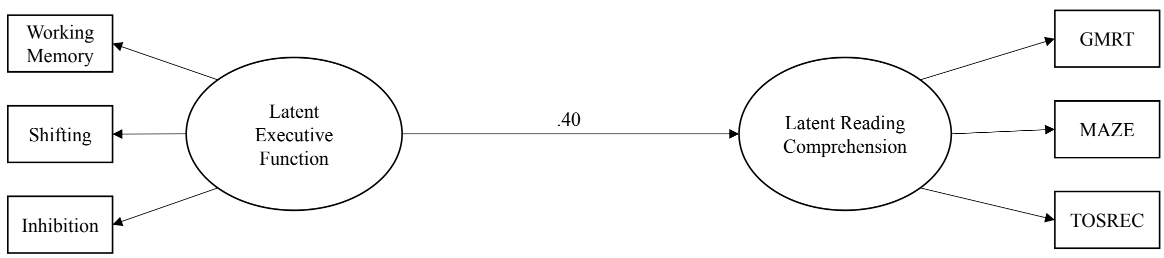
	Component Score Coefficient
Working Memory	0.618
Shifting	-0.204
Inhibition	0.587

Table 4.9. *Tests of Measurement Invariance by Time For Latent RC at Pretest And Posttest.*

Model	x^2	*df*	*p*	CMIN/*df*	CFI	RMSEA
Time Invariance						
Unconstrained	4.994	10	.89	.499	1.000	0.000
Measurement weights	4.994	14	.99	.357	1.000	0.000
Measurement intercepts	4.994	20	1.00	.250	1.000	0.000
Structural covariances	4.994	23	1.00	.227	1.000	0.000
Measurement residuals	4.994	32	1.00	.156	1.000	0.000

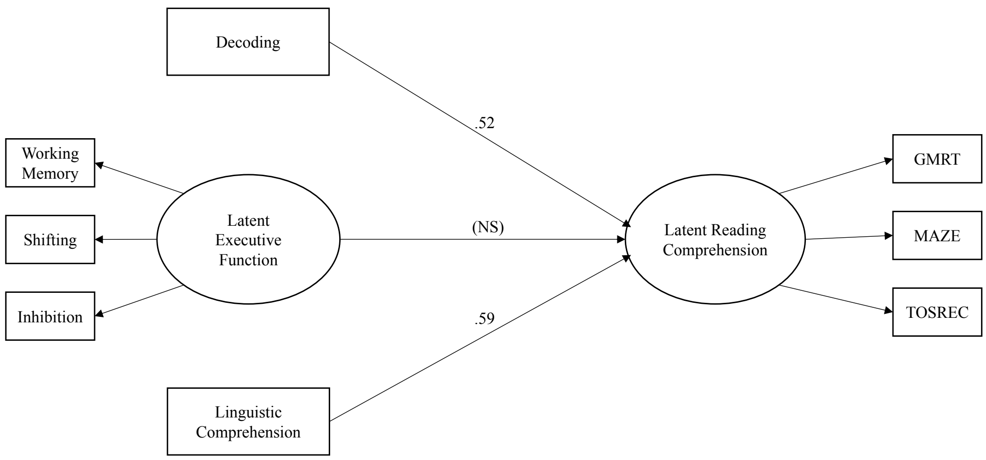
Figure 2.1 *Domains of EF in Frequently Cited Literature.*

Figure 3.1 *RQ 1, Model 1 Direct Latent EF to Latent Reading Comprehension*



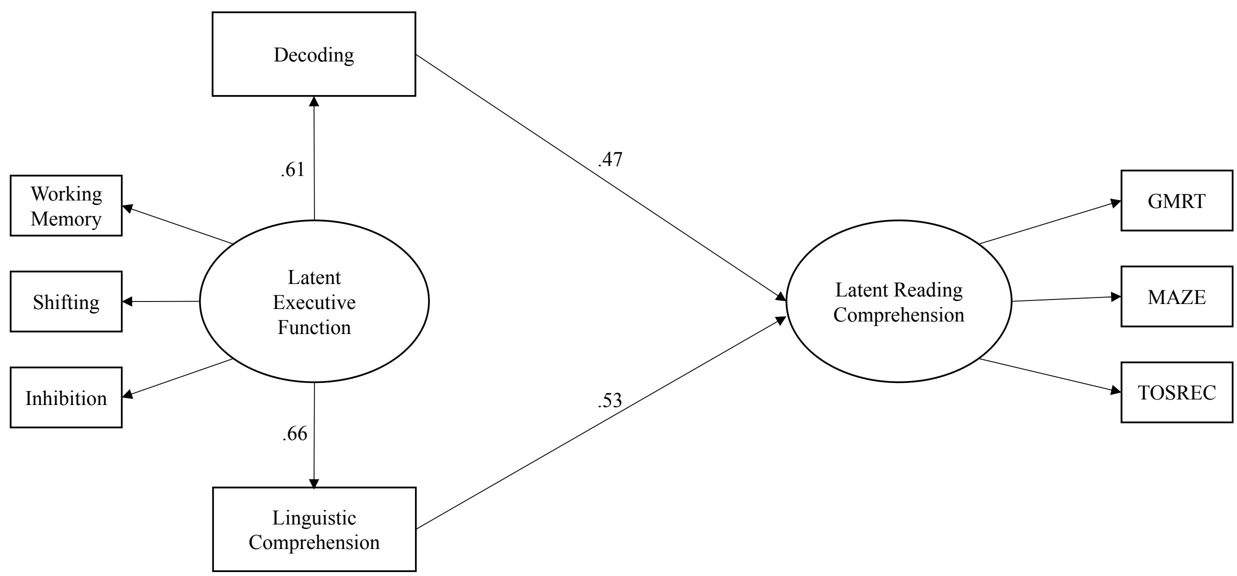
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 3.2 *RQ 1, Model 2 Direct Latent EF to Latent Reading Comprehension with Decoding and Linguistic Comprehension*



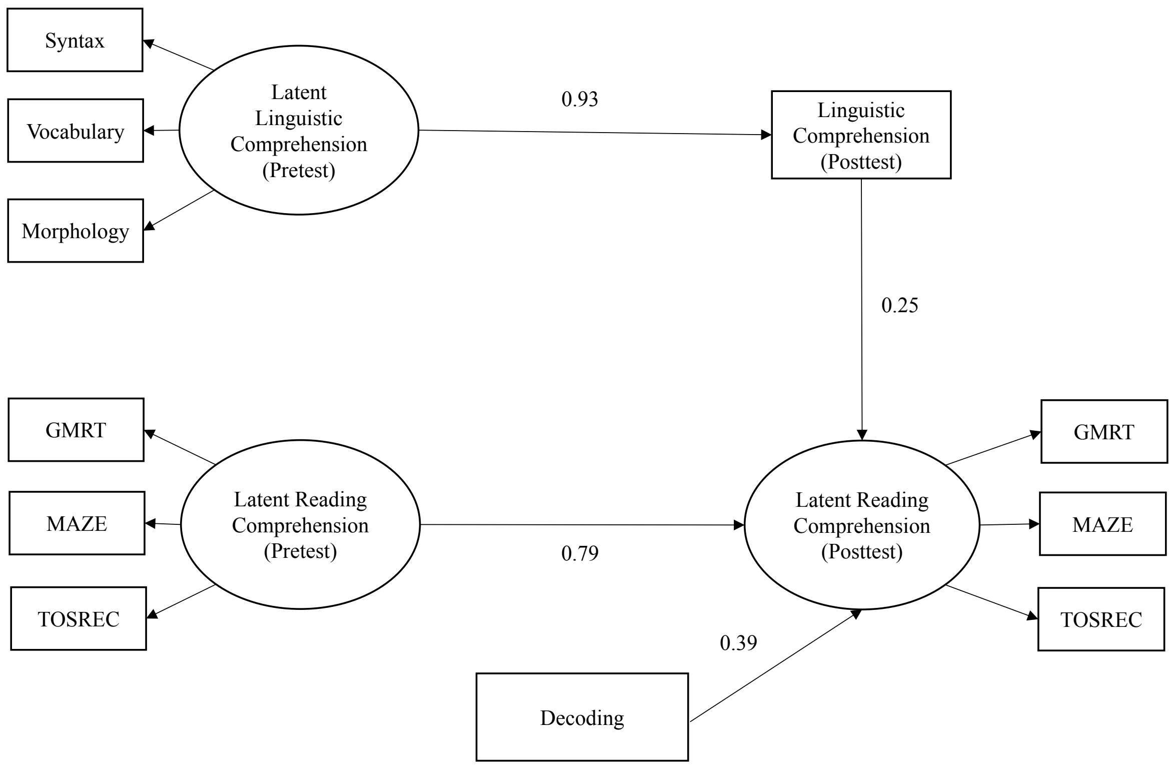
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 3.3 *RQ 1, Model 3 Indirect Latent EF to Latent Reading Comprehension through Decoding and Linguistic Comprehension*



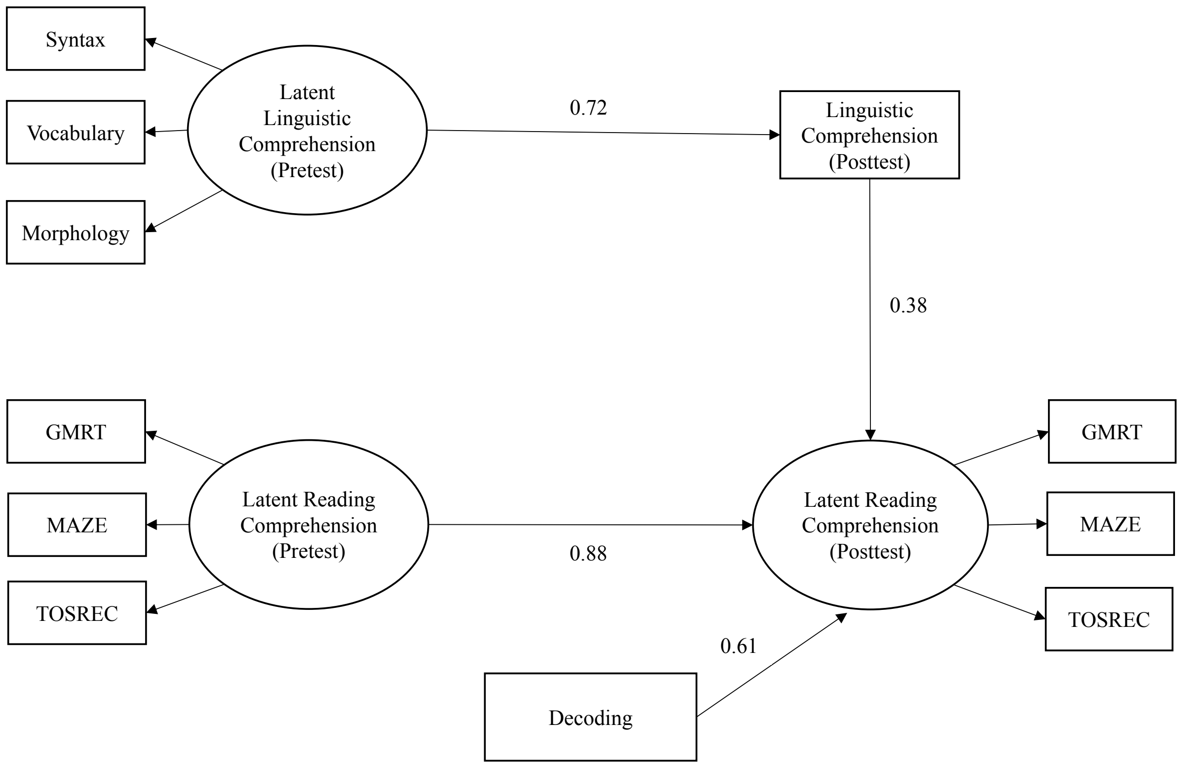
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.1 *RQ 1, Model 1 Control Standardized Regression Weights*



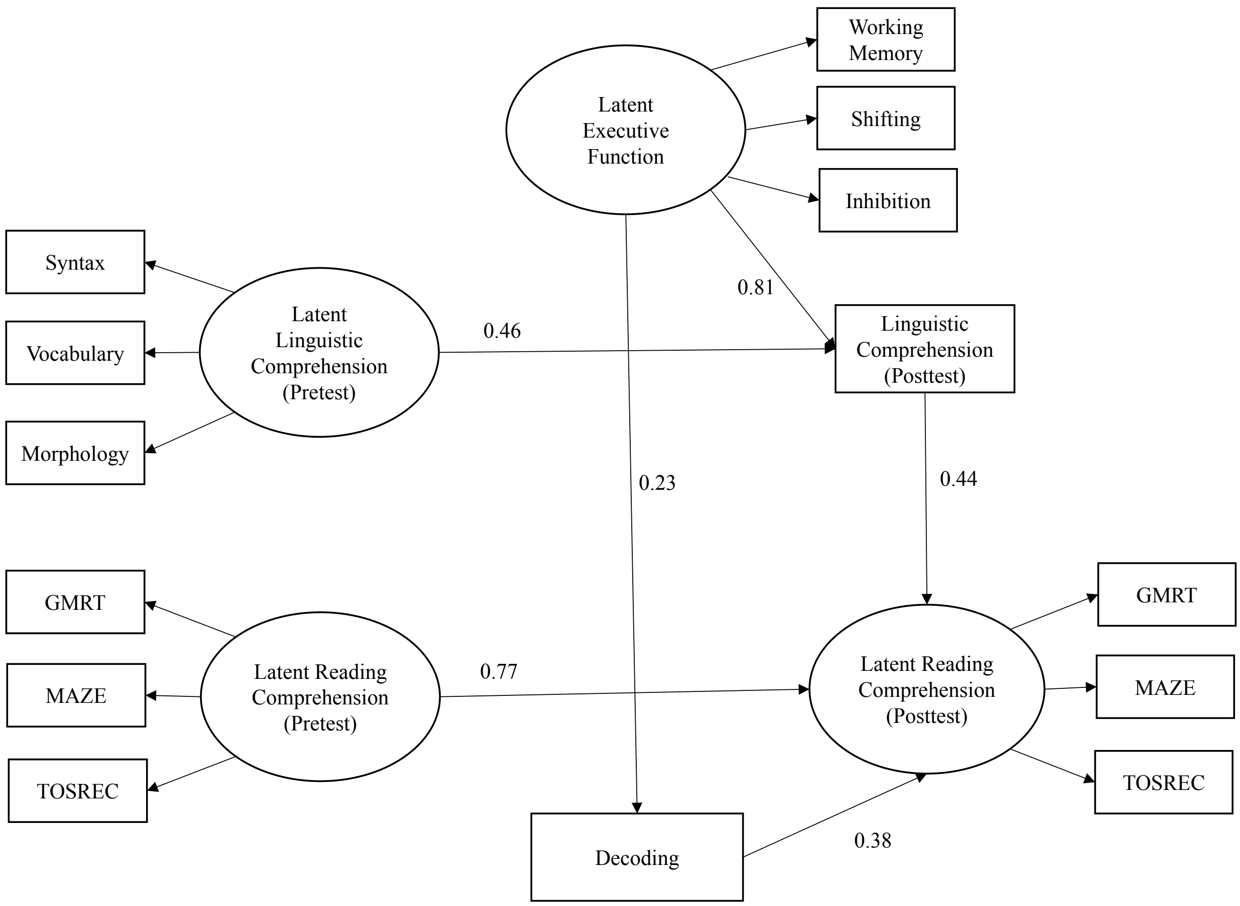
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.2 *RQ 1, Model 1 Intervention Standardized Regression Weights*



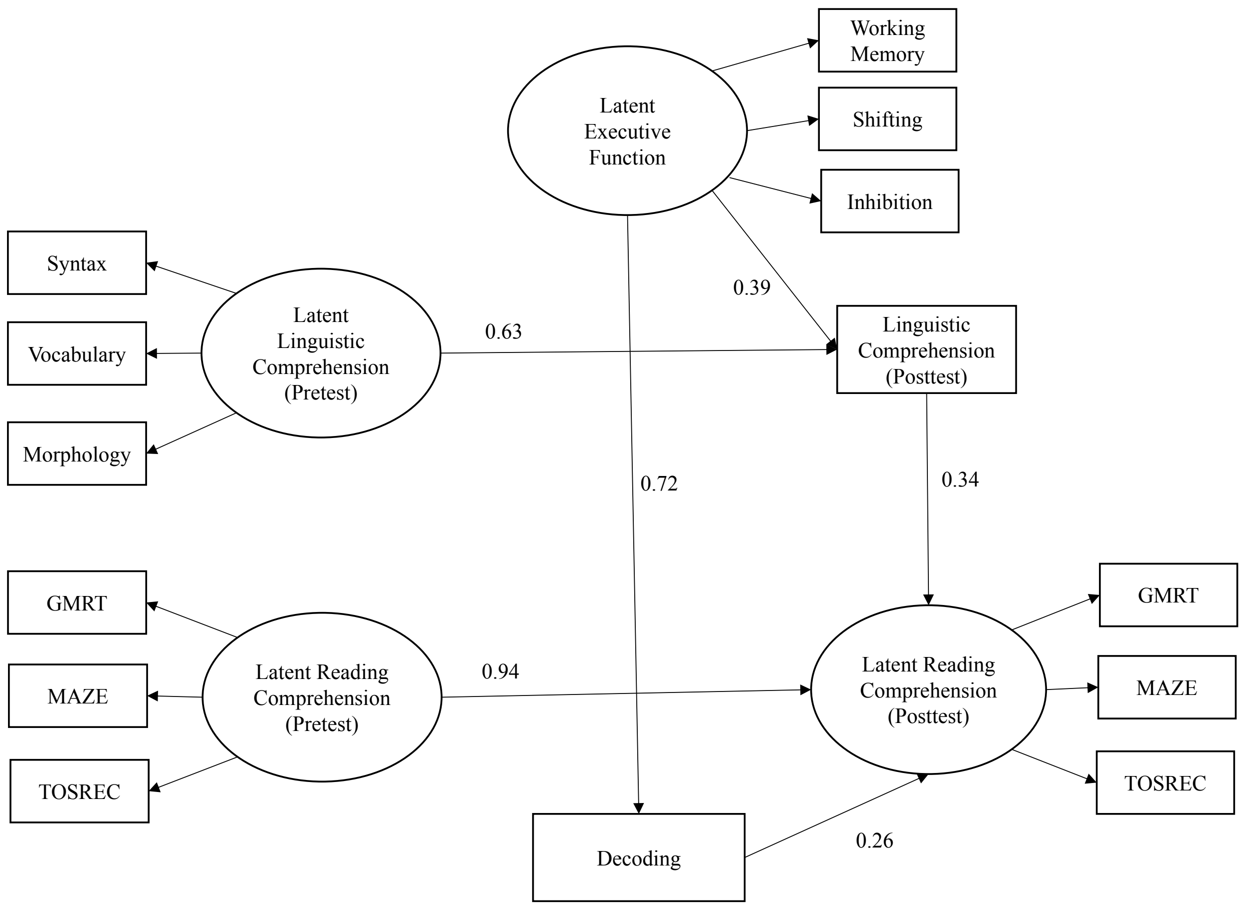
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.3 *RQ 1, Model 3 Control Standardized Regression Weights*



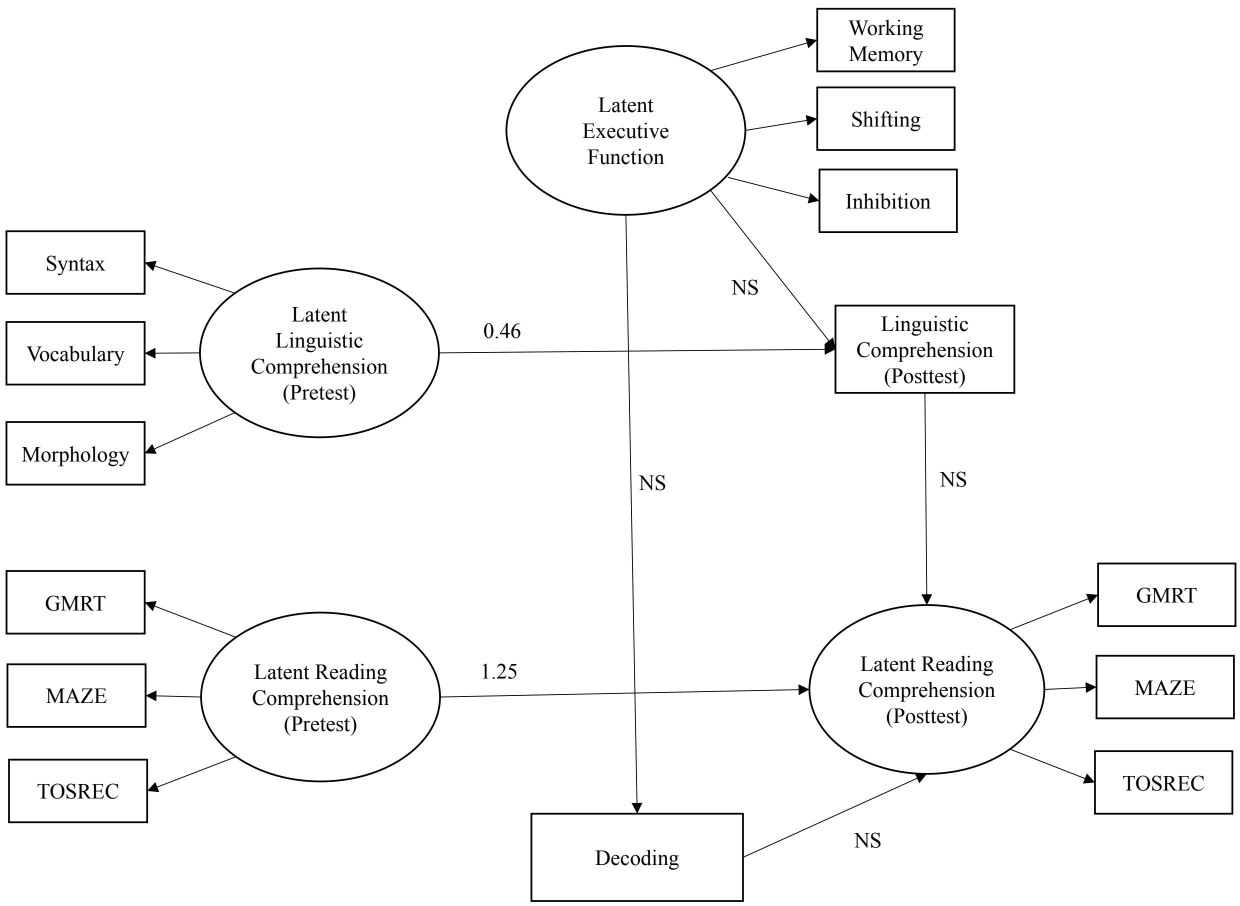
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.4 *RQ 1, Model 3 Intervention Standardized Regression Weights*



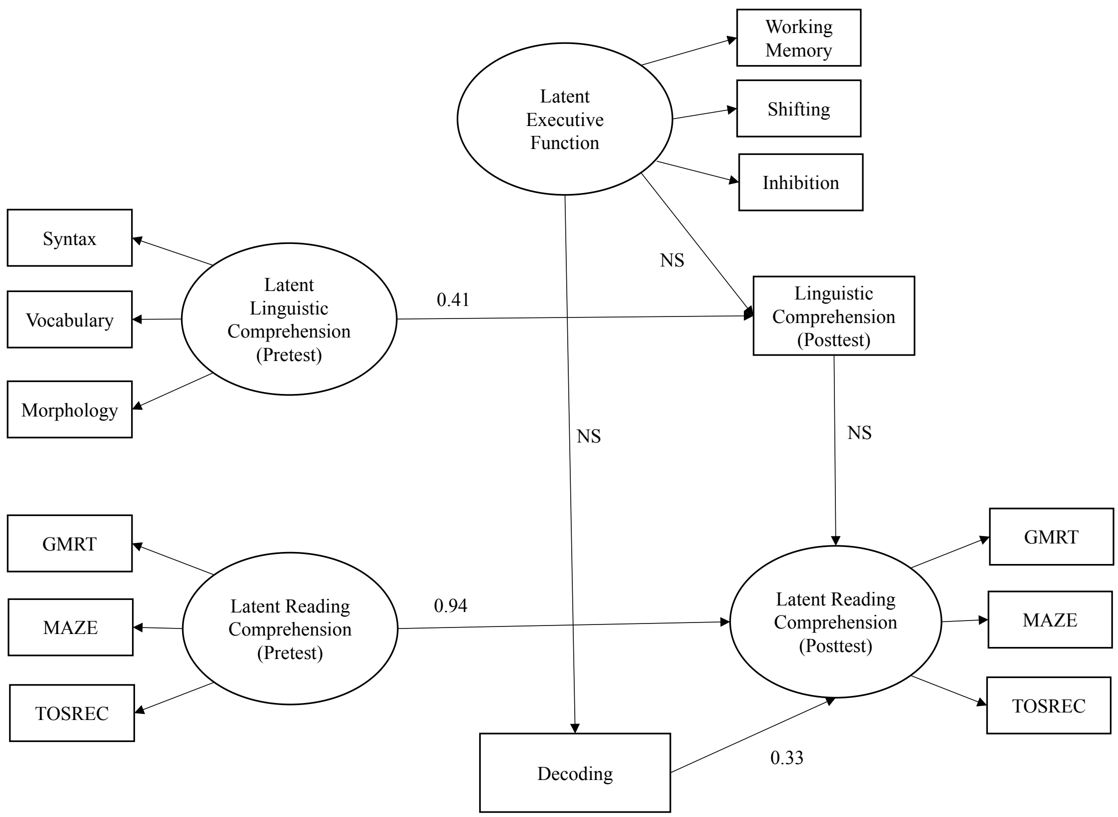
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.5 *RQ 2, Model 3 Control*EL Standardized Regression Weights*



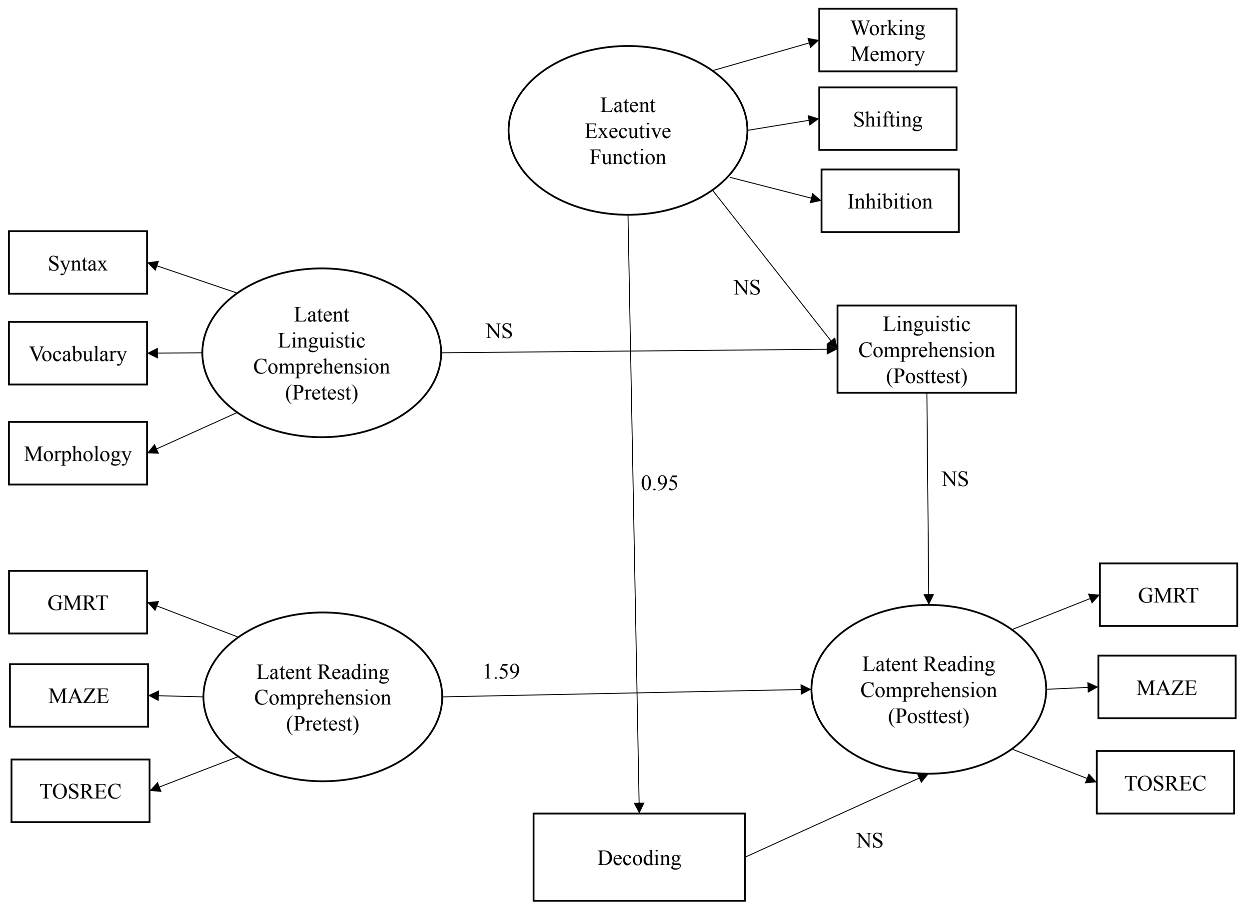
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.6 *RQ 2, Model 3 Control*REL Standardized Regression Weights*



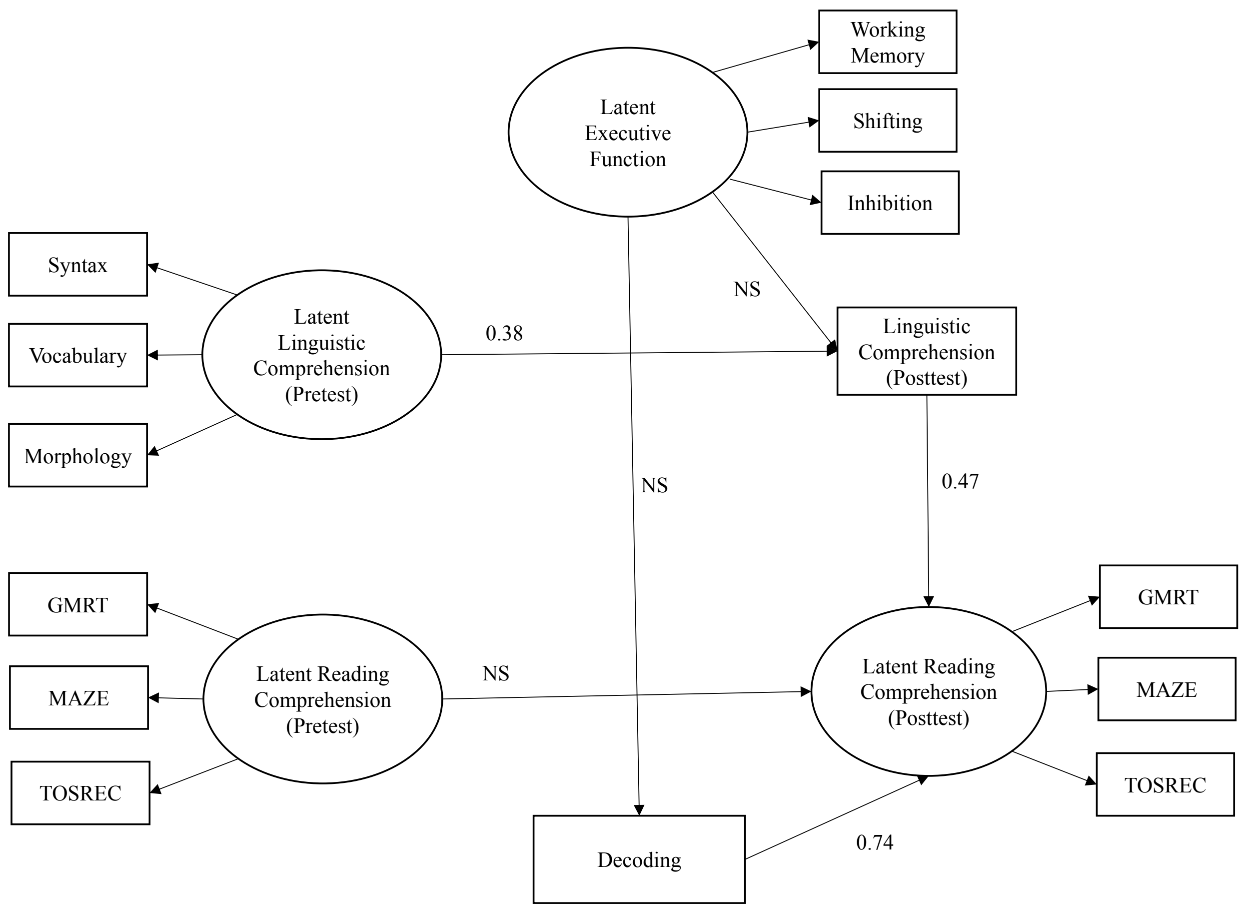
Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.7 *RQ 2, Model 3 Intervention*EL Standardized Regression Weights*



Note: Standardized regression path weights shown significant at the p = .05 level.

Figure 4.8 *RQ 2, Model 3 Intervention*REL Standardized Regression Weights*



Note: Standardized regression path weights shown significant at the p = .05 level