A TERTIARY LEVEL INTERVENTION FOR
AT-RISK ENGLISH LANGUAGE LEARNERS: PROOF-OF-CONCEPT

A Dissertation in
Special Education

by
Michelle J. Cook

© 2021 Michelle J. Cook

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2021
The dissertation of Michelle J. Cook was reviewed and approved by the following:

Elizabeth Hughes  
Associate Professor of Education (Special Education)  
Dissertation Advisor  
Chair of Committee

Paul Riccomini  
Associate Professor of Education (Special Education),  
Coordinator of Online Programs

Tracy Raulston  
Assistant Professor of Education (Special Education)

Amy Crosson  
Assistant Professor of Curriculum and Instruction

Pamela Wolfe  
Associate Professor of Special Education  
Professor-In-Charge
ABSTRACT

Reading is a pivotal skill for all students. However, there is a dearth of research examining interventions for adolescent ELLs who are struggling readers in their L2. The purpose of this study was to examine what are the changes in undergraduate ELLs’ knowledge of English morphology, decoding, and general reading comprehension in response to a reading intervention grounded in morphological, alphabetic, and phonological awareness. This AB (pretest-posttest) design study involved three participants from two different L1 backgrounds (Arabic and Chinese). Descriptive statistics are reported according to both proximal and distal variables. Results indicated that: (a) all three participants experienced growth in decoding skills, (b) two participants experienced growth in knowledge of English morphology, and (c) one participant experienced growth in general reading comprehension. The intervention was found to be socially valid by the participants. Implications related to intervention dosage and experimental design are discussed in relation to future iterations of this study.

Keywords: morphological awareness, alphabetic principle, phonological awareness, reading comprehension, response-to-intervention, ELLs, adolescent, undergraduate
# TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. viii

LIST OF TABLES .................................................................................................................. ix

ACKNOWLEDGEMENTS ....................................................................................................... x

FUNDING ............................................................................................................................... xi

INTRODUCTION ..................................................................................................................... 1

- Special education and ELLs .............................................................................................. 1
- Special Education and ELLs in the Elementary Grades .................................................. 2
- Special Education and ELLs in Middle and High School ............................................ 2
- Post-Secondary Transition and ELLs ............................................................................... 3
- Students with Interrupted Formal Education .................................................................. 5

Research on Reading Specific to ELLs ................................................................................ 7

- Learning to Read: A Developmental Trajectory ............................................................ 7
- Reading Interventions: Language of Instruction .......................................................... 8
- Reading Interventions for ELLs: Review of the Research Literature ......................... 8
- Reading Interventions for ELLs: Elementary and Middle School ............................... 10
- Reading Interventions for ELLs: High School ............................................................. 11

What Reading Intervention Components are Effective for use with ELLs at Risk of Reading Difficulties? ........................................................................................................... 13

A Tertiary Level Reading Intervention for At-Risk ELLs: Proof-of-Concept .................. 14

Theoretical Underpinnings ............................................................................................... 14
Intervention Components………………………………………………………..15

Decoding…………………………………………………………………………15

Phonological Awareness…………………………………………………………16

Morphological Awareness………………………………………………………17

Response-to-Intervention………………………………………………………..18

Digital Delivery of Reading Interventions……………………………………..19

Rationale for Participant Selection……………………………………………19

Contribution to the Literature……………………………………………………20

METHOD………………………………………………………………………………………23

Participants……………………………………………………………………………..23

Setting………………………………………………………………………………….24

Independent Variable………………………………………………………………25

Dependent Variable…………………………………………………………………25

Measures (Proximal DVs)……………………………………………………………..25

Knowledge of English Morphology………………………………………………25

Decoding…………………………………………………………………………………26

Measures (Distal DVs)…………………………………………………………………26

General Reading Comprehension…………………………………………………26

Experimental Design…………………………………………………………………27

Procedures………………………………………………………………………………27

Pretest……………………………………………………………………………………27

Test of Knowledge of English Morphology………………………………………28

Decoding…………………………………………………………………………………28

TOSCRF-2………………………………………………………………………………..28
Intervention…………………………………………………………………………………...28
Posttest……………………………………………………………………………………..31
Test of Knowledge of English Morphology………………………………………31
Decoding…………………………………………………………………………………...32
TOSCRF-2………………………………………………………………………………...32
Interrater Reliability and Treatment Integrity……………………………………32
Social Validity………………………………………………………………………………33
Descriptive Statistics……………………………………………………………………33
RESULTS……………………………………………………………………………………34
Knowledge of English Morphology…………………………………………………34
Decoding…………………………………………………………………………………...34
Reading Comprehension……………………………………………………………35
Social Validity………………………………………………………………………………36
DISCUSSION……………………………………………………………………………….37
Implications for Practice…………………………………………………………….41
Implications for Research……………………………………………………………41
Limitations…………………………………………………………………………………..42
CONCLUSION……………………………………………………………………………..44
REFERENCES……………………………………………………………………………45
Appendix A: Reading Interventions for English Language Learners at Risk of Reading Failure………………………………………………………………………………...58
Appendix B: Is a Phonologically Based Reading Intervention Functionally Related to Improved Reading Comprehension for At-Risk Undergraduate English Language Learners?…………………………………………………………………………104
LIST OF FIGURES

Figure 1-1: Student view of reading intervention (morphological analysis)..........................30

Figure 2-1: Student view of reading intervention (phonological analysis) ..........................31
LIST OF TABLES

Table 1-1: Summary of Descriptive Statistics. ................................................................. 145
Completing this dissertation and doctoral program has been a lifelong dream and it would not have been possible without the guidance and support of many people.

To my husband: Thank you! You have always been a tireless supporter of my dreams and I cannot imagine having been able to reach this milestone without you in my corner. You have been there from my undergraduate studies, to my years in the classroom, and now my graduate work. It has been a wonderful adventure.

To my chair, Dr. Elizabeth Hughes: Your kindness, support, and unwavering willingness to share your expertise has been fundamental to my successful completion of the doctoral program. I will forever be grateful to you for all that you have done to support me throughout this process. You exemplify the meaning of a gifted and dedicated educator.

To my committee members, Drs. Riccomini, Raulston, and Crosson: Thank you so much for serving on my dissertation committee. Your advice, feedback, and support are very much appreciated! Completing a dissertation during a pandemic presented unique challenges and I cannot have asked for more encouraging mentors to help me navigate this process.

To my family: Thank you! The support and encouragement that you have shown me is what has allowed me to continue on this journey.

Finally, to the many students that I have been fortunate to work with over the years. You made being an educator an incredibly rewarding profession. It is because of your desire to learn that I have been motivated to continue with my studies.
FUNDING

Assessment materials for this study were partially funded by the Penn State College of Education Dissertation Initiation Grant.
INTRODUCTION

Reading comprehension is a pivotal skill for all students. Students who become competent readers have access to content knowledge (Carnine & Carnine, 2004; Vaughn, Fall, et al., 2019) and are better positioned to succeed academically (Hammer et al., 2014; Klass et al., 2020). Indeed, studies have demonstrated a direct and significant relation between reading proficiency and postsecondary academic achievement (e.g., Cox et al., 2014; Vaughn, Fall, et al., 2019). Conversely, students who struggle to become capable readers are not only at-risk in terms of educational outcomes, but also face precarious futures beyond the classroom, extending into their ability to function in society at large (Alexander, 2005; Klass et al., 2020). This situation is even more acute for ELLs who are struggling with reading comprehension as research indicates that they may represent one of the most at-risk subgroups of students (Vaughn, Martinez et al., 2019). Therefore, when examining how to efficaciously provide reading remediation for English language learners (ELLs) who are struggling with reading in their second or additional language (L2), the issue is closely tied to equity in terms of access to both curriculum content and post-graduation success.

Special Education and ELLs

There is significant debate in the research literature relating to the relative representation of minority students in special education where some scholars have reported an over-representation of minority students in special education programs (Harry & Fenton, 2016; Maydosz, 2014) while others have argued that minority students have experienced reduced access to the specialized education services that they require (Morgan et al., 2015). Furthermore, in a study examining disproportionality specifically in relation to ELLs in special education, Sullivan (2011) noted that ELLs were at a greater risk of being overrepresented in high-incidence disability categories (i.e., specific learning disability, mild intellectual disability, and speech or
language impairment). When examining placement patterns, scholars have noted that students for whom English is not their first language are frequently overlooked for special education services in the early years (i.e., K-2) and over-represented starting in the third grade (Samson & Lesaux, 2009). This is problematic given that ELLs who begin schooling (i.e., kindergarten) with low language proficiency in their L2 frequently experience an achievement gap in reading in the early grades that grows over time (Kieffer, 2008).

**Special Education and ELLs in the Elementary Grades**

The accurate assessment of reading skills and the timely placement of ELLs in special education remains problematic (Solari et al., 2014). Given that 56% of ELLs considered for special education services seek remediation for reading problems (Gyovai et al., 2009), it is concerning that ELLs are frequently not afforded special education services prior to the third grade (Samson & Lesaux, 2009). This delay in the provision of targeted services is problematic because the focus of a significant amount of research has been reading interventions for students in the early elementary grades (Vaughn, Martinez, et al., 2019; Wanzek et al., 2016).

**Special Education and ELLs in Middle and High School**

When considering assertions that ELLs may be overrepresented in special education in the older grades (Samson & Lesaux, 2009), it is disconcerting that there is a lack of interventions that are efficacious with adolescent struggling readers (Brasseur-Hock et al., 2011). Further to this point, Vaughn, Martinez et al. (2019) reported that ELLs in high school “with significant reading difficulties may represent one of the most at-risk subgroups in American schools” (p. 373). Yet, despite their high-risk status, these students have not been the focus of reading intervention research (Vaughn, Martinez et al., 2019). Vaughn, Martinez et al. (2019) argued that the need for efficacious interventions for these students is all the more pressing given the limited time remaining for academic support. Additionally, many ELLs may enroll in an English-based education system later in their academic careers as a result of factors such as immigration or
refugee status (Oxley & de Cat, 2019). As such, these students may not have had access to reading interventions during their elementary school years (Oxley & de Cat, 2019) and may experience additional risk factors such as low socioeconomic status (Kieffer, 2011).

**Post-Secondary Transition and ELLs**

Post-secondary outcomes for ELLs are an important consideration as their academic trajectories provide information regarding the efficacy of their K-12 educational experience. Additionally, having access to post-secondary education is a key factor in terms of economic prosperity and stability for individuals including ELLs (Flores & Drake, 2014). Herein, data shared pertains to all ELLs and not solely those who have been identified as having a disability. This more global approach is necessitated by the fact that Federal law does not require state educational agencies “to collect or report data for ELLs with disabilities as a separate reporting category” (Bethea & Stevenson, 2017, p. 59). Instead, data regarding ELLs with disabilities are divided between the disparate categories of Limited English Proficiency and Students with Disabilities (Bethea & Stevenson, 2017).

Prior to discussing post-secondary transition for ELLs, it is important to acknowledge that 17% of ELLs do not complete high school (Kanno, 2018) and thus face the significant obstacle of navigating the labor market without a high school diploma. However, of those ELLs who do graduate from high school, Kanno (2018) stated that they have the possibility of pursuing one of three paths: (1) college, (2) career, or (3) neither. Although it is reasonable to assert that the K-12 educational system likely did not meet the needs of students who did not complete high school, an examination of how ELLs fair within each path following graduation is also necessary to develop an understanding of whether the K-12 education system is meeting their academic needs.

Firstly, many researchers have argued that ELLs’ access to post-secondary education is limited (Kanno, 2018; Kanno & Cromley, 2015; Kanno & Kangas, 2014). This lack of access is
evident in the admissions data for two-year colleges and even more pronounced for four-year colleges (Kanno & Kangas, 2014). Indeed, Kanno and Cromley (2015) highlighted the contrasting situation where the K-12 ELL population has continued to increase, however, this growth has not translated to admissions into four-year college programs. In fact, monolingual students have far exceeded ELLs in terms of admission to four-year colleges with rates of 43% versus 18% respectively (Kanno & Cromley, 2015). This inequitable access is exacerbated by the fact that there has been a lack of research into the disparity described above (Oropeza et al., 2010). Additionally, Kanno and Cromley (2015) reported that colleges have not been transparent in terms of their reporting of ELL admissions and overall success in their programs. However, relevant to the research presented herein, lack of access has been attributed in part to deficits in reading proficiency (Kanno & Cromley, 2015) in addition to an overall deficient academic preparation provided in high school (Kanno & Kangas, 2014). Even for ELLs who do successfully gain access to post-secondary education, challenges remain. In a study involving ELLs who transitioned from surrounding school districts to a local Canadian university, it was determined that these students were generally “inadequately prepared for the literacy demands of university” (Roessingh & Douglas, 2011, p. 285). This lack of preparedness in the area of literacy prompted the researchers to speculate that the ELLs in their study faced “immediate academic risk” (Roessingh & Douglas, 2011, p. 285). Indeed, Roessingh and Douglas (2011) noted that conditions in institutions of higher education such as a lack of differentiated instruction and scaffolded supports combined with large class sizes and reading materials that are advanced and require a mastery of academic language placed ELLs at risk. In fact, Roessingh and Douglas (2011) postulated that the average freshman textbook is written at a grade equivalent reading level of 20 whereas the average freshman ELL is reading at a grade equivalent level of nine. Most notably, Roessingh and Douglas (2011) argued that this discrepancy is not unique to the geographic location of their study, but is common among most developed nations. Additionally,
given that Roessingh and Douglas (2011) argued that age of arrival for ELLs does not appear to impact the challenges ELLs face when transitioning from high school to university, such findings have implications for all of K-12 programming.

In terms of being career-ready following graduation from high school, Kanno (2018) explained that some ELLs opt to pursue career and technical education. Concerns have been raised that many ELLs arrive at career and technical schools lacking the language and literacy skills required to succeed (Hernández-Gantes & Blank, 2009). However, Kanno (2018) expressed an even greater level of concern for ELLs who are not prepared to transition either to a two-year or four-year college nor a training program oriented toward a career or technical education. Deil-Amen and Deluca (2010) termed this group of students as the “underserved third.” These researchers posited that ELLs frequently comprise this group of students who finish high school with very low academic or marketable skills and lamented that they are likely to be poorly positioned to enter the labor force (Deil-Amen & Deluca, 2010).

**Students with Interrupted Formal Education**

When examining the literacy outcomes and transition data for ELLs, it is also important to understand that some within this group may be students with interrupted formal education (SIFE). Gutierrez (2018) defined SIFE as follows:

A group of English language learners (ELLs) who arrive at schools in the United States after the second grade with a limited or sporadic education due to a variety of factors that have impeded their development of literacy skills which commonly causes them to perform at least two years below their expected grade levels. (p. 4735)

Custodio and O’Loughlin (2020) estimated that between 10% and 20% of recently arrived ELLs may be SIFE. The challenge for educators is that SIFE may have reduced literacy skills in their L1, deficits in academic knowledge, and social or emotional concerns (Custodio & O’Loughlin, 2020, p. 9). Thus, SIFE represent a subpopulation of ELLs who have unique educational needs
(Cohan & Honigsfeld, 2017) and meeting these needs may be particularly challenging for educators (Custodio & O’Loughlin, 2020). As such, SIFE are at an increased academic risk (Hos, 2020). However, there is no formal category for SIFE students in most state departments of education and the research base targeting this specific subpopulation of ELLs is extremely limited (Cohan & Honigsfeld, 2017; Gutierrez, 2018). Exacerbating the issue is the fact that parents of students who might qualify as SIFE may be reticent to disclose this information as their interrupted schooling may have been the consequence of sensitive matters such as poverty, learning issues which precluded the student from attending school (Custodio & O’Loughlin, 2020), or war (Cohan & Honigsfeld, 2017). Gutierrez (2018) suggested that SIFE can be identified in the following manner: (a) lack of prior school transcripts, (b) gaps in education records, and (c) low literacy levels in L1. Cohan and Honigsfeld (2017) argued that SIFE have distinct “academic, linguistic, cultural, and socioeconomic challenges as newcomers to the United States” (p. 168) which may necessitate additional considerations when providing educational interventions for this group of students. It is important to note that SIFE students may require social-emotional support in addition to academic interventions (Hos, 2020). However, for the purposes of this study, the focus is placed on academic considerations. Academic considerations recommended by Cohan and Honigsfeld (2017) for this subgroup of ELLs include individualized instruction, the use of baseline data to inform instruction, access to content area curriculum, personally motivating instruction that engages with students’ lived experience, and extended graduation timelines (Cohan & Honigsfeld, 2017). Additionally, Gutierrez (2018) emphasized the need to capitalize on students’ funds of knowledge and engage in culturally responsive teaching to best meet the needs of SIFE. Of particular relevance to the study described herein, Gutierrez (2018) recommended “methods that support phonemic awareness and reading comprehension within meaningful contexts” (p. 4737) when planning academic interventions for SIFE.
Research on Reading Specific to ELLs

Learning to Read: A Developmental Trajectory

Prior to examining research relating to reading interventions for ELLs, it is important to establish whether the developmental trajectory for learning to read is the same for ELLs in their L2 when compared with their native English peers who are learning to read in their L1 (first language). According to the research literature, early reading skills including word and text reading competence are equally attainable for both ELLs and native English speaking populations and achievement levels are similar between the two groups in response to effective instruction (Lovett et al., 2008). As such, Lipka et al. (2005) asserted that the mere fact that ELLs are learning to read in their non-native language does not represent a risk factor for reading disability. Further to this point, Kieffer (2011) found that language status did not predict risk profile for reading outcomes, but rather low socioeconomic status (SES).

Some divergence in the findings does exist with regard to the development of early phonological processing skills when comparing ELLs to their native English-speaking peers (Lovett et al., 2008). According to Lovett et al. (2008), some studies have shown that ELLs in the early elementary grades achieve comparable levels of phonological skills when compared to their native English-speaking peers (Chiappe et al., 2002; Lipka et al., 2005). Whereas Lovett et al. (2008) reported that other studies have demonstrated a delay in the emergence of phonological processing skills for ELLs (e.g., Lipka & Siegel, 2007). However, despite a possible lag in phonological and linguistic processing, a parallel developmental process in terms of the acquisition of basic literacy skills was found (Chiappe et al., 2002). Notably, Chiappe et al. (2002) asserted that for both ELLs and native English speakers, alphabetic knowledge and phonological processing remained salient early literacy skills and concluded that literacy acquisition is likely to follow a similar developmental trajectory for both groups of students. One exception for ELLs may be that alphabetic knowledge may emerge before phonological
awareness and in turn facilitate the development of the latter (Chiappe et al., 2002).

Generally, the research indicates that there is parity in terms of the development of early reading skills in ELLs and native English speakers (Lesaux & Siegel, 2003) and English proficiency has not been shown to be a crucial factor in terms of developing early reading skills (Lesaux & Siegel, 2003). With regards to instructional approaches, Tam et al. (2006) noted that there is evidence to suggest that pedagogical practices that have been shown to be efficacious in terms of teaching reading to native English speakers are also effective for ELLs. However, these authors caution that some modifications might be necessary to address the unique learning needs of ELLs (Tam et al., 2006).

**Reading Interventions: Language of Instruction**

It is also important to address the issue of language of instruction when delivering reading interventions. Although some schools whose ELL populations are homogeneous (e.g., Spanish speaking) have employed bilingual instructional models, schools where the students identify as speaking a vast array of native languages, have primarily provided instruction in English (Calderon et al., 2011). Fitton et al. (2018) argued that due to a lack of empirical data, it is not possible to make conclusive assertions related to optimal intervention delivery in terms of language of instruction (i.e., English, L1, or bilingual). However, Kamps et al. (2007) reviewed the extant research and found evidence to support the use of English interventions. They asserted that instruction of reading skills delivered in English is associated with positive reading outcomes for ELLs, irrespective of their L1.

**Reading Interventions for ELLs: Review of the Research Literature**

Sullivan (2011) postulated that inadequate instructional and assessment practices are contributory causes of inequities related to special education placement. Furthermore, Sullivan (2011) argued that faulty instructional practices can impact the validity of assessment measures used to determine student eligibility for special education.
Given the significant consequences related to the successful acquisition of reading skills, educational research has focused on determining the best pedagogical practices to teach literacy. However, the majority of early literacy research has focused on teaching reading to native English speakers (August et al., 2009; Gyovai et al., 2009; Hur et al., 2020). The findings of such studies were synthesized by the National Reading Panel (NRP, 2000) and the following five areas of knowledge were determined to be fundamental to reading success: (a) phonological awareness (PA), (b) alphabetical principle (AP), (c) fluency, (d) text comprehension strategies, and (e) vocabulary knowledge (Ehri et al., 2001). Additionally, extant literature indicates that explicit instruction represents a favored pedagogical approach when working with students who are struggling to read (Tam et al., 2006).

Although much is known about teaching reading to native English speakers, there is a dearth of knowledge relating to the instructional components that contribute to positive reading outcomes for ELLs (Hammer et al., 2014; Linan-Thompson et al., 2003; Vaughn, Martinez, et al., 2019). This insufficiency exists despite the fact that ELLs represent the fastest growing subgroup within the student population in North America (Wagner et al., 2005) and that it is predicted that by 2030, up to 40% of the student population in the United States may be ELLs (U.S. Department of Education, 2003). According to the U.S. Department of Education National Center for Education Statistics (2017), in 2014-2015 approximately 665,000 ELLs were identified as students with disabilities, amounting to 13.8% of the total ELL population enrolled in U.S. public elementary and secondary schools. Given that 56% of ELLs considered for special education services seek remediation for reading problems (Gyovai et al., 2009), there is a need to extend the research to address the learning of ELLs who are at risk of reading failure.

Two meta-analyses have been conducted to date in order to determine the effects of reading interventions for ELLs. The first was conducted by Richards-Tutor et al. (2016) and involved a total of 12 studies. These researchers included only studies that were published in
peer reviewed journals that involved participants who were deemed to be at risk of reading difficulties. These researchers reported a small effect for phonics/word reading (Hedge’s $g = 0.33$), reading fluency (Hedge’s $g = 0.28$), and passage comprehension (Hedge’s $g = 0.22$).

Ludwig et al. (2019) published a more recent meta-analysis to determine the efficaciousness of reading interventions for ELLs. These authors located a total of 26 studies that included both pretest and posttest data. It is worth noting that 27% of the included studies were unpublished dissertations. Ludwig et al. (2019) reported that the effect on reading accuracy and reading fluency for the reading interventions was large ($d = 1.221$ and $d = 0.802$ respectively). With regards to reading comprehension, the researchers calculated a moderate effect ($d = 0.449$). The higher effect sizes reported by Ludwig et al. (2019) may be due to the fact that, unlike Richards-Tutor et al. (2016), these researchers did not limit the inclusionary criteria to at-risk ELLs.

Oxley and de Cat (2019) conducted a systematic review of language and literacy interventions published since 2014 that have been used with either children or adolescents who were ELLs. However, they did not emphasize studies with at-risk participants. Of the 25 included studies, 15 focused on literacy. In terms of reading comprehension, the effects reported in these studies ranged from none to medium. These researchers postulated that the reduced effect sizes may be due either to the short duration of the interventions or the complex nature of the outcome variable (i.e., reading comprehension).

**Reading Interventions for ELLs: Elementary and Middle School**

Hur et al. (2020) conducted a systematic review of English early literacy interventions that have been used with students who were designated as dual language learners (DLLs). These authors defined DLLs as children who are learning a second language and their L1 either simultaneously or sequentially (i.e., after the age of 3). Hur et al. (2020) located a total of 25 studies that fit their inclusionary criteria. These researchers reported that the 96% of the
participants were either 4 or 5 years old and that only a fraction of the group experimental design studies (25%) included participants who were identified as at risk of either speech or language delays or disabilities. However, although representing far fewer participants overall, a greater percentage (78%) of the single case experimental design studies included students with delays or disabilities. Generally speaking, the participant pool was largely homogeneous in that 23 of the 25 included studies involved participants who reported Spanish as their L1. Additionally, the majority of the studies (72%) included participants whose families were identified as experiencing low SES. Hur et al. (2020) focused their synthesis on four outcomes: (a) receptive vocabulary, (b) expressive vocabulary, (c) language, and (d) code-related skills. They determined that both the bilingual and English-only interventions reviewed had a positive effect on participant outcomes. However, they noted that the findings should be interpreted with caution due to the homogeneous nature of the L1 status of the overall participant pool.

When examining the effects of reading interventions for ELLs in elementary and middle school, it is noteworthy that of the 26 studies analyzed by Ludwig et al. (2019), 54% (n = 14) involved students in kindergarten and Grade One. Thus, corroborating assertions of an emphasis in the research literature on the earlier grades with regards to reading intervention studies. Furthermore, the extant research on the effects of reading interventions for ELLs in Grades 4 through 8 shows that their overall effect on reading comprehension is not overly positive (Vaughn, Martinez, et al., 2019).

**Reading Interventions for ELLs: High School**

Vaughn, Martinez et al. (2019) reported a lack of research examining reading interventions for ELLs in the upper grades (i.e., Grades 9 through 12). In fact, Vaughn, Marinez et al. (2019) noted that in a guide presenting the research on reading interventions for ELLs produced by the Institute of Education Sciences, the high school grades were omitted in response to a lack of data (see Baker et al., 2014). Similarly, Oxley and de Cat (2019) concluded that their
literature review highlighted the dearth of interventions for this student population (i.e., between the ages of 14 and 18) as only four of the total 25 studies involved students in this older age bracket. Additionally, none of the four interventions explicitly focused on reading. Indeed, two of the studies focused on language including academic vocabulary instruction with and without a focus on morphology (Crosson & Moore, 2017) and Word Generation with 5 academic words taught per week (Hwang et al., 2015). The other two studies focused on professional development that provided explicit guidance for educators to enhance student writing (Matuchniak et al., 2014; Olson et al., 2017).

A recent study by Vaughn, Martinez et al. (2019) attempted to address this deficit in the research literature by examining the effects of reading interventions for upper high school students who were ELLs. The study involved a total of 358 participants for whom English was not their first language. In terms of the participant demographic information, 89.4% reported that Spanish was their native language and 75.4% reported experiencing low SES. Additionally, 44 students qualified for special education with 93% having been identified as having an LD. The study occurred over a two-year period (Grades 9 and 10) and 26.3% of participants (n = 94) were lost to attrition by the end of grade 10. Students in the experimental condition received 50 minutes of intervention five days per week for two years. The intervention consisted of two phases. The first phase included explicit instruction in word study, fluency, and vocabulary. This phase incorporated instruction in decoding multisyllabic words by analyzing prefixes, suffixes, and vowels. Thus, phase one instruction included vocabulary development using morphological analysis. The second phase began in the second semester of year one and involved both vocabulary and comprehension skills. This phase involved units that were based on themes from science and social studies where participants were encouraged to apply the skills taught in the first phase. Graphic organizers and learning logs were also employed. The counterfactual group experienced “business as usual” instruction. Significant effects were found for sentence-level
fluency and comprehension \( (g = 0.18) \) and on a proximal measure of vocabulary \( (g = .41) \). However, minimal to no effects were found on standardized measures for word reading (sight word reading \( [g = -.02] \) and phonemic decoding \( [g = .06] \)), vocabulary \( (g = .00) \), and reading comprehension \( (g = -0.09) \). Vaughn et al. (2019b) noted the high attrition rate and the low fidelity of implementation scores as limitations of the study.

**What Reading Intervention Components are Effective for use with ELLs At Risk of Reading Difficulties?**

As noted previously, the majority of the research examining reading interventions for use with ELLs at risk of reading difficulties has been conducted with participants in the elementary to middle school grades. As such, I conducted a review of the literature (see Appendix A) to determine what researchers have reported about the components of reading interventions delivered in English associated with positive outcomes for ELLs identified as at risk of reading difficulties in elementary and middle school in order to inform the development of a reading intervention for older students with similar demographic and reading risk statuses. The findings suggest that a comprehensive instructional package that reflects the essential reading skills identified by the NRP (2000) may be effective. According to the extant literature, components of a comprehensive instructional package for ELLs at risk of failure should include (a) PA, (b) AP, (c) comprehension strategies, (d) fluency, and (e) vocabulary instruction. This finding suggests that intervention packages that have been found to be successful with native English readers are also promising practices for use with ELLs at risk of reading failure. This also applies to a delivery of instruction that is explicit. However, modifications to address the specific needs of ELLs (e.g., vocabulary instruction that includes concrete examples, visuals, and facial expressions) may be advisable in order to differentiate instruction for this population of students.

Pertinent to the research proposed herein, it is important to note that post intervention reading performance outcomes for older ELLs (i.e., middle school and higher) are mixed. For
example, two studies involving middle school students (Lovett et al., 2008; Tam et al., 2006) reported positive effects, whereas the results from a study by Denton et al. (2008) suggested that ELLs at risk of reading failure may become less responsive to intervention in the later grades. A discussion of these divergent findings is warranted given that Foorman et al. (2003) synthesized the research and found evidence to support the assertion that early intervention is most effective and that older students may become less responsive to intervention. Accordingly, Denton et al. (2008) offered an important glimpse at the resistance to treatment that may develop for ELLs who receive reading interventions in the later grades. Although the study’s data for ELLs and non-ELL participants were only disaggregated for pre-test measures and not post-test measures, it is relevant to this discussion as it emphasizes the importance of providing timely interventions for ELLs at risk of reading failure. In their study, the authors obtained disappointing results in terms of response to intervention for struggling eighth-grade ELLs. In fact, despite employing an intervention that included all of the components identified by the NRP (2000) for ELLs at risk of reading failure, outcomes for participants in the treatment condition for word recognition, comprehension, and fluency were not significantly greater than the those achieved by the control group. Moreover, they reported that over the course of the study, neither the treatment nor the control group experienced significant growth. The authors emphasized that on all reading measures, students in the treatment group did not show significantly higher results than their peers in the control condition. These findings suggest that ELLs with risk indicators who fail to receive effective early remediation are likely to become more impervious to remediation in the later grades (Gyovai et al., 2009).

A Tertiary Level Reading Intervention for At-Risk ELLs: Proof-of-Concept

Theoretical Underpinnings

The Simple View of Reading (SVR: Gough & Tunmer, 1986; Hoover & Gough, 1990) is an explanatory model for the mechanisms that are responsible for reading comprehension. As
such, reading comprehension is the result of effective decoding and listening comprehension skills (Hoover & Gough, 1990). In this model, both components are necessary to achieve reading comprehension. In light of this model for understanding reading comprehension, it is necessary to acknowledge the possible difference in the relative importance of these skills between English L1 and L2 students. Pasquarella et al. (2012) noted that research with English L1 students shows change over time in terms of the relative importance of these two skills in relation to reading comprehension. For English L1 students, listening comprehension surpasses decoding skills as an explanatory factor of reading comprehension levels toward the end of middle school. However, this situation may not be applicable to adolescent L2 reading comprehension. Pasquarella et al. (2012) examined the factors related to L2 reading comprehension in adolescent ELLs who were just beginning to learn English. These researchers conducted hierarchical regression analyses and determined that decoding, vocabulary, and the interaction between these two factors were significant predictors of reading comprehension for this population of students. Furthermore, the researchers concluded that the predictive model for reading comprehension for adolescent English L2 differed from that of adolescent English L1 students as vocabulary knowledge was the only significant predictor of reading comprehension for the latter group (Pasquarella et al., 2012). Thus, the study described herein involves a reading intervention that is grounded in the Simple View of Reading and recognizes that decoding and listening comprehension may differently impact the reading comprehension of adolescent ELLs who are struggling with reading in their L1 when compared with their English L1 peers.

Intervention Components

Decoding. Decoding is the ability to pronounce pseudowords correctly (Gough & Turner, 1986). Decoding is an important skill for ELLs as Pasquarella et al. (2014) examined factors related to reading comprehension in adolescents and determined that significant predictors of reading comprehension for ELLs included: (a) decoding, (b) vocabulary, and (c) the interaction
of decoding and vocabulary. Similarly, Brasseur-Hock et al. (2011) noted that adolescents with significantly low reading comprehension levels generally demonstrate issues in areas such as decoding and fluency.

**Phonological Awareness.** Phonological awareness is a predictor of reading disability (Geva et al., 2000). Despite the many studies that have uncovered the significant relationship between phonological awareness and reading comprehension (Wanzek et al., 2016), there is a dearth of studies investigating this relationship in older students with reading disabilities (Swanson et al., 2005). Interestingly, Swanson et al. (2005) described fMRI research (see Brookheimer [2003]) which demonstrated that similar to younger readers, brain patterns in older students with reading difficulties responded to explicit instruction in phonological awareness.

Sparks et al. (2009) conducted a longitudinal study over the course of 10 years to explore the long-term crosslinguistic transfer of skills that included phonological awareness from L1 to L2. These researchers assessed the participants’ L1 skills in the first, second, third, and fifth grades and L2 skills in the ninth grade (following two years of instruction in L2). The researchers reported that L1 and L2 skills were related and suggested that these findings represented evidence of cross-linguistic transfer of reading skills (Sparks et al., 2009).

Swanson et al. (2005) found that explicit instruction involving phonological awareness with struggling readers in the seventh grade ($N = 35$) resulted in significant improvement in terms of reading comprehension. The intervention duration was 12 weeks which consisted of daily 45-minute sessions for a total of 45 hours of time in intervention in a small group setting (i.e., 2-3 students per group). Swanson et al. (2005) determined that the phonologically-based intervention resulted in a statistically significant effect (ES = .15) and concluded that instruction in phonological awareness can lead to improved reading comprehension outcomes for ELLs in their L2.
Morphological Awareness. For ELLs, vocabulary knowledge supports reading comprehension (Aryadoust & Baghaei, 2016; August et al., 2005; Li & Kirby, 2014) and improved reading comprehension leads to the acquisition of more vocabulary (Bowers & Kirby, 2010; Stanovich, 1986). Schmitt et al. (2011) determined that a linear relationship exists between vocabulary knowledge and reading comprehension. Indeed, these researchers postulated that language learners need to be familiar with 98% of the text vocabulary in order to achieve solid comprehension of the reading (Schmitt et al., 2011).

Morphological analysis has been proposed as an efficacious strategy to enhance students’ vocabulary (Bowers & Kirby, 2010; Crosson & Moore, 2017; Pressley et al., 2007). As such, Bowers and Kirby (2010) hypothesized that the use of morphological analysis may allow for exponential vocabulary growth as it can be applied to novel vocabulary especially when compared with alternatives such as the direct instruction of individual vocabulary words. These authors defined morphological analysis as a process “in which learners break complex words into constituent meaning elements called morphemes” (Bowers & Kirby, 2010, p. 517). According to these researchers, it is through an analysis of the morphemes that make up words (i.e., bases, prefixes, and suffixes) that students can derive meaning.

Research shows that morphemic analysis increases people’s vocabulary knowledge (Baker et al., 2014; Anglin, 1993) and that vocabulary knowledge improves reading comprehension (Schmitt et al., 2011). Carlisle (2000) determined that for English L1 students in grades 3 and 5, morphological analysis skills accounted for 43% and 55% of reading comprehension variance respectively. Further to this point, in the ESL context, knowing morphemes can help young adult readers identify words and determine their meaning (Varatharajoo, 2015; Crosson & Moore, 2017). In a study of 1238 students in the sixth grade (309 of whom were ELL), Siegel (2008), determined that “morphological awareness was significantly related to reading and spelling over and above the contribution of phonological
awareness and oral language skills” (p. 15). Similarly, Kieffer et al. (2013) examined the effects of morphological analysis on reading comprehension in a study involving ELLs (N = 101) reading in their L2 in Grades 6, 7, 8. These researchers reported a “significant unique contribution of morphological awareness, controlling for phonemic decoding, listening comprehension, reading vocabulary, word reading fluency, and passage reading fluency” (p. 697).

In another study, Crosson and Moore (2017) determined that an intervention based on bound Latin roots in order to support academic word learning in three grade bands (i.e. Grades 6-8, 9-10, and 11-12) had positive results. Significant to the intervention for the study described herein, Crosson and Moore (2017) found that the largest effects were found for the students in the highest grade band.

Finally, in a quasi-experimental study by Carlo et al. (2004), 254 bilingual and monolingual students in the fifth grade participated in a vocabulary intervention that included morphological awareness. Although morphological analysis constituted only a part of the intervention which also included strategies such as the explicit instruction of vocabulary, using context, looking for L1 cognates, and knowledge of synonyms/antonyms and polysemy, the intervention was effective at improving the vocabulary of both the English only and ELL participants and improved reading comprehension outcomes for both groups of participants. These researchers hypothesized that strategies (such as morphological analysis) could help support reading comprehension when students encounter novel vocabulary (Carlo et al., 2004).

Findings such as those described above led Siegel (2008) to advise that morphological awareness interventions are required when working with students who have reading difficulties. This recommendation was echoed by Kieffer and Lesaux (2007) in relation to reading instruction for ELLs.

Response-to-Intervention
Gischlar et al. (2019) defined response-to-intervention (RTI) as “a multi-tiered service delivery model that utilizes performance data to evaluate instructional effectiveness for all students” (p. 201). Research shows that the identification of ELLs in need of Special Education services is not stable and that RTI is likely the best approach (Park & Thomas, 2012). Indeed, according to Samson and Lesaux (2009), there is a need to develop evidence-based teaching approaches to use in RTI in order to accurately identify ELLs in need of special education services.

**Digital Delivery of Reading Interventions**

Moran et al. (2008) conducted a meta-analysis which included 20 studies and 89 effect sizes to determine the overall impact of technology on reading performance of middle-school students. The researchers determined that digital tools and learning platforms can have a positive effect on reading comprehension with a weighted effect size of 0.489. For example, Dalton et al. (2011) designed a digitally delivered reading intervention that included both English L1 \(n = 75\) and L2 \(n = 31\) participants. These authors reported that the digital intervention positively impacted vocabulary development for both groups of participants and that the scaffolding of the intervention made possible by the digital learning tool allowed the English L2 participants to perform at the same level as their English-only counterparts (Dalton et al., 2011). In the current Covid-19 educational environment, educators are challenged to provide students with effective interventions that can be delivered remotely. This intervention is designed to respond to this pressing need.

**Rationale for Participant Selection**

Given that postsecondary institutions are recruiting are greater number of students who are considered to be academically underprepared (Mulvey, 2008) including immigrant and ELL students who arrived in the US as refugees (Kanno & Varghese, 2010), concerns relating to how to effectively address their learning needs are mounting. For example, Perin (2013) reviewed the
extant literature relating to the literacy skills of postsecondary academically underprepared students and concluded that there remains a need to learn more about how to address the literacy skills of these students. Additionally, Perin (2013) posited that in terms of reading comprehension, both decoding and linguistic comprehension strategies needed to be explicitly taught to postsecondary students who are academically underprepared such as those with low English language proficiency. Given that this need mirrors the assertion of Brasseur-Hock et al. (2011) who determined that adolescent struggling readers require interventions that target both word-level (e.g., decoding, fluency) and comprehension (e.g., vocabulary, facts and concepts, language structures) skills, the use of freshmen ELLs may represent an appropriate proxy participant pool to pilot the proof-of-concept study described below. In fact, Brasseur-Hock et al. (2011) argued that “Given that young adult populations often overlap with older adolescents, studies conducted with adults may be helpful in understanding reading proficiency” (p. 441).

**Contribution to the Literature**

Sullivan (2011) proposed that the disproportionality of ELLs in special education would be less concerning if questions regarding appropriate placement could be resolved. Furthermore, Solari et al. (2011) described a seemingly intractable situation where reading problems are often the primary issue for ELLs who are struggling academically and LD is frequently associated with reading problems. The development of research-based practices to provide efficacious interventions within an RTI framework for adolescent ELLs who are struggling with reading would help to address this issue. This need is all-the-more pressing in light of the research by Solari et al. (2014) which found that ELLs with LD are significantly more at risk of reading failure than any other subgroup of student (LD, ELL, and GenEd). Further to this point, Solari et al. (2014) noted that early interventions have been the focus of most research and that this does not respond to their findings that the need remains constant across grade levels. This call to action was also echoed by Pasquarella et al. (2012) who advocated the development of
empirically based instructional programs for this population of students. Thus, there is an immediate need to develop evidence-based interventions for use within an RTI framework in order to accurately respond to the learning needs of adolescent struggling readers who are ELLs. Additionally, students who enter an English-based education system (such as that of the U.S.) later in their academic careers due to immigration or refugee status may not have had access to intensive early literacy interventions that include alphabet knowledge and phonological awareness (Oxley & de Cat, 2019).

As such, a proof-of-concept investigation into the effects of a reading intervention for undergraduate ELLs (who self-identify as experiencing difficulty in the area of reading comprehension) represents an initial exploratory examination. This study builds upon a previous pilot study (Cook & Hughes, in preparation) in which a reading intervention grounded in alphabetic and phonological awareness was implemented to address reduced reading comprehension for first year postsecondary at-risk ELLs when reading in their L2 (see Appendix B). In this previous study, there was a small effect found for the participants’ decoding skills, however, improvements in decoding did not result in a discernable improvement in reading comprehension. This present study proposes the addition of a morphological awareness component to address the role of novel vocabulary in relation to reading comprehension.

In this study, a reading intervention grounded in morphological, alphabetic, and phonological awareness was implemented to address reduced reading comprehension for first year postsecondary ELLs when reading in their L2. The research questions are as follows:

**RQ1:** What are the changes in undergraduate ELLs’ knowledge of English morphology in response to a reading intervention grounded in morphological, alphabetic, and phonological awareness?
**RQ2:** What are the changes in undergraduate ELLs’ decoding ability in response to a reading intervention grounded in morphological, alphabetic, and phonological awareness?

**RQ3:** What are the changes in undergraduate ELLs’ general reading comprehension in response to a reading intervention grounded in morphological, alphabetic, and phonological awareness?
METHOD

Participants

Participant recruitment occurred in two sections of a foundational freshman English course at a midsized undergraduate university campus. Both sections of the course were taught by the same Associate Teaching Professor. The study occurred during the Covid-19 pandemic and both sections were taught remotely via Zoom. The delivery mode for the classes was a combination of synchronous and asynchronous instruction. I introduced the study to the students live via Zoom during their regular instructional time. Participant recruitment was conducted in accordance with the university’s Institutional Review Board’s procedures as follows:

1. In order to ensure that students did not feel pressure to participate in the study, students were informed that their decision to participate would not impact their grades in the course.

2. Incentive to participate was provided as participants who completed their respective components of the study (e.g., viewing the intervention videos, completing assessments) were entered into a draw for two $25.00 gift cards to Best Buy.

3. Students were informed that if they opted to participate and were randomly assigned to the control condition, they would receive all intervention materials (i.e., modules) following the final posttest (due to changes in the research design, all students received treatment).

4. Prospective participants were provided with my contact so that they could ask any outstanding questions.

5. Implied consent was obtained. Following the in-class (i.e., Zoom) recruitment presentation, the course instructor sent the recruitment information via Penn State email to all of the students’ university email addresses. Potential participants were instructed to email the investigator who would then email them the consent forms and information
about the first time they would meet. Consent forms stated that by participating, the
students were giving implied consent to be a part of the study. As such, when the
potential participants set up a time to meet with the investigator, this was considered as
them having given implied/verbal consent.

6. All identifying information was disaggregated from participant data through the use of
identification codes before the data was analyzed.

Inclusionary criteria were as follows:

1. The student is enrolled in an undergraduate foundational English course at Penn State.
2. The student is 18 years of age or older.
3. The student is an ELL (i.e., the student’s L1 or home language is a language other than
English).
4. The student self-reports as experiencing difficulty with English reading comprehension.
5. Be exposed to English at school for more than 3 years.
6. The student has consented to participate in the research study.

The Covid-19 pandemic had a significant impact on participant recruitment efforts as a
total of three students consented to participate in the present study. In terms of the participants’
L1, Participant A spoke Arabic and Participants B and C spoke Mandarin. The participants were
male and they ranged in age from 18 to 23. The three participants self-reported as experiencing
difficulty with reading comprehension in English.

Setting

This study was conducted through a university campus in the northeastern part of the
United States. The campus is located just on the periphery of a medium-sized urban center in a
suburban setting. The university campus offers both undergraduate and master level courses. At
the time of the study, the campus was offering a mix of in-person and online courses in response
to the Covid-19 pandemic. Participant A remained within the U.S. throughout the study,
Participant C started the intervention while in the U.S., but left the U.S. partway through the study, and Participant B remained outside of the U.S. for the duration of the study. Participants viewed the reading intervention videos online outside of regularly scheduled class hours. Thus, students reported participating in the reading intervention using either (a) their personal computers or IPads in their dorm rooms or other private residences or (b) the campus computers located in the library or computer laboratory.

**Independent Variable**

The reading intervention consisted of prerecorded PowerPoint lessons and narration. The intervention videos were recorded using Kaltura technology and emailed directly to the participants’ Penn State email addresses. The lessons consisted of two parts: (a) morphemic analysis and (b) alphabetic and phonological awareness. Morphemic analysis instruction followed procedures recommended by Kieffer and Lesaux (2007) and alphabetic and phonological awareness procedures were similar to those recommended by the Corrective Reading program (Englemann et al., 2008).

**Dependent Variable**

The proximal dependent variables (DVs) were decoding and knowledge of English morphology. The distal DV for this study was general reading comprehension. General reading comprehension was operationally defined as the ability to extract meaning from English text.

**Measures (Proximal DVs)**

**Knowledge of English Morphology.** In order to assess the participants’ knowledge of English morphology both at pretest and posttest, a Test of Knowledge of English Morphology adapted from Carlo et al. (2004) and Carlisle’s (1988) Extract-the-Base task was administered. Similar to Carlo et al. (2004) the assessment consisted of 27 items (with less than a third of the words drawn from the intervention) and participants were asked to identify the base form of a derived word. Following the procedures described by Carlo et al. (2004), participants were told
the derived word (e.g., discussion) followed by a sentence providing context (e.g., What did he want to ______?) Participants were required to write their answer in the space provided. The items varied as follows:

- No phonological or orthographic change
- A phonological change but no orthographic change
- An orthographic change but no phonological change
- Both changes

A coding scheme from Carlo et al. (2004) and informed by Bear, Invernizzi, Templeton, and Johnston’s (1996) spelling rubrics was used to score student responses. Scores ranged from 5 points for answers that included the correct base and were correctly spelled to 0 for completely incorrect responses.

Decoding. Pre and posttest materials were drawn from the *Woodcock Johnson IV Tests of Achievement* (Schrank et al., 2014). In particular, the subtest *Test 7 Word Attack* was used as a measure of decoding. Pretest consisted of *Form A Standard* and Posttest consisted of *Form B Standard*. Participants were asked to read nonwords presented in printed form aloud. A cut-off rule of six consecutive errors was applied. The participants’ scores were the number of correct items. This subtest was selected to assess participant decoding skills because it is well standardized and extensively used in educational settings (Lovett et al., 2008).

Measures (Distal DVs)

General Reading Comprehension. Pre and posttest measures of reading comprehension and general reading ability consisted of the Test of Silent Contextual Reading Fluency-Second Edition (TOSCRF-2; Hammill et al., 2014). Specifically, the TOSCRF-2 measures participants’ contextual reading abilities (e.g., word identification, word meaning, word building, sentence structure, comprehension, and fluency). Participants are given short passages that consist of rows of contextually related words that are ordered by increasing levels of reading difficulty. All
words are printed in uppercase without any spaces or punctuation between words (e.g., ACATWITHWHITESTRIPES). Students are asked to draw a line separating as many recognizable words as possible within 3 minutes (e.g., A/CAT/WITH/WHITE/STRIPES). The passages increase in difficulty level (i.e., content, vocabulary, and grammar). Success on the test requires that students read the meaning of the text. The test yields: (a) raw scores, (b) standard scores, (c) percentiles, and (d) age/grade equivalents. The TOSCRF-2 has been validated for a wide variety of subgroups, as well as for the general population (ages 7 years 0 months to 24 years 11 months). Four equivalent test forms (A, B, C, D) are available for pre and post testing.

**Experimental Design**

An AB (pretest/posttest) research design was employed. All participants received four 30-minute intervention sessions over the course of approximately six weeks. Pretesting occurred before treatment sessions began with all participants and post-testing occurred following approximately a total of 120 minutes of intervention per student. All pretests and posttests were administered by the interventionist. Additionally, an intervention video comprehension quiz was administered online following each of the four intervention sessions. Intervention sessions were prerecorded and administered online at a time of the participants’ choosing during the assigned week. Treatment integrity data were collected for 20% of sessions.

**Procedures**

**Pretest**

All participants completed the pretest. These assessment activities took place via Zoom. Testing materials were displayed using a document camera and the screen share feature. When completing assessments that required the participants to write down their answers (i.e., Test of Knowledge of English Morphology and TOSCRF-2), participants were assigned remote control and used the digital pencil tool to record their responses. The sessions were recorded in order to
provide a permanent record of the participants’ responses and to permit scoring after the testing sessions ended. The assessments were administered to each student individually.

**Test of Knowledge of English Morphology.** The assessment consisted of 27 items. Participants were asked to identify the base form of a derived word. Following the procedures described by Carlo et al. (2004), participants were told the derived word (e.g., discussion) followed by a sentence providing context (e.g., What did he want to _______?) The first author read each derived word and sentence aloud in order to avoid introducing reading ability as a confounding variable. Participants were required to write their answer in the space provided.

**Decoding.** Participants were asked to read nonwords presented in printed form aloud. A cut-off rule of six consecutive errors was applied.

**TOSCRF-2.** Students were asked to read the passage silently and to place a dot between words. The students were given three minutes to complete the assessment.

**Intervention**

Students in enrolled in the foundational English course are required to read three essays upon which major class writing assignments are based. As such, a firm understanding of these essays is essential for student success in the course. Each reading intervention video targeted prefixes, suffixes, and sound blends associated with key vocabulary in each essay. In order to ensure that students viewed the reading interventions, participants were asked to complete a brief quiz following the viewing of each intervention video to ensure that they were familiar with the videos’ content. Intervention sessions took place online and were emailed weekly to the participants. The reading intervention videos were created using Kaltura technology and consisted of a PowerPoint presentation with narration. Each intervention video followed the same format:

a) Introduction to morphology.

b) Application of morphology to key vocabulary in the essay.
c) Introduction to word-attack.

d) Application of word-attack skills to vocabulary found in the essay.

Subsequent reading intervention videos followed the following format:

a) A review of morphology.

b) Application of morphology to key vocabulary in the essay the students were expected to read.

c) A brief review of word-attack.

d) Application of word-attack skills to vocabulary found in the essay.

As such, the reading intervention videos consisted of the following skills: morphemic analysis and alphabetic and phonological awareness. In the first section, students learned that morphology is the study of meaningful units of language and how those units are combined in words. Students were presented with a brief summary of research indicating that morphemic analysis skills can help build vocabulary knowledge which in turn can increase reading comprehension. Next, I modeled how to conduct a morphemic analysis using sample vocabulary. This was followed by a presentation of key vocabulary from the target essay where I modeled the steps for conducting a morphemic analysis. For example, the prefix “auto” appeared often in an essay and students learned that the prefix “auto” meant “by oneself or itself” and then reviewed key vocabulary such as “autoethnographic” and “autobiography.” A similar process was used to introduce important suffixes in the reading. Finally, I modeled a full morphemic analysis where key vocabulary was analyzed according to the prefix, root, connector, and suffix. Students were then presented with a flow chart outlining the steps to follow when reading the essay:

1. Keep the two tables of prefixes and suffixes handy.
2. When you encounter vocabulary that is new, ask yourself: (a) Can I conduct a morphemic analysis of this word and (b) Is there a prefix, root, or suffix that could help me understand the meaning of the word?

<table>
<thead>
<tr>
<th>Prefixes</th>
<th>Meaning</th>
<th>Examples from Pratt</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>“not” or “without”</td>
<td>asymmetrical</td>
</tr>
<tr>
<td>auto</td>
<td>“by oneself or itself”</td>
<td>autographic, autobiography</td>
</tr>
<tr>
<td>ethno</td>
<td>“race” or “nation”</td>
<td>ethnographic, ethnographers</td>
</tr>
<tr>
<td>re</td>
<td>“again” or “back”</td>
<td>revisionist</td>
</tr>
<tr>
<td>co</td>
<td>“together” or “with”</td>
<td>collaboration</td>
</tr>
<tr>
<td>trans</td>
<td>“across” or “connecting”</td>
<td>transcultural, transmitted,</td>
</tr>
<tr>
<td>inter</td>
<td>“between”</td>
<td>intercultural</td>
</tr>
<tr>
<td>mono</td>
<td>“one”</td>
<td>monolingual</td>
</tr>
<tr>
<td>bi</td>
<td>“two” or “twice”</td>
<td>bilingual</td>
</tr>
<tr>
<td>Homo</td>
<td>“same” or “alike”</td>
<td>homogeneous</td>
</tr>
<tr>
<td>multi</td>
<td>“many” or “more than one”</td>
<td>multicultural</td>
</tr>
<tr>
<td>Con</td>
<td>“with” or “thoroughly”</td>
<td>contact</td>
</tr>
<tr>
<td>Sub</td>
<td>“under” or “below” or “beneath”</td>
<td>subordination</td>
</tr>
</tbody>
</table>

Figure 1-1: Student view of reading intervention (morphological analysis).

The next exercises involved the introduction of recurrent sound combinations from the target essay. For example, participants were told that the letters A-I go together and make the sound áááá. The participants were then presented with a series of words containing the sound combination A-I which was underlined. The first author read the words aloud, asked the participants to repeat the words, then provided a brief pause for the student to repeat the word containing the target sound combination. Students were encouraged to discriminate between the sounds in the spoken words. This process was repeated for other sound combinations, including sound combinations that make the same sound such as I-R, U-R, and E-R.
Posttest

Procedures for the posttest mirrored those followed for the pretest. All participants completed the posttest activities. These assessment activities took place via Zoom and materials were displayed using a document camera and the screen share feature. When completing assessments that required the participants to write down their answers (e.g., Test of Knowledge of English Morphology and TOSCRF-2), participants were assigned remote control and used the digital pencil tool to record their responses. The sessions were recorded in order to provide a permanent record of the participants’ responses and to permit scoring after the testing sessions ended. The assessments were administered to each student individually.

Test of Knowledge of English Morphology. The assessment consisted of 27 items. Participants were asked to identify the base form of a derived word. Following the procedures described by Carlo et al. (2004), participants were told the derived word (e.g., discussion) followed by a sentence providing context (e.g., What did he want to _______?) The first author read each derived word and sentence aloud in order to avoid introducing reading ability as a confounding variable. Participants were required to write their answer in the space provided.
**Decoding.** Participants were asked to read nonwords presented in printed form aloud. A cut-off rule of six consecutive errors was applied.

**TOSCRF-2.** Students were asked to read the passage silently and to place a dot between words. The students were given three minutes to complete the assessment.

**Interrater Reliability and Treatment Integrity**

Interrater reliability was calculated for 30% of student assessments. Pretest and posttest assessments were scored by both the interventionist and the second rater. The second rater, who held a PhD, was trained in scoring the assessments according to the guidelines for each assessment prior to scoring the assessment samples. Practice scoring sessions were scheduled to allow the observer the opportunity to request clarifications. Interobserver agreement (IOA) between the first and second rater was 100%.

A measure of treatment integrity was developed for both intervention sessions and pretest/posttest procedures. To assess treatment integrity for intervention sessions, a checklist was developed by the interventionist (see figure Appendix E). Treatment integrity data was obtained for 20% of the digitally recorded intervention sessions. Treatment integrity was measured using a point-by-point format where the independent observer scored treatment integrity on four items that corresponded with the methods section. Items could be scored as either ‘observed’ (1 point) or ‘not observed’ (0 points). The total number of possible points per session was four. To obtain a treatment integrity score, the independent observer divided the total number of points by the total number of items. An independent scorer, who held a PhD, was trained in data collection prior to viewing the intervention sessions. Practice data collection sessions were scheduled to allow the observer the opportunity to request clarifications. Average treatment integrity was 100% for all observed sessions.

To assess treatment integrity for pretest and posttest procedures, a checklist was developed by the interventionist (see figure Appendix E). Treatment integrity data was obtained
for 20% of pretest and posttest administrations. Treatment integrity was measured using a point-by-point format where the independent observer scored treatment integrity on 10 items. Items could be scored as either ‘observed’ (1 point) or ‘not observed’ (0 points). The total number of possible points per pretest or posttest administration was 10. To obtain a treatment integrity score, the independent observer divided the total number of points by the total number of items. An independent scorer was trained in data collection prior to viewing the pretest and posttest procedures. Practice data collection sessions were scheduled to allow the observer the opportunity to request clarifications. Average treatment integrity was 100% for all observed sessions.

Additionally, data regarding student interaction with the intervention videos was collected via the administration of four intervention video comprehension quizzes. Data were collected and analyzed for treatment integrity purposes. All three participants scored 100% on Intervention Video Quiz 1. For Intervention Video Quiz 2, Participants A and C obtained a score of 100% and Participant B obtained a score of 85.71%. On Intervention Video Quiz 3, Participants A and C obtained a score of 100% and Participant B did not submit the quiz. Finally, on Intervention Video Quiz 4, all three participants achieved a score of 100%.

Social Validity

A questionnaire was developed to assess the participants’ perceptions of the social validity of the reading intervention (see Appendix C for the participant social validity form). The questionnaire included both multiple choice and open-ended questions in order to allow participants the opportunity to provide feedback. The questionnaire was emailed to the participants following the posttest.

Descriptive Statistics

Due to the reduced number of participants, descriptive statistics were calculated according to the dependent variable measured. In terms of knowledge of English morphology,
scores were reported in terms of percent correct and a calculation of change between pretest and posttest was calculated. For decoding, raw scores as well as age equivalents (AE) and grade equivalents (GE) were reported. Mean raw scores as well as change in scores between pretest and posttest were also calculated and reported in terms of GE. Finally, for reading comprehension, raw scores, age and grade equivalents, percentile rank, and a standard score reported as an index score ($M = 100, SD = 15$) were calculated. A change in index scores was also calculated.

**RESULTS**

This study involved a digitally delivered reading intervention that targeted undergraduate ELLs’ knowledge of English morphology, decoding, and general reading comprehension. Results will be presented both on a group and individual level.

**Knowledge of English Morphology**

The mean for the three participants on the Test of English Morphology at pretest was 59.01% and 58.83% at posttest. Thus, the mean change for the participants was a negative difference of .18% between pretest and posttest. However, this measure of central tendency should be viewed with caution given the variability in the data. Participant A saw a small positive improvement and obtained a score of 63.7% at pretest and 65.38% at posttest (indicating a positive change of 1.68%). Participant B saw a larger positive improvement and obtained a score of 60.74% at pretest and 77.03% at posttest (indicating a positive change of 16.29%). Conversely, Participant C saw a negative change with a pretest score of 52.59% and a posttest score of 34.07% (indicating a negative change of 18.52%).

**Decoding**

The Word Attack mean of at pretest of 17.67 for the participants corresponds roughly to an AE of 8-0 and a GE of 2-5. The Word Attack mean at posttest of 20 for the participants corresponds roughly to an AE of 9-0 and a GE of 3-6. Thus, indicating a change in GE of 1-1.
Given the limited number of participants, individual data is presented to facilitate a more targeted discussion. Participant A obtained a pretest score of 17 which corresponds to an AE of 8-0 and a GE of 2-5. At posttest, Participant A obtained a score of 18 which corresponds to an AE of 8-3 and a GE of 2-8. Thus, indicating a change in GE of 0-3. Participant B obtained a pretest score of 19 which corresponds to an AE of 8-7 and a GE of 3-2. At posttest, Participant B obtained a score of 21 which corresponds to an AE of 9-7 and a GE of 4-1. Thus, indicating a change in GE of 0-11. Participant C obtained a pretest score of 17 which corresponds to an AE of 8-0 and a GE of 2-5. At posttest, Participant C obtained a score of 21 which corresponds to an AE of 9-7 and a GE of 4-1. Thus, indicating a change in GE of 1-8.

**Reading Comprehension**

The mean raw score on the TOSCRF-2 for the three participants was 57.67 at pretest and 55.33 at posttest indicating a decrease in the mean raw score of 2.34. However, due to the variability in the scores, a measure of central tendency is not likely appropriate. As such, an examination of participants’ achievement on an individual level is warranted. On the pretest, Participant A obtained a raw score of 16 which equates to an AE of < 7-0, a GE of < 1-0, and a percentile rank of < 1. Participant A’s index score was 40 at pretest which is considered “very poor.” At posttest, Participant A obtained a raw score of 33 which equates to an AE of < 7-0, a GE of 1-0 and a percentile rank of < 1. Participant A’s index score was 42 at posttest which is considered “very poor.” Participant A’s results indicated a positive change in index score of 2. Participant B obtained a raw score of 78 which equates to an AE of 9-6, a GE of 4-5, and a percentile rank of 1. Participant B’s index score was 67 at pretest which is considered “very poor”. At posttest, Participant B obtained a raw score of 62 which equates to an AE of 8-6, a GE of 3-0, and a percentile rank of < 1. Participant B’s index score was 59 at posttest which is considered “very poor.” Participant B’s results indicated a negative change in index score of 8. Participant C obtained a raw score of 79 at pretest which equates to an AE of 9-6, a GE of 4-5,
and a percentile rank of 3. Participant C’s index score was 72 at pretest which is considered “poor.” At posttest, Participant C obtained a raw score of 71 which equates to an AE of 9-0, a GE of 3-8 and a percentile rank of < 1. Participant C’s index score was 63 at posttest which is considered “very poor.” Participant C’s results indicated a negative change in index score of 9.

Social Validity

The participants were asked to complete a short Social Validity Questionnaire following the final intervention session (see Appendix C). Six questions on the Participant Feedback Questionnaire provided information with regards to the social validity of the intervention. All three participants indicated that they found the video lessons helped them to better understand the related readings. Participants were then provided with an open-ended question where they could provide more information. The participants responded as follows: (a) “Because it made me understand the article better,” (b) “The prefix root part, help me understand the word better,” and (c) “How words are combined together. I can tell meanings of the words. Knowing roots, prefixes gives me better understanding of what it means.” Participants were then asked to indicate what parts of the intervention videos they found helpful and explain why. The participants responded as follows: (a) “the part that I found to be helpful is when the instructor talked about the suffixes and the prefixes and roots of word that is amazing,” (b) “The word learning part, you can understand the word better,” and (c) “All of it.” For the final three questions, the participants did not indicate that there were any elements of the intervention videos that they found to be less helpful or would change. Finally, when provided with the opportunity to make any additional comments, one participant indicated that he would recommend the reading intervention to friends and another asserted that “really it has been helpful.” As such, the results of the participant questionnaire support the social validity of a reading intervention grounded in morphological, alphabetic, and phonological awareness to improve the reading comprehension skills of at-risk adult ELLs within the university context.
Discussion

The objective of this study was to determine what changes a reading intervention grounded in morphological, alphabetic, and phonological awareness had on the reduced reading comprehension of first year postsecondary ELLs when reading in their L2. An examination of the participants’ results across proximal and distal variables can provide insights into the impact of the intervention on an individual level and can inform future iterations of the research. Additionally, a discussion of the findings in light of research relating to cross-linguistic transfer may elucidate factors that may be contributing to participants’ results as cross-linguistic transfer may assist in the acquisition of a second language (Yang et al., 2017).

In terms of knowledge of English morphology, Participant A saw a small increase of 1.68% between pretest and posttest. Notably, when considering this participant’s L1 (Arabic) researchers have found no correlation in morphological awareness between English and Arabic (Saiegh-Haddad et al., 2007). In fact, these researchers reported that morphological awareness plays a greater role in reading in what they termed as morphologically transparent English versus morphologically opaque Arabic (Saiegh-Haddad et al., 2007). Furthermore, Saiegh-Haddad et al. (2007) asserted that in the case of English-Arabic bilingualism “morphological awareness is primarily a language-specific linguistic skill that emerges as a function of language proficiency” (p. 495). Thus, explicit instruction in morphological awareness in English may still be warranted for this participant despite the limited progress reported herein. The insubstantial increase in knowledge of morphological awareness observed in Participant A’s results may be attributed to the truncated duration of the intervention and reflect the need for a more intensive approach to morphological awareness instruction for this participant.

In terms of knowledge of English morphology for the remaining two participants; Participant B saw an increase of 16.29% and Participant C saw a decrease of 18.52%. 
Given that both Participants B and C spoke Chinese as their L1, it may be helpful to consider their results in light of the research related to cross-linguistic transfer between Chinese and English. Lin et al. (2017) reported that in terms of morpheme awareness, there is evidence of transfer between L1 Chinese and L2 English. These authors speculated that “The abundance of compound words in Chinese may result in children’s heightened awareness of morphemes which is particularly useful in reading comprehension” (Lin et al., 2017, p. 121). These researchers asserted that transfer of morpheme awareness from Chinese L1 to English L2 is most likely when the following conditions are met: (a) well-established morpheme awareness and (b) automaticity. Similarly, Yang et al. (2017) conducted a meta-analysis and reported an overall correlation between Chinese and English morphological awareness that is small, \((r = .37, 95\% \text{ CI } [.31, .42])\), but significant, \((z(8) = 12.35, p < .01)\). That Participant B saw an improvement in knowledge of English morphology may speak to well-established previous knowledge in his L1 that may have been activated by the reading intervention. Conversely, Participant C obtained results at posttest that reflected negative growth in morphological awareness. Although it is not possible to speculate, it would be informative to assess Participant C’s morpheme awareness in his L1 to understand the role of previous knowledge on the acquisition of morphological awareness as targeted in this reading intervention. However, the most plausible explanation for Participant C’s results may be related to fatigue and differences in time zones as this participant was the only participant who experienced a significant change in time zone toward the end of the intervention.

In terms of decoding, Participant A’s results demonstrated a change in GE of 0-3. This participant’s results are of significance given that Abu-Rabia and Siegel (2002) noted that in their research, the scores on English nonsense word reading and spelling assessments for students with reading problems who were bilingual (English-Arabic) were higher when compared with English-only students with reading disabilities. These researchers postulated that this may be due to “positive transfer from the regular nature of Arabic orthography” (Abu-Rabia & Siegel, 2002, p.
As such, these authors suggested that for Arabic L1 students, cross-linguistic transfer may facilitate decoding when negotiating the deep orthography of English. Thus, a more intensive intervention may be necessary to address the decoding skills of this participant.

Participants B and C obtained posttest results that demonstrated some improvement in decoding skills with a change in GE of 0-11 and 1-8 respectively. It is worthwhile to observe that English contains more phonemes than Chinese and that in terms of orthography, Chinese and English written forms are disparate where the former is a logographical language and the latter is an alphabetical language (Yang et al., 2017). In their meta-analysis of research investigating the cross-linguistic transfer between Chinese and English, Yang et al. (2017) found a mean correlation between Chinese and English decoding skills that is moderate (r = .44, 95% CI[.41, .47]), and significant, (z (28) = 26.89, p < .01). The modest improvements in decoding skills obtained for Participants B and C may indicate a need for a more intensive decoding intervention.

In terms of general reading comprehension, Participant A saw an increase of 2 index scores. Participant B saw a decrease of 8 index scores, and Participant C saw a decrease of 9 index scores. Firstly, it should be noted that based on anecdotal observations, the declining performance of Participants B and C may be attributable to confounding factors such as fatigue and time zone differences. However, the overall lack of improvement in the reading comprehension results obtained in this study are similar to the findings reported by Vaughn, Martinez et al. (2019). These researchers found that adolescent ELLs with reading difficulties in the treatment group did not perform differentially better than participants in the business-as-usual group on standardized measures of reading comprehension despite an intensive 2-year intervention. Further to this point, Herrera et al. (2016) reviewed the research on the effectiveness of adolescent literacy interventions and determined that in the area of reading comprehension, none of the studies with rigorous research designs conducted in high schools were found to have had either positive or potentially positive effects. As such, the need to
continue research investigating efficacious supports for ELLs with reading difficulties in high school remains salient.

The participants’ results on the distal dependent variable of reading comprehension warrant further discussion in terms of accurate self-assessment of reading comprehension skills. Given that the TOSCRF-2 assesses multiple reading skills (e.g., word identification, comprehension, fluency), the assessment can be used to evaluate general reading ability. The participants in this study obtained index scores, both at pretest and posttest, that are considered to range between “very poor” and “poor.” In fact, index scores below 90 are considered an indication that the students “will likely have difficulty in all kinds of reading skills, including decoding, word identification, and comprehension skills” (Hammill et al., 2014, p. 18). Notably, limited English proficiency is listed as a factor that may contribute to a student’s low scores on the TOSCRF-2. However, given the students’ enrolment at a university, it is reasonable to assume that they possess adequate English proficiency to take the TOSCRF-2. The index scores obtained by the participants in this study indicate that the recruitment procedures which relied on participant self-assessment of difficulties with reading comprehension were accurate. Similar to the findings in a previous study by Cook and Hughes (in preparation), student performance on standardized measures of reading comprehension reflected an accurate self-assessment of reading comprehension on the part of the participant. This finding is supported by the work of Parrila et al. (2007) who determined that the use of student self-reports is a viable method for locating university students with a history of reading problems.

Another notable result is the participants’ normative scores on the TOSCRF-2 in terms of AE and GE. However, it is important to recognize that Aiken (2003) advised that age and grade norms should be interpreted with caution. The issue is that growth rate is not consistent across ages and grades in the areas of cognitive, psychomotor, and affective development and thus the use of these scores can be misleading (Aiken, 2003). Similarly, Reynolds (1981) warned that the
use of GE can overemphasize concerns on reading assessments at the upper grades. Furthermore, given the inconsistency across ages and grades described by Aiken (2003), changes in performance between pretest and posttest that involved calculating the difference in AE or GE should be understood in light of the limitations of such calculations.

Finally, it is worthwhile to note that participants’ responses on the Social Validity Questionnaire indicated that the intervention was helpful in supporting their comprehension when reading the essays that were targeted by the intervention videos. As such, the addition of a specific reading comprehension quiz related to each essay within the context of a group experimental design study might provide more insight into how this reading intervention supports specific reading comprehension for struggling ELLs.

**Implications for Practice**

The need for efficacious interventions for high school and undergraduate ELLs to address concerns related to reading comprehension is well established. Although further research must be conducted to determine a clear effect for the reading intervention piloted in this study, the components of the intervention are supported by the research literature. Additionally, the social validity of the reading intervention may encourage students in the upper grades to participate in the intervention. The fact that this intervention can be tailored to meet the immediate educational needs of students to support their reading comprehension of contextually appropriate material makes this intervention a classroom appropriate tool that can be leveraged within an RTI approach.

**Implications for Research**

The Covid-19 pandemic had a significant impact on the research design employed in this study. Future iterations of this study will involve a group experimental design with randomization. Furthermore, to address the need for more intensive supports (as suggested within the discussion section) an increase in intervention dosage will be integrated. The addition
of a control group along with specific reading comprehension quizzes related to the target readings will also help to measure more proximal increases in specific reading comprehension. Future iterations of this study should also include the collection of participant information related to SIFE in recognition of the unique learning needs of this population of learners. The inclusion of such information may help to determine the efficaciousness of the reading intervention for use with this subgroup of ELLs.

**Limitations**

This study is not without limitations. Firstly, the reduced number of participants places significant restrictions on the ability to arrive at any conclusions with regards to the effects of the intervention. It is also necessary to recognize that “To serve effectively in the interpretation of test scores, norms must be appropriate for the group or individual to be evaluated” (Aiken, 2003, p. 74). The normative sample for the TOSCRF-2 consisted of 2,375 students from 29 states (including the state in which this research study was conducted). Although the test publishers noted that the normative sample is representative of the United States in terms of gender, race, geographic region, Hispanic status, exceptionality status, household income, and parental educational attainment, it is plausible that the participants in this study deviated from the normative sample. Thus, further caution should be used when interpreting the results.

Additionally, the Covid-19 pandemic necessitated the online administration of assessments that are designed to be administered in paper and pencil format. It is necessary to recognize that this change in administration may have impacted student performance. Efforts were made to ensure that the administration of the assessments closely replicated the original administration guidelines for the individual assessments. For example, the TOSCRF-2 is designed to be administered using a paper and pencil format where the participant is provided with a Student Record Form consisting of a connected series of words upon which the participant draws dividing lines according to the instructions. In order to closely simulate the original
administration of the assessment, I held a trial assessment with an associate professor of education to determine the conditions that would be necessary to administer the assessment via Zoom using the screen share and remote control features. By conducting a trial session for the administration of the assessments, I attempted to determine the level of magnification necessary for the participant to clearly see the assessment and receive feedback from the practice participant (the associate professor of education) in terms of the experience of using the digital pencil through remote control. Although this simulation helped me to replicate the necessary testing conditions using a digital platform, research does call into question the actual equivalency of this testing approach. Hosseini et al. (2014) reported that in their study evaluating the difference in student performance on a multiple-choice assessment between equivalent pencil-based tests and computer-based tests, ELLs performed better on the pencil-based tests. Thus, although the pandemic necessitated this form of computer-based test delivery, a replication of the study in person might be warranted to confirm the results.

Finally, the delivery of an intervention via a digital platform can solicit questions related to treatment integrity and whether the interventionist can be certain that the participants viewed the intervention videos in their entirety. Contrary to an in-person intervention, the very nature of the digital delivery can render the participant experience and level of engagement with the intervention opaque. In an effort to measure treatment integrity, the Intervention Video Quizzes were designed to monitor student engagement with the interventions. Participants A and C achieved scores that indicated a significant level of engagement with the videos and comprehension of the content. Participant B obtained scores that indicated a comparatively lower level of engagement and comprehension of the content. Most notably, Participant B did not return the completed Intervention Video Quiz 3 to the interventionist despite several reminder emails. Although the quizzes cannot certify that participants viewed all interventions in their entirety, the participants’ overall performance is a positive indicator related to treatment integrity.
Conclusion

This study represents a proof-of-concept investigation into the changes in undergraduate ELLs’ knowledge of English morphology, decoding, and general reading comprehension in response to a reading intervention grounded in morphological, alphabetic, and phonological awareness. It builds upon previous research conducted by Cook and Hughes (in preparation) that supported the use of a reading intervention that included alphabetic and phonological awareness and includes the addition of a morphological component to address the role of vocabulary in reading comprehension. The findings were mixed and weighted in favor of the proximal variables. However, given the limited number of participants and the short intervention period, future iterations of this research involving both a group experimental design with randomization and an increase in the intensity of the intervention may be warranted. The targeted nature of this reading intervention may provide educators working with adolescent ELLs who are struggling with reading comprehension a contextually relevant intervention that can be integrated within their pedagogical practice.
REFERENCES


doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1016/j.lindif.2011.01.008


doi: https://doi.org/10.1353/foc.2011.0007


doi:10.1017/S0142716400007839


e-learning applications for career and technical education: Technologies for vocational training (pp. 263-275). IGI Global. doi: 10.4018/978-1-60566-739-3


doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1080/10888438.2011.593066


https://doi.org/10.1016/0022-4405(81)90029-7


doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1177/0022219412463435


https://doi.org/10.1177/001440291107700304


https://doi.org/10.1177/0022219418775117


Appendix A

Literature Review and Supporting Information

Reading Interventions for English Language Learners at Risk of Reading Failure

Michelle J. Cook

Pennsylvania State University
Abstract

Reading proficiency represents a pivotal skill for students that extends beyond the academic realm. Despite the ever-increasing number of students enrolling in the American educational system for whom English is not their first language, there exists a paucity of research relating to interventions for use with English language learners (ELLs) who are struggling to become proficient readers. I conducted this literature review (a) to examine what researchers have reported about the components of reading interventions delivered in English associated with positive outcomes for ELLs identified as at risk of reading failure in elementary and middle school, (b) to assess the strength of the research designs of included studies according to the quality indicators described by Gersten et al. (2005) and Horner et al. (2005), and (c) to determine effect sizes for studies where a common reading assessment was applied in order to facilitate a comparison of intervention component composition across grade levels. The initial search located twelve studies involving 769 participants ranging from kindergarten to the eighth grade. The articles reviewed provided evidence supporting the inclusion of the following components in a reading intervention package for at-risk ELLs: (a) phonological awareness, (b) alphabetical principle, (c) comprehension strategies, (d) fluency, and (e) vocabulary instruction. Additionally, instruction that is explicit, provided to small groups, and includes modifications for ELLs was determined to be advantageous. Six of the studies used a common standardized reading assessment and a homogeneous participant pool permitting comparison of effect sizes across grade level bands. Implications are discussed.

Keywords: reading interventions, English language learners, at-risk, special education
Reading Interventions for English Language Learners at Risk of Reading Failure

Reading ability is a pivotal skill for all students. Students who become competent readers have access to content knowledge (Carnine & Carnine, 2004) and are better positioned to succeed academically (Hammer et al., 2014). Conversely, students who struggle to become capable readers are not only at-risk academically, but face precarious futures beyond the classroom, extending into their ability to function in society at large (Alexander, 2005).

Research on Reading Interventions for English Language Learners

Given the significant consequences related to the successful acquisition of reading skills, educational research has focused on determining the best pedagogical practices to teach literacy. The majority of early literacy research has focused on teaching reading to native English speakers (August, Shanahan, & Escamilla, 2009; Gyovai, Cartledge, Kourea, Yurick, & Gibson, 2009). The findings of such studies were synthesized by the National Reading Panel (NRP, 2000) and the following five areas of knowledge were determined to be fundamental to reading success: (a) phonological awareness (PA), (b) alphabetical principle (AP), (c) fluency, (d) text comprehension strategies, and (e) vocabulary knowledge (Ehri et al., 2001). Additionally, extant literature indicates that explicit instruction represents a favored pedagogical approach when working with students who are struggling to read (Tam, Heward, & Heng, 2006).

Although much is known about teaching reading to native English speakers, there is a dearth of knowledge relating to the instructional components that contribute to positive reading outcomes for English Language Learners (ELLs) (Hammer et al., 2014; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzakanani, 2003). This insufficiency exists despite the fact that ELLs represent the fastest growing subgroup within the student population in North America (Wagner, Francis, & Morris, 2005) and that it is predicted that by 2030, up to 40% of the student population in the United States may be ELLs (U.S. Department of Education, 2003). According to the U.S. Department of Education National Center for Education Statistics (2017), in 2014-2015 approximately 665,000 ELLs were identified as students with disabilities, amounting to 13.8% of the total ELL population enrolled in U.S. public elementary and
secondary schools. Given that 56% of ELLs considered for special education services seek remediation for reading problems (Gyovai et al., 2009), there is a need to extend the research to address the learning of ELLs who are at risk of reading failure.

**Learning to Read: Developmental Trajectory**

Prior to examining research relating to reading interventions for ELLs, it is important to establish whether the developmental trajectory for learning to read is the same for ELLs when compared with their native English peers. According to the research literature, early reading skills including word and text reading competence are equally attainable for both ELL and native English speaking populations and achievement levels are similar between the two groups in response to effective instruction (Lovett et al., 2008). As such, Lipka et al. (2005) asserted that the mere fact that ELLs are learning to read in their non-native language does not represent a risk factor for reading disability.

Some divergence in the findings does exist with regard to the development of early phonological processing skills when comparing ELLs to their native English speaking peers (Lovett et al., 2008). According to Lovett et al. (2008), some studies have shown that ELLs in the early elementary grades achieve comparable levels of phonological skills when compared to their native English speaking peers (Chiappe, Siegel, & Wade-Wolley, 2002; Lipka et al., 2005). Whereas Lovett et al. (2008) reported that other studies have demonstrated a delay in the emergence of phonological processing skills for ELLs (Lipka & Siegel, 2007). However, despite a possible lag in phonological and linguistic processing, a parallel developmental process in terms of the acquisition of basic literacy skills was found (Chiappe et al., 2002). Notably, Chiappe et al. (2002) asserted that for both ELLs and native English speakers, alphabetic knowledge and phonological processing remained salient early literacy skills and concluded that literacy acquisition is likely to follow a similar developmental trajectory for both groups of students. One exception for ELLs may be that alphabetic knowledge may emerge before phonological awareness and in turn facilitate the development of the latter (Chiappe et al., 2002). Generally, the research indicates that there is parity in terms of the development of early reading skills in ELLs and native English speakers (Lesaux & Siegel, 2003) and English proficiency has not been shown to be a crucial
factor in terms of developing early reading skills (Lesaux & Siegel, 2003). In terms of instructional approaches, Tam et al. (2006) noted that there is evidence to suggest that pedagogical practices that have been shown to be efficacious in terms of teaching reading to native English speakers are also effective for ELLs. However, these authors caution that some modifications might be necessary to address the unique learning needs of ELLs (Tam et al., 2006).

**Reading Interventions: Language of Instruction**

It is also important to address the issue of language of instruction when delivering reading interventions. Although some schools whose ELL populations are homogeneous (e.g., Spanish speaking) have employed bilingual instructional models, schools where the students identify as speaking a vast array of native languages (L1s), have primarily provided instruction in English (Calderon, Slavin, & Sanchez, 2011). In fact, Kamps et al. (2007) reviewed the extant research and found evidence to support the use of English interventions. They asserted that instruction of reading skills delivered in English is associated with positive reading outcomes for ELLs, irrespective of their L1. For the purpose of this literature review, only English reading interventions for ELLs at risk of reading failure will be reviewed.

**Purpose and Research Question**

Given the need to provide effective reading interventions for ELLs, the identification of what constitutes a promising intervention for ELLs at risk of reading failure has been presented as a pressing issue in need of attention (Kamps et al., 2007). Previous researchers have conducted syntheses of the research (Baker, Richards-Tutor, Sparks, & Canges, 2018; Richards-Tutor, Baker, Gersten, Baker, & Smith, 2016). However, this present synthesis will take a divergent approach by focusing uniquely on the components of interventions that are effective for use with ELLs at risk of reading difficulties with the ultimate goal of extracting data from the research that will enable a comparison of mean effect sizes across studies in order to analyze the effects of the distribution of various components across grade levels. As such, unlike previous literature reviews, only reading interventions delivered in English will be
examined and both group experimental and single case research designs will be considered for inclusion. Therefore, in this literature review, I aimed to address the following research questions:

RQ1. What have researchers reported about the components of reading interventions delivered in English for ELLs identified as at risk of reading failure in elementary and middle school?

RQ2. Do the included studies represent quality research designs according to quality indicators described by Horner et al. (2005) and Gersten et al. (2005) for single-subject research and group experimental designs respectively?

RQ3. What are effects across bands of grade levels? How can patterns of distribution across grade levels relating to reading intervention components be interpreted according to aggregated effect sizes?

Method

Selection Criteria

To determine what authors have reported about the components of reading interventions delivered in English associated with positive outcomes for ELLs identified as at risk of reading failure in elementary and middle school: (a) only experimental or quasi-experimental studies, published in refereed journals, were reviewed; (b) included studies examined English reading interventions as the independent variable; (c) reading performance of at-risk ELL students was the dependent variable; and (d) study participants were in either elementary or middle school.

Search Procedures

This literature review was conducted in two phases. The first phase involved a primary literature search in order to locate all applicable studies to answer R1. The second phase involved a secondary literature search conducted within the located studies in order to answer R3.

Primary literature search. I identified studies using a multistep process. First, I searched the electronic databases Education Resources Information Center (ERIC) and PsychINFO using descriptors in various forms. Descriptors pertaining to the target participants and settings were: ELL, ESL, EFL, L2,
FL, English language learner, English as a second language, English as a foreign language, linguistically diverse, special education, at-risk, elementary school, and middle school. Descriptors pertaining to the independent and dependent variables were: multicomponent, intervention, phonological awareness, difficulties, and reading. Second, an ancestral search of the articles retrieved from the databases produced additional articles. A cross-check using Google Scholar was conducted to ensure that all relevant articles had been considered for inclusion. The procedures yielded 198 articles; however, examination resulted in a final count of 12 articles for this literature review (See Figure 1). Most studies were excluded because the independent variable was not a reading intervention, the intervention was delivered in a language other than English, the ELL participants were not identified as at risk of reading difficulties, or the study design lacked a control group.

**Evaluation of Design Quality of Included Studies**

Procedures similar to those employed by Sreckovic, Common, Knowles, and Lane (2014) were used to evaluate the quality of the included studies (see Tables 4 and 5). As such, a binary scale (met = 1; not met = 0) was applied to each component of the quality indicators (QIs). In order for an absolute indicator score to be one, all associated components had to be met otherwise the indicator was assigned an absolute score of 0. However, a weighted score for each indicator was calculated by dividing the number of components met by the total number of components associated with each indicator. As such, each associated component contributed equally to the total.

**Single-case research designs: Coding procedures for quality indicators.** Horner et al. (2005) developed a set of QIs that can be applied to single-subject research studies in order to evaluate the strength of the design. These seven QIs comprised of 21 components formed the basis of this assessment of the single-case research designs included in this review.

**Group experimental designs: Coding procedures for quality indicators.** Gersten et al. (2005) developed a set of QIs that can be applied to group design studies in order to evaluate the strength of the design. The QIs are divided into two overarching categories: essential quality indicators (EQIs) and
desirable quality indicators (DQIs). The former includes 4 EQIs comprised of 21 components and the latter consists of 8 DQIs.

**Secondary literature review.** Following the primary literature search and review, the 12 studies were coded in terms of the standardized assessments applied at pre and posttest to determine reading outcomes. This process yielded seven studies which were then coded to identify homogeneous participant pools according to grade level. This procedure yielded six articles that utilized both comparable subtests of a standardized reading assessment and homogeneous participant pools in terms of grade level.

**Effect size for group studies**

In order to calculate effect size for the subtests utilized by the studies that administered the same standardized reading assessment, I replicated the procedures employed by Hwan and Riccomini (2016). Hedges’ $g$ was calculated for these studies to estimate the effect size related to specific measures of reading skill: word-letter identification, word attack, and passage comprehension. Given that these studies included a control group, effect size was calculated as $E_{\text{IGRM}} = \frac{(X_{\text{post.ctr}} - X_{\text{pre.ctr}})}{\sqrt{\frac{(n_{\text{trt}}-1)s_{\text{trt}}^2 + (n_{\text{ctr}}-1)s_{\text{ctr}}^2}{n_{\text{trt}} + n_{\text{ctr}} - 2}}}$. For the respective treatment and control groups, $X_{\text{post.trt}}$ and $X_{\text{post.ctr}}$ were unadjusted post-test means, $X_{\text{pre.trt}}$ and $X_{\text{pre.ctr}}$ were unadjusted pre-test means, $n_{\text{trt}}$ and $n_{\text{ctr}}$ were sample sizes, and $s_{\text{trt}}$ and $s_{\text{ctr}}$ were unadjusted posttest standard deviation. Following the recommendations of Hedges and Olkin (1985), an unbiased estimator was obtained by multiplying all of the effect sizes by $(1 - \frac{3}{4(n_{\text{trt}}+ n_{\text{ctr}})-9})$. As per the recommendations of Turner and Bernard (2006), when values were missing (i.e. means, standard deviations, sample size) study authors were contacted. When values could not be obtained, the effect size for the group comparison was set equal to zero (Turner & Bernard, 2006). Standard errors and confidence intervals were calculated and reported with the corresponding effect sizes.

Prior to calculating an average effect size across studies within a grade level band, an average effect size within studies was calculated in order to avoid a violation of the assumption of statistical independence which can occur when the same comparison on multiple outcomes during analysis is
utilized (Turner & Bernard, 2006). I followed the procedures recommended by Turner and Bernard (2006) and found the sum of the effect sizes and divided by that number of effect sizes.

Finally, I averaged effect sizes across grade level bands. In order to resolve the confounding that can result from an unequal n (i.e. differences in sample size), a weighted mean for Hedge’s g was calculated to pool effect sizes across studies using the following formula: \[ Hedges' s \ g = \frac{\sum Hedges's \ g(i) \times IVWHedges's \ g(i)}{\sum IVWHedges's \ g(i)} \] In this formula, “i” was the number of study effect sizes and IVW represented the inverse variance weight for Hedges’ g.

**Results**

This review was designed to examine the research on components of reading interventions delivered in English that are associated with positive outcomes for ELLs identified as at risk of reading failure in elementary and middle school. A summary of the grade level, sample size, L1, and method used to determine at-risk status for each of the 12 studies is provided in Table 1. A summary of the instructional components contained in each study’s reading intervention and the delivery of interventions is provided in Table 2. Information regarding setting, experimental design, and results is provided in Table 3. Results for coding of single-case research design studies according to the quality indicators (QIs) described by Horner et al. (2005) are provided in Table 4. Results for coding of group design studies according to the QIs described by Gersten, et al. (2005) are provided in Table 5. Results of the secondary literature review and calculations of effect size are provided in Table 6.

**Primary Literature Search**

**Participants and Settings**

The studies reviewed included 769 participants ranging in grade level from kindergarten to grade 8. With the exception of two studies (Lovett et al., 2008; Tam et al., 2006) that included a combined number of 81 participants in middle school (10.5% of the total number of participants), 89.5% of the studies focused solely on the early elementary grades. The studies included students who spoke a variety of different languages as their L1 including Spanish, Somali, Sudanese, Vietnamese, Portuguese,
Tagalog, Italian, Polish, Arabic, Syrian, Urdu, Khmer, Chinese, and Amharic. Participants were designated as at risk of reading failure using a variety of different identification methods ranging from standardized tests, government assessment protocols, experimenter developed screening tools, specific diagnoses, teacher referral, and considerations of general risk factors.

Seven studies indicated that they occurred in multicultural urban communities (Gyovai et al., 2009; Kamps et al., 2007; Lovett et al., 2008; O’Connor, Bocian, Beebe-Frankenberger, & Linklater, 2010; Vadasy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b). Additionally, several study settings were identified as low-income, including four studies which occurred in Title 1 schools (Filippini, Gerber, & Leafstedt, 2012; Leafstedt, Richards, & Gerber, 2004; Solari & Gerber, 2008; Vadasy & Sanders, 2010). Other studies explicitly selected a setting where the school was identified as effective for ELLs in an effort to control for the possibility of student low achievement in reading resulting from a lack of exposure to adequate educational practices (Linan-Thompson et al., 2006; Vaughn et al., 2006a; Vaughn et al., 2006b). All of the studies provided the reading interventions as a supplemental component to the participants’ regular education. In terms of the control condition, the school’s traditional instruction was provided to the comparison group for the majority of the included studies.

**Experimental Design**

The 12 reviewed studies employed two different experimental designs to determine the effects of instructional interventions on the reading performance of ELLs at risk of reading failure. Two studies utilized a multiple baseline across participants design (Gyovai et al., 2009; Tam et al., 2006). This experimental single-subject design allows the participants in baseline to act as his or her own control because the onset of the intervention is staggered across time. Conversely, a group experimental study design (i.e., randomized control trial) was utilized for the remaining ten studies (Denton, Wexler, Vaughn, & Bryan, 2008; Filippini et al., 2012; Kamps et al., 2007; Leafstedt et al., 2004; Linana-Thompson et al., 2006; Lovett et al., 2008; O’Connor et al., 2010; Solari & Gerber, 2008; Vadasy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b).
**Independent Variable: Reading Intervention**

The independent variable for all of the included studies was a reading intervention designed to remediate reading deficits of ELLs at risk of reading failure. Despite the common objective of each study, the instructional components of the 12 intervention packages varied as follows: (a) 91.6% contained a PA component (Filippini et al., 2012; Gyovai et al., 2009; Kamps et al., 2007; Leafstedt et al., 2004, Linan-Thompson et al., 2006; Lovett et al., 2008; O’Connor et al., 2010; Solari & Gerber, 2008; Vadasy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b), (b) 83.3% contained an AP component (Filippini et al., 2012; Gyovai et al., 2009; Kamps et al., 2007; Linan-Thompson et al., 2006; Lovett et al., 2008; O’Connor et al., 2010; Solari & Gerber, 2008; Vadasy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b), (c) 33.3% had a fluency component (Kamps et al., 2007; Linan-Thompson et al., 2006; Tam et al., 2006; Vaughn et al., 2006a; Vaughn et al., 2006b), (d) 33.3% had a comprehension component (Kamps et al., 2007; Linan-Thompson et al., 2006; Vaughn et al., 2006a; Vaughn et al., 2006b), and (d) 41.6% included a vocabulary component (Filippini et al., 2012; Kamps et al., 2007; Tam et al., 2006; Vaughn et al., 2006a; Vaughn et al., 2006b). Additionally, some studies included supplemental components including (a) oracy instruction (Vaughn et al., 2006a; Vaughn et al., 2006b), (b) listening comprehension (Solari & Gerber, 2008; Vaughn et al., 2006a), and (c) error correction (Tam et al., 2006). A component designed to differentiate instruction for the target group of participants was the inclusion of ELL modifications in some studies such as the use of concrete examples and scaffolding to reduce the level of challenge presented by abstract tasks (O’Connor et al., 2010; Vaughn et al., 2006a; Vaughn et al., 2006b).

All included studies used explicit instruction when delivering their respective reading intervention. Additionally, 11 of the included studies administered their intervention to small groups of students ranging in size from individual instruction to a maximum number of eight students. The only study that did not employ small group sizes for the delivery of the intervention relied on a push-in model where the interventionist provided supplemental instruction to the entire class (Filippini et al., 2012). However, the authors did note that opportunities for differentiation of instruction were provided.
The intervention was administered by individuals with varying levels of qualifications and experience across the included studies such as (a) a researcher who was a certified special education teacher (Leafstedt et al., 2004), (b) a reading specialist (O’Connor et al., 2010), (c) special education teachers (Lovett et al., 2008), (d) general education teachers (Vaughn et al., 2006a; Vaughn et al., 2006b), (e) teaching assistants (O’Connor et al., 2010; Vadasy & Sanders, 2010), (f) school personnel (Kamps et al., 2007), (g) the first author of the study (Tam et al., 2006), and (h) trained undergraduate and graduate researchers (Filippini et al., 2012; O’Connor et al., 2010; Solari & Gerber, 2008). Two studies did not explicitly identify the credentials of the individuals who administered the intervention (Gyovai et al., 2009; Linan-Thompson et al., 2006).

The overall duration of the studies and dosage in terms of length of treatment sessions varied significantly ranging from 10 weeks of 15 minute sessions two times per week on the low end (Leafstedt et al., 2004) and 32 weeks of 50 minute sessions five times per week on the high end (Vaughn et al., 2006b). The lowest average time participants spent in intervention ranged between 300 and 480 minutes (Filippini et al., 2012; Leafstedt et al., 2004; Solari & Gerber, 2008). By contrast, participants spent between 1200 and 2160 minutes in intervention in two studies (Gyovai et al., 2009; Vadasy & Sanders, 2010). The greatest amount of time participants spent in intervention was between 5750 and 7000 minutes (Linan-Thompson et al., 2006; Lovett et al., 2008; Vaughn et al., 2006a; Vaughn et al., 2006b).

Due to the nature of the interventions (e.g. movement between intervention and regular class instruction in response-to-intervention), it was not possible to calculate accurate estimates of time spent in intervention for three studies (Kamps et al., 2007; O’Connor et al., 2010; Tam et al., 2006).

**Dependent Variable: Reading Performance**

The dependent variable for all of the included studies was reading performance. However, the included studies contained different combinations of instructional components in order to target some or all of the five essential skills identified by the NRP (2000). Subsequently, each study measured the effect of their intervention package on the reading performance of ELLs at risk of reading failure accordingly.
As such, the results in terms of the effect of the independent variable on the dependent variable of the included studies are grouped in terms of intervention component composition.

**Fluency and vocabulary development.** A single study (Tam et al., 2006) examined the effectiveness of an intervention that included only the components of fluency and vocabulary development. Measures of reading fluency and comprehension were used to assess participant reading performance in response to the independent variable in this study. The authors reported that for both interventions, participants demonstrated an increase in the number of words read correctly per minute. The authors noted that progress was more distinct during same passage to criterion interventions than new passage each session interventions. With regard to words read incorrectly per minute, all participants in both intervention conditions showed a decrease in the number of incorrect reading rates. Moreover, the authors reported that all participants demonstrated an increase in their comprehension scores on taught passages during the two intervention conditions and that improvements were retained during maintenance. The authors proposed that improvements in oral reading rates and reading comprehension demonstrated in both conditions by 100% of participants were the result of vocabulary instruction and error correction as part of the fluency training.

**Phonological awareness.** The intervention administered by Leafstedt et al. (2004) was centered on a single component where the researchers focused on PA to assess its effect on reading performance. This study determined that an emphasis on PA for kindergarten students improved both PA and word reading skills as students in the intervention condition made greater gains than those in the control condition (Leafstedt et al., 2004). The authors noted that pre-intervention skills appeared correlated with post intervention outcomes. For example, low achieving participants maintained their achievement status relative to their peers in the intervention group as did the pre-study high achieving participants (Leafstedt et al., 2004). The authors also found that the lowest achieving students in the intervention condition made fewer gains in late PA skills. However, they attributed this result to the successful differentiation of instruction in the intervention condition which required a greater amount of instructional time spent with
the lowest achieving group dedicated to the acquisition of early PA skills at the expense of instruction in late PA skills (Leafstedt et al., 2004).

**Phonological awareness and alphabetical principle.** Four studies examined the effects of PA and AP on the reading performance of at-risk ELLs. According to Gyovai et al. (2009), improvements over baseline in both phoneme segmentation fluency and nonsense word fluency (NWF) were found for all participants. The authors reported effect sizes (ES) of the intervention on NWF ranging from 0.9 to 1.9 with only one student falling below 1.0. Pretest/posttest scores on the Woodcock Johnson assessment tool showed that group 1 experienced larger gains in letter-word identification than groups 2 and 3. Word attack (WA) subtests showed a similar pattern of achievement for the three groups in terms of mean group gain of raw score points. The authors noted that this pattern of achievement was anticipated given that the multiple baseline across participants design resulted in group 1 participants remaining in the instruction condition for a greater period of time. The authors reported that on all measures, lower performing students showed greater gains than their higher performing counterparts. Similarly, Lovett et al., 2008, found significant fixed effects for the intervention condition on all measures except for one (WRMT-R Passage Comprehension). These authors noted that intervention participants showed greater rates of growth over time in their reading and reading related skills and scored higher on all measures than those in the control condition (Lovett et al., 2008). Specifically, on Test of Transfer, the participants in the intervention condition scored higher on posttest ($M = 69.5/100$) than control ($M = 55.2/100$) and intervention participants increased their accuracy by 12.1 words at each testing point compared to the control students who increased their accuracy by 5.0 accordingly. O’Connor et al. (2010) obtained similar results and found effects that were significant for PA. Additionally, O’Connor et al. (2010) reported that participants who received the intervention earlier in the kindergarten year benefited more than students who received a delayed intervention package starting in January of the same academic year. Vadasy and Sanders (2010) reported similar findings given that participants in the treatment groups outperformed those in the control groups on measures of AP, word reading, spelling, passage reading,
fluency, and comprehension. However, these authors did report that ELLs were less responsive to treatment on word reading when compared to their non-ELL peers in the treatment condition.

**Phonological awareness, alphabetical principle, and vocabulary instruction or listening comprehension.** Two studies added either a component of vocabulary instruction or listening comprehension to PA and AP to examine the effects on the dependent variable (Filipini et al., 2012; Solari & Gerber, 2008). In the latter study, measures of vocabulary knowledge and phonological decoding skills were used to assess participant reading performance in response to the intervention (Filipini et al., 2012). The researchers found that all participants in the experimental conditions demonstrated statistically reliable growth from pretest to posttest on measures of NWF. In terms of phonological decoding, the authors reported that students who received the lowest pretest scores on NWF demonstrated greater gains in their posttest scores in both treatment conditions when compared to their peers. Furthermore, the researchers found that students who received a combined PA and vocabulary intervention showed the most improvement on measures of vocabulary without any notable loss in PA when compared to the PA-only control group. Similarly, Solari and Gerber (2008) found that an intervention package consisting of 70% listening comprehension, 20% PA, and 10% alphabetical knowledge produced positive effects in listening comprehension without negatively impacting PA and alphabetical knowledge outcomes.

**Phonological awareness, alphabetical principle, fluency, and comprehension.** Another included study administered a reading intervention that contained four components: PA, AP, fluency, and comprehension (Linan-Thompson et al., 2006). In this case, measures of decoding and comprehension were used to determine participant reading performance in response to the independent variable. Researchers categorized students as either responders or non-responders to instruction. The authors reported that a greater number of participants in the treatment condition responded to instruction than did those in the control condition. For example, at the end of first grade, 59% of control students remained at-risk of reading failure. Additionally, four of the non-responders from the control group were available for assessment at the end of the second grade and of these four participants, three remained at-risk.
Conversely, 91% of the first-grade students in the intervention condition responded to instruction and were no longer considered at risk at the end of Grade 1. Additionally, of those students in the treatment group who remained available for assessment in second grade, only one student was found to have returned to an at-risk status (Linan-Thompson et al., 2006).

**Phonological awareness, alphabetical principle, fluency, comprehension, and vocabulary.**

Finally, three included studies used interventions that contained all five components of PA, AP, fluency, comprehension, and vocabulary (Kamps et al., 2007; Vaughn et al., 2006a; Vaughn et al., 2006b). In their study, Kamps et al., 2007 used measures of early literacy skills as revealed by DIBELS assessments for decoding using nonsense word and oral reading fluency to assess student reading performance in response to the independent variable. They reported that the ELLs in the experimental condition experienced the greatest improvements in early literacy skills, where participants in the secondary-tier of the intervention who received direct instruction in small groups demonstrated the greatest outcomes. In terms of NWF, the researchers found significant differences between the experimental and control groups (df, 1; $F = 10.800, p = .001$). They reported that NWF means calculated in the fall, winter, and spring for the ELL experimental group were 22.7, 47.8, and 60.5 whereas, for the comparison group, the ELL group means were 17.5, 29.2, and 35.2. ES using Cohen's $d$ formula for the ELL group was 0.70 (Kamps et al., 2007). In the study conducted by Vaughn et al. (2006a), results showed that intervention students benefited most in areas related to early reading skills such as PA, letter-sound identification, decoding, and word reading efficiency given that ESs were approximately 0.40. Participants also gained eight standard score points in terms of measures of reading comprehension. The degree of response to intervention varied where spelling showed significant effects and text fluency only showed small effects. Finally, Vaughn et al. (2006b), measured reading outcomes in terms of PA, WA, fluency, comprehension, and oral language. These researchers reported that students in the treatment condition achieved higher outcomes than their peers in the control condition on all measures. Moreover, they noted that the lowest achieving participants in all reading skills on pretest in the treatment group scored (a) above average on
WA and comprehension, (b) average on dictation, and (c) near average in terms of fluency and language composite following intervention (Vaughn et al., 2006b).

**Evaluation of Design Quality of Included Studies**

**Single-Case Research Designs: Quality Indicators**

**Quality indicator 1: Describing participants and setting.** One of the two single-case research design studies included in this review met all of the components of QI 1 as defined by Horner et al. (2005). Gyovai et al. (2009) adequately described the participants, the selection process, and the setting. Tam et al. (2006) indicated that two included participants had a specific learning disability and one participant had a developmental disability, however, they did not specify the disability or the instrument/process employed to diagnose these students’ disabilities. Tam et al. (2006) reported that the school district’s privacy regulations meant that they had to rely on teacher referrals for participant selection.

**Quality indicator 2: Dependent variable.** Both included studies met all of the components for QI 2 as identified by Horner et al. (2005). Although both studies involved a reading intervention, the dependent variables for Gyovai et al. (2009) were phoneme segmentation fluency and nonsense word fluency, whereas the dependent variables for Tam et al. (2006) were oral reading rate and comprehension.

**Quality indicator 3: Independent variable.** Both included studies met all of the components for QI 3 as identified by Horner et al. (2005). Gyovai et al. (2009) and Tam et al. (2006) provided detailed descriptions of the independent variable to allow for both accurate interpretations of the results and replication. Gyovai et al. (2009) assessed fidelity of implementation for 21.6% of treatment sessions using both a quantitative and a qualitative scale. They obtained a mean of 100% and 88% respectively. Tam et al. (2006) assessed fidelity of implementation for 20% of all intervention sessions and found that procedural reliability was 100%.

**Quality indicator 4: Baseline.** Both included studies met all of the components for QI 4 as described by Horner et al. (2005) including repeated measurement, an established pattern, and a description of baseline that provided enough detail to permit replication. Gyovai et al. (2009) described
baseline as the regular classroom reading program using Trophies (Beck, Farr, & Strickland, 2003). Tam et al. (2006) also provided a detailed description of the reading methodology without feedback that formed the basis of the baseline procedures implemented prior to the onset of the reading intervention.

**Quality indicator 5: Experimental control/internal validity.** Both included studies met all of the components for QI 5 as described by Horner et al. (2005) including a range of 10 to 12 demonstrations of experimental effect (Gyovai et al., 2009; Tam et al., 2006).

**Quality indicator 6: External validity.** Both included studies met QI 6 in accordance to the guidelines determined by Horner et al. (2005) as the experimental effects were replicated across 4 participants in Tam et al. (2006) and 11 participants in Gyovai et al. (2009).

**Quality indicator 7: Social validity.** Both included studies met all of the components for QI 7 as described by Horner et al. (2005). Most notably, the change in the dependent variable of improved reading performance was socially important as participants in both studies showed notable improvement on targeted measures of reading performance. Additionally, both reading interventions were practical and cost effective as they occurred within the school setting and involved low-tech resources such as paper-based reading materials. As such, the interventions described by Gyovai et al. (2009) and Tam et al. (2006) represented viable options in typical contexts.

**Group Experimental Designs: Quality Indicators**

**Essential quality indicator 1: Describing participants.** Seven of the 10 included group design studies met all of the components of EQI 1 as described by Gersten et al. (2005). Two studies met one of the three components as they provided sufficient information to enable a determination of comparable conditions which could inform replication (Fillipini et al., 2012; Kamps et al., 2007). However, it was not possible to ascertain whether the participants demonstrated a disability/or difficulty as Fillipini et al. (2012) identified participants as at risk of reading difficulties based on their enrollment in a Title 1 school and the percentage of students receiving free or reduced-price lunches. Similarly, Kamps et al. (2007) relied solely on tier 2 RTI assignment in order to determine the at-risk of reading failure status for
participants. Another study met two of the three components in terms of describing participant disability/or difficulty and comparable interventions across groups (Linan-Thompson et al., 2006).

**Essential quality indicator 2: Implementation of the intervention.** Seven of the 10 studies met all of the components for EQI 2, whereas three studies did not adequately describe and assess treatment fidelity (Leafstedt et al., 2004; Linan-Thompson et al., 2006; Lovett et al., 2008). In the case of one study, a second researcher observed and videotaped the delivery of the intervention, however, no specific treatment fidelity scoring was reported (Leafstedt et al., 2004). Treatment fidelity may have been precluded from the Linan-Thompson et al. (2006) study because it was part of a much larger multistate, multisite longitudinal study. Similarly, the Leafstedt et al. (2004) study may have excluded treatment fidelity for logistical reasons as the participant pool was substantial (166 participants) and spread over 16 schools in a large urban school district. This may also explain why both Leafstedt et al. (2004) and Linan-Thomson et al. (2006) did not clearly document practices in the comparison group. Finally, one study did not provide sufficient detail regarding the intervention such as teacher language, instructional procedures, and materials to enable replication which may again be attributed to the fact that participating schools represented a subsample of a larger study (Linan-Thompson et al., 2006).

**Essential quality indicator 3: Outcome measures.** All 10 studies met EQI 3 given that they employed multiple measures that were aligned with the interventions’ research questions and measures were completed at appropriate times. All 10 of the group design studies reviewed utilized a pretest/posttest assessment schedule to determine the effects of the independent variable.

**Essential quality indicator 4: Data analysis.** Nine of the 10 studies met all of the components for EQI 4. One study linked the research question and unit of analysis with data analysis techniques, but did not report effect sizes (Linan-Thompson et al., 2006). Instead, these researchers relied on the standard score performance of participants on the Woodcock Language Proficiency Battery-Revised (WLPB-R; Woodcock, 1991; Woodcock & Munoz-Sandoval, 1995) to categorize participants as responders and non-responders.
Desirable quality indicators. The DQIs of validity and clarity of results were met by all included group design studies. Attrition was below 30% in each condition for six of the 10 studies (Fillipini et al., 2012; Leafstedt et al., 2004; Linan-Thompson et al., 2006; Lovett et al., 2008; O’Connor et al., 2010; Solari & Gerber, 2008). The DQI of reliability of estimates was met in seven of the 10 studies (Fillipini et al., 2012; Linan-Thompson et al., 2006; O’Connor et al., 2010; Solari & Gerber, 2008; Vadassy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b). Only Kamps et al. (2007) measured outcomes beyond the immediate posttest as multiple measures were taken over a two-year period. Fidelity was assessed beyond surface features in seven of the 10 studies (Fillipini et al., 2012; Kamps et al., 2007; O’Connor et al., 2010; Solari & Gerber, 2008; Vadassy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b). The DQI of documentations of the comparison conditions was met by eight of the 10 studies (Fillipini et al., 2012; Kamps et al., 2007; Lovett et al., 2008; O’Connor et al., 2010; Solari & Gerber, 2008; Vadassy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b). Finally, audio and video exerts of the interventions were provided in four of the 10 studies (Leafstedt et al., 2004; Vadassy & Sanders, 2010; Vaughn et al., 2006a; Vaughn et al., 2006b).

Secondary Literature Review

Initially, the secondary literature review resulted in a subsample representing 58% of the original study pool, involving 63% of the initial participant sample. As such, seven studies (Leafstedt et al., 2004; Linan-Thompson et al., 2006; Lovett et al., 2008; Solari & Gerber, 2008; Vadasy & Sanders, 2010; Vaught et al., 2006a; Vaughn et al., 2006b) utilized comparable subtests of a standardized reading assessment including (a) Woodcock Reading Mastery Tests - Revised-Normative Update (WRMT-R/NU, 1998), (b) Woodcock Language Proficiency Battery-Revised (WLPB-R, 1991), and (c) Woodcock Johnson 111 Tests of Achievement 111 (2000). The Letter Word Identification and Word Attack subtests were employed to assess decoding and the Passage Comprehension subtest was used to evaluate reading comprehension. However, the study by Lovett et al. (2008) was eliminated from the secondary literature review during the second phase of coding because the participants spanned a relatively large grade level range (i.e. second to the eighth grade).
Efficacy of Interventions within and Across Studies Relative to Grade Level Bands

The effect of the interventions was measured and represented by Hedge’s g for the group design studies. The effect sizes for each study are presented in Table 6 and summarized herein. The three studies involving participants in kindergarten (Leafstedt et al., 2004; Solari & Gerber, 2008; Vadasy & Sanders, 2010) reported results relating to the decoding domain (i.e. Word Identification and Word Attack Subtests). In terms of the Word Identification and Word Attack subtests, the intervention implemented by Leafstedt et al. (2004) was largely effective at p = 0.05 with a magnitude per subtest of ES = 1.288 and ES = 1.106 respectively. In terms of the Word Identification subtest, the intervention implemented by Solari and Gerber (2008) was more variably effective at p = 0.05 with a magnitude of ES = 0.563 (PA concentration) and ES = -0.187 (LC concentration). Similarly, variable effectiveness at p = 0.05 was found in terms of the Word Attack subtest with a magnitude of ES = -0.069 (PA concentration) and ES = -0.427 (LC concentration). Vadasy and Sanders (2010) reported the mean of the Word Identification and Word Attack subtest standard scores and the intervention implemented had a medium effect at p = 0.05 with a magnitude of ES = 0.601.

In terms of a weighted mean effect, reading interventions targeting kindergarten students resulted in an improvement in reading skills in the sample by approximately 0.697 SD. However, when this improvement is generalized to the sample from which study participants were drawn, this improvement could be as high as 1.0001 SD or as low as 0.393 SD. Despite this variability in the estimate, the conclusion from this bare-bones meta-analysis is that one can be 95% confident that the effect of the reading interventions is positive.

The studies involving participants in grade one (Linan-Thompson et al., 2006; Vaughn et al., 2006a; Vaughn et al., 2006b) reported results relating to the decoding domain (i.e. Letter-Word Identification and Word Attack Subtests) and the reading domain (i.e. Passage Comprehension Subtest). Standard errors and confidence intervals were calculated and reported with the corresponding effect sizes. It should be noted that because the posttest values could not be obtained for the study by Linan-Thompson et al. (2006), the effect size for the group comparison was set equal to zero as recommended
by Turner and Bernard (2006). In terms of the Letter-Word Identification Subtest, the intervention implemented by Vaughn et al. (2006a) had a small effect at p = 0.05 with a magnitude of ES = 0.353. In terms of the Word Attack Subtest, the interventions implemented by Vaughn et al. (2006a) and Vaughn et al. (2006b) had a medium to large effect at p = 0.05 with a magnitude of ES = 0.467 and ES = 1.541 respectively. In terms of Passage Comprehension, the intervention implemented Vaughn et al. (2006a) had a small effect at p = 0.05 with a magnitude of ES = 0.13. With regards to Passage Comprehension, the intervention implemented by Vaughn et al. (2006b) had a large effect at p = 0.05 with a magnitude of ES = 1.325.

The weighted mean effect for the reading interventions reported in this secondary review targeting grade one students resulted in an improvement in reading skills in the sample by 0.433 SD. However, when this improvement is generalized to the sample from which study participants were drawn, this improvement could be as high as 0.751 SD or as low as 0.115 SD. Despite this variability in the estimate, the conclusion from this bare-bones meta-analysis is that one can be 95% confident that the effect of the reading interventions is positive.

**Discussion**

**Primary Literature Search**

In reviewing the literature to determine what researchers have reported about the components of reading interventions delivered in English associated with positive outcomes for ELLs identified as at risk of reading failure in elementary and middle school, the findings suggest that a comprehensive instructional package that reflects the essential reading skills identified by the NRP (2000) may be effective. According to the extant literature, components of a comprehensive instructional package for ELLs at risk of failure should include (a) PA, (b) AP, (c) comprehension strategies, (d) fluency, and (e) vocabulary instruction. This finding suggests that intervention packages that have been found to be successful with native English readers are also promising practices for use with ELLs at risk of reading failure. This also applies to a delivery of instruction that is explicit. However, modifications to address
the specific needs of ELLs (e.g., vocabulary instruction that includes concrete examples, visuals, and facial expressions) may be advisable in order to differentiate instruction for this population of students.

**Participants and Settings.** The results indicated that elementary and middle school students benefited from the reading interventions included in this review in terms of reading performance outcomes. Although both included studies involving middle school students (Lovett et al., 2008; Tam et al., 2006) reported positive effects, it should be noted that this is in contrast with another study where the results suggested that ELLs at risk of reading failure may become less responsive to intervention in the later grades (Denton et al., 2008). A discussion of these divergent findings is warranted given that Foorman, Breier, & Fletcher (2003) synthesized the research and found evidence to support the assertion that early intervention is most effective and that older students may become less responsive to intervention. Accordingly, Denton et al. (2008) offered an important glimpse at the resistance to treatment that may develop for ELLs who receive reading interventions in the later grades. Although the study was excluded from this literature review because data for ELL and non-ELL participants were only disaggregated for pre-test measures and not post-test measures, it is relevant to this discussion as it emphasizes the importance of providing timely interventions for ELLs at risk of reading failure. In their study, the authors obtained disappointing results in terms of response to intervention for struggling eighth-grade ELLs. In fact, despite employing an intervention that included all of the components that were shown to have positive outcomes for ELLs at risk of reading failure in the studies reviewed herein, outcomes for participants in the treatment condition for word recognition, comprehension, and fluency were not significantly greater than the those achieved by the control group. Moreover, they reported that over the course of the study, neither the treatment nor the control group experienced significant growth. The authors emphasized that on all reading measures, students in the treatment group did not show significantly higher results than their peers in the control condition. These findings suggest that ELLs with risk indicators who fail to receive effective remediation are likely to become more impervious to remediation (Gyovai et al., 2009).
In terms of whether students from different L1 backgrounds can be grouped together for intervention purposes, the experimental studies reviewed demonstrated that positive effects of the English reading interventions for ELLs have been found for students with varying L1 backgrounds. Specifically, the studies included in this review were not limited to a participant pool of one specific L1 group, but instead included students from many L1 populations and all students responded to the interventions. This finding is important given the implications for heterogeneous student populations.

Finally, the reading interventions described were shown to be effectual in varying settings including low-income, urban, multicultural, Title 1, and effective schools. These findings suggest that the interventions for ELLs at risk of reading failure included in this review can be successfully implemented in varying school environments. Considering the disparate nature of the American academic landscape, these results may be useful in informing teacher practice given that educators require efficacious interventions that can be delivered in a variety of settings.

**Independent Variable: Reading Intervention.** In terms of the independent variable, all of the included studies involved the manipulation of a reading intervention in order to measure its effect on reading performance. Only one reading intervention did not include PA. This is not surprising given that a student’s aptitude in terms of the manipulation of speech sounds is extremely indicative of early reading performance (Ehri et al., 2001). Additionally, the results show that ELLs may benefit from the inclusion of supplementary components that are responsive to their language needs in order to remediate reading deficits when delivering the independent variable. Most notably, modifications suitable for ELLs such as the use of concrete examples, gestures, facial expressions, and scaffolding to reduce abstract task requirements may be advisable. Moreover, small group size may also have functioned to enhance the effects of the independent variable. Given that principles of behavior stipulate that learning is enhanced when individuals are provided with increased opportunities to respond paired with positive reinforcement to increase desired behavior (Cooper, Heron, & Heward, 2007), small group size may indeed function to support the effective delivery of the independent variable.
**Dependent Variable: Reading Performance.** In all of the experimental studies, the dependent variable was positively impacted by the intervention and had favorable results for students in terms of reading performance. Based on the above findings, educators should consider reading intervention packages that contain the components included in this review to be efficacious in terms of improving the reading performance of ELLs at risk of reading failure. Although each included study measured different elements of student reading performance in order to determine the impact of the intervention on reading achievement, each aspect of reading performance measured constituted a fundamental element of reading competency.

**Secondary Literature Search**

The secondary literature search involved coding the 12 studies located during the primary literature search according to (a) shared standardized assessments applied at pre and posttest to determine reading outcomes and (b) homogeneity in terms of grade level of participants. This procedure yielded six articles where the results for the Woodcock Johnson subtests could be analyzed statistically to determine intervention effects across bands of grade levels. Additionally, the secondary literature search allowed for patterns of distribution across grade levels relating to reading intervention components to be interpreted according to aggregated effect sizes.

**Woodcock Johnson assessment data.** The decision to isolate data (means and standard deviations) related to versions of a particular standardized reading assessment was twofold. Firstly, it represented the most commonly used assessment amongst the included studies identified during the primary literature search and, secondly, the assessment in question is highly regarded in terms of efficacy in evaluating student reading skills.

Torgesen (1998) described diagnostic measures used to identify and assess student achievement to prevent reading failure. Notably, Torgesen (1998) posited that adequate reading comprehension represents the foremost important outcome related to reading instruction. He stipulated that two factors are essential to reading comprehension:

- General language comprehension ability.
• The ability to accurately and fluently identify words in print.

As such, Torgesen (1998) affirmed that measures that assess a student’s ability to decode words are important with respect to identifying a student’s reading skill level. In his discussion of reading assessment tools, Torgesen (1998) identified the Woodcock assessment measures as suitable for assessing decoding skill development in students. Specifically, Torgesen (1998) noted that the Letter Identification subtest of the Woodcock Reading Mastery Test-Revised (Woodcock, 1987) is an effective measure that presents letters in different fonts and allows students to provide either the letter name or the sound the letter makes in words. Torgesen (1998) also identified the Word Identification subtest from the Woodcock Reading Mastery Test-Revised (Woodcock, 1987) as a measure that is widely used and appropriate for determining student sight word reading ability. He noted that this assessment’s strengths include the fact that it has been normed nationally and it does not place strict time pressure on students, thus allowing for both the use of phonetic decoding and sight word processes. Additionally, given that a list format is used to present words, context cannot be used to support reading. Torgesen (1998) also identified the Word Attack Subtest of the Woodcock Reading Mastery Test-Revised (1987) as a sound example of a diagnostic assessment of phonetic reading ability. According to Torgesen (1998), research by Share and Stanovich (1995) indicated that measures of nonword reading are the most efficacious way of determining students’ knowledge of letter-sound correspondences in decoding words.

In light of the above discussion, the decision to base an evaluation of treatment effectiveness on aggregated mean effect sizes of data resulting from Woodcock reading assessment subtests is grounded in the assertion that they represent a reliable evaluation of foundational reading skills.

**Weighted Mean Effect Sizes: Analysis**

The secondary literature search and analysis revealed that the reading interventions targeting kindergarten students resulted in an improvement in reading skills in the sample by approximately 0.697 SD. The components included in the reading interventions for the kindergarten participants were alphabetical principle and phonological awareness. This reflects developmental and curricular outcomes for this grade level band. As such, students were only evaluated in terms of Letter Identification and
Word Attack skills. This may explain why the effect size related to these interventions was higher than the reported effect size for the grade one level band, given that language proficiency is not a determining factor in terms of the acquisition of early reading skills (Lovett et al., 2008).

With regards to the grade one level band, the secondary literature search and analysis determined that reading interventions targeting grade one students resulted in an improvement in reading skills in the sample by 0.433 SD. All five components identified in the primary literature search (PA, AP, comprehension strategies, fluency, and vocabulary instruction) were included in the reading interventions for the grade one participants. This reflects the more complex developmental and curricular expectations for this grade level band. As such, these participants were also given the Passage Comprehension subtest to assess student reading skill development. The nature of the higher order tasks involved in reading comprehension may explain why this grade level band had a lower effect size in terms of intervention effect compared to the kindergarten grade level band. Additionally, as noted by Torgesen (1998), reading comprehension involves the interplay between two key subsets of skills: general language comprehension ability and the ability to accurately and fluently identify words in print. The latter skill is related to AP and PA and is less impacted by a participant’s mastery of English vocabulary (Lovett et al., 2008). However, the former would be highly impacted by ELLs’ grasp of vocabulary in English.

The findings of this review add to the emerging literature base that seeks to determine what constitutes best practice in terms of remediating the reading deficits of ELLs at risk of reading failure. These particular findings are meaningful as they provide educators with classroom-ready strategies in order to address the needs of ELLs who are struggling to become capable readers in English. Thus, these results add to the existing knowledge base by demonstrating that the components of reading intervention packages that have been found to be effective with native English language students, can be seen as efficacious for use with ELLs.

**Limitations**

This review revealed a lack of uniformity in terms of how ELLs are identified for inclusion in the research. Studies varied in terms of how they identified ELLs including criteria such as: (a) parent report,
(b) teacher report, or (c) performance on a standardized test. This variation may limit the ability of researchers to draw comparisons across studies and a universal definition for ELLs within the research literature might help alleviate this issue.

Similarly, the task of identifying ELLs who are at-risk of reading failure due to a learning problem rather than simple exposure to the language remains difficult. In fact, the issue of correctly identifying learning disabilities and reading disabilities among ELLs remains a highly-debated topic (Gerber & Durgunogly, 2004). However, this shortcoming in the research literature does not negate the need for effective interventions for ELLs who are showing signs of being at-risk of reading failure. Given the importance of early intervention, these students should receive effective remediation as we await more accurate identification procedures.

The secondary literature search was necessary due to statistical concerns and limitations related to aggregating results obtained using different measures. Given that Hedge’s g is a measure for describing the magnitude of an intervention’s effect (Turner & Bernard, 2006), if either the participants or procedures used to measure study outcomes differ significantly, then aggregating their respective effect sizes is problematic. Calculating weighted mean effect sizes across grade level bands using data resulting from versions of a single standardized reading assessment was an attempt to control for extraneous factors introduced by study procedures (Hunter & Schmidt, 2004). Specifically, the measurement process can cause variation on the dependent variable (Hunter & Schmidt, 2004). By controlling for the actual measure used to determine the effect of the independent variable on the dependent variable as well as grouping participants according to grade level, I endeavored to reduce statistical concerns relating to calculating aggregated effect sizes across studies. Thus, a dearth of common standardized assessments used across studies as well as the heterogeneity of participant pools limited the extent that aggregated effect sizes could be used to make comparisons across studies according to either grade level or the component make-up of the studies. The more uniform application of similar standardized outcome measures and the inclusion of studies with more homogenous participant pools would increase the types of statistical analyses that could be conducted.
Future Research and Implications for Practice

The search conducted for this review confirmed the assertion by Vaughn et al. (2006) that there is an overall paucity in the extant literature of experimental studies relating to ELLs who are at-risk academically. Studies specifically examining reading interventions for this growing student population are required in order to adequately address the learning needs of ELLs and subsequently prepare future educators who are likely to work with these students in a 21st century classroom. Additionally, although this review allowed for an examination of the components of reading interventions for at-risk ELLs, it was not possible to determine the individual contribution value of each component. Studies designed to determine the relative value of each component will better inform pedagogical practice and help researchers to determine the amount of instructional time that should be allotted to each component.

Despite the need to conduct more research related to reading interventions for at-risk ELLs, educators should look to the findings of this literature to inform their practice. Pedagogical approaches to literacy instruction for ELLs at-risk of reading failure should include PA, AP, comprehension strategies, fluency, and vocabulary instruction.

Conclusion

The primary review of 12 experimental studies found that the research supports the inclusion of the following components in a reading intervention package for at-risk ELLs: (a) PA, (b) AP, (c) comprehension strategies, (d) fluency, and (e) vocabulary instruction. Additionally, delivery of instruction that is explicit and delivered to small groups of students is likely to be advantageous. Modifications to address the specific needs of ELLs are advisable in that they allow educators to scaffold learning. The secondary review of six experimental studies confirmed that the reading interventions described above can result in positive outcomes for students in kindergarten and grade one. The findings of this review have implications for classroom teachers seeking effectual strategies aimed at providing efficacious reading interventions for at-risk ELLs.
References


Issues in Communication Science and Disorders, 33, 42-55.


Table 1

Participant Characteristics

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Sample Size</th>
<th>Grade Level</th>
<th>ELL primary language (L1)</th>
<th>Determination of at-risk status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fillipini et al. (2012)</td>
<td>60</td>
<td>1</td>
<td>Spanish (n=59); Undetermined (n=1)</td>
<td>Not described</td>
</tr>
<tr>
<td>Gyovai et al. (2009)</td>
<td>12</td>
<td>K-1</td>
<td>Somali (n=10); Spanish (n=1); Vietnamese (n=1)</td>
<td>DIBELS (low performance on the fall benchmark standardized subtests); WJTA (below grade-level performance)</td>
</tr>
<tr>
<td>Kamps et al. (2007)</td>
<td>170</td>
<td>1-2</td>
<td>Spanish (n = 99); Somali/Sudanese/Vietnamese (n = 71)</td>
<td>DIBELS (at risk for reading failure based on screening in the fall of first and second grade)</td>
</tr>
<tr>
<td>Leaflstedt et al. (2004)</td>
<td>62</td>
<td>k</td>
<td>Spanish</td>
<td>General risk factors: low income, limited English, low parent education level, and limited literacy resources in the home</td>
</tr>
<tr>
<td>Linan-Thompson et al. (2006)</td>
<td>39</td>
<td>1</td>
<td>Spanish</td>
<td>WLPB-R: LWID in English and Spanish (&lt; 25th percentile); Experimenter word reading list (raw score &lt; 2)</td>
</tr>
<tr>
<td>Lovett et al. (2008)</td>
<td>76</td>
<td>2-8</td>
<td>Portuguese (n = 37); Spanish (n = 16); Tagalog (n = 6); Italian (n = 5); Polish (n = 3); Arabic (n = 2); Syrian (n = 1); Urdu (n = 1); Undetermined (n = 7)</td>
<td>Teacher referral WRMT: LWID and WA (SS &lt; 85) WRAT-3 (SS &lt; 85)</td>
</tr>
</tbody>
</table>

Note: DIBELS = Dynamic Indicators of Basic Early Literacy Skills; WJTA = Woodcock-Johnson Tests of Achievement-Third Revision; WLPB-R = Woodcock Language Proficiency Battery Revised; LWID = letter word identification; WRMT = Woodcock Reading Master Test; WA = word attack; WRAT-3 = Wide Range Achievement Test, Third Edition; LNF = letter naming fluency; ISF = initial sound fluency; PPVT = PeabodyPicture Vocabulary Test; PA = phonological awareness; LD = learning disability; ADD = attention deficit disorder; ESL = English as a second language; LN = letter naming; LS = letter sounds; CTOPP = Comprehensive Test of Phonological Processing; WLPB = Woodcock Language Proficiency Battery; RDWRL = researcher developed word reading list.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Sample Size</th>
<th>Grade level</th>
<th>ELL primary language (L1)</th>
<th>Determination of at-risk status</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Connor et al. (2010)</td>
<td>35</td>
<td>K</td>
<td>Spanish</td>
<td>DIBELS (LNF &lt; 9 and ISF &lt; 7); PPVT (SS &lt; 85)</td>
</tr>
<tr>
<td>Solari &amp; Gerber (2008)</td>
<td>82</td>
<td>K</td>
<td>Spanish</td>
<td>PA; Vocabulary (specific criteria not reported)</td>
</tr>
<tr>
<td>Tam et al, (2006)</td>
<td>5</td>
<td>3-5</td>
<td>Khmer (n = 2); Spanish (n = 2); Amharic (n =1)</td>
<td>Specific LD (n = 1); Specific LD and ADD (n= 1); Developmental disabilities (n = 1); No specific diagnosis (n = 2); School district office of ESL assessment guidelines</td>
</tr>
<tr>
<td>Vadasy &amp; Sanders (2010)</td>
<td>84</td>
<td>K</td>
<td>28 languages: Spanish (49%); Vietnamese (15%); Somali (6%); Chinese (6%); Tagalog (6%)</td>
<td>In bottom half of class on LN, LS; CTOPP Sound Matching (composite-score z score)</td>
</tr>
<tr>
<td>Vaughn et al. (2006a)</td>
<td>96</td>
<td>1</td>
<td>Spanish</td>
<td>WLPB LWD (&lt; 25th percentile); RDWRL (raw score &lt; 2)</td>
</tr>
<tr>
<td>Vaughn et al. (2006b)</td>
<td>48</td>
<td>1</td>
<td>Spanish (n = 48)</td>
<td>WLPB LWD (&lt; 25th percentile); RDWRL (raw score &lt; 2)</td>
</tr>
</tbody>
</table>

Note: DIBELS = Dynamic Indicators of Basic Early Literacy Skills; WJTA = Woodcock-Johnson Tests of Achievement-Third Revision; WLPB-R = Woodcock Language Proficiency Battery Revised; LVID = letter word identification; WRMT = Woodcock Reading Master Test; WA = word attack; WRAT-3 = Wide Range Achievement Test, Third Edition; LNF = letter naming fluency; ISF = initial sound fluency; PPVT = Peabody Picture Vocabulary Test; PA = phonological awareness; LD = learning disability; ADD = attention deficit disorder; ESL = English as a second language; LN = letter naming; LS = letter sounds; CTOPP = Comprehensive Test of Phonological Processing; WLPB = Woodcock Language Proficiency Battery; RDWRL = researcher developed word reading list.
Table 2

Breakdown of Intervention Packages

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Instructional Components</th>
<th>Delivery of Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA</td>
<td>AP</td>
</tr>
<tr>
<td>Fillipini et al. (2012)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Gyovai et al. (2009)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Kamps et al. (2007)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Leafstedt et al. (2004)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Linan-Thompson et al. (2006)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Lovett et al. (2008)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>O’Connor et al. (2010)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Solari &amp; Gerber (2008)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Tam et al. (2006)</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Vadasy &amp; Sanders (2010)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Vaughn et al. (2006a)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Vaughn et al. (2006b)</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Note. PA = phonological awareness; AP = alphabetical principle; Fl = fluency; Comp = comprehension; Vocab = vocabulary instruction; Misc = miscellaneous; LC = listening comprehension.
### Study Features and Results

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Setting</th>
<th>Design</th>
<th>Results</th>
</tr>
</thead>
</table>
| Fillipini et al. (2012) | Elementary school | GED | -100% of participants demonstrated statistically reliable growth from pretest to posttest on NWF.  
-Students in the PA + Vocabulary groups demonstrated greater average growth on the vocabulary measure compared with the PA only group.  
-Participants in the PA + Vocabulary groups achieved greater scores on vocabulary measures without experiencing any reduction in PA skills when compared to the control PA only group. |
| Gyovai et al. (2009) | Elementary school | MBAPD | -Improvements over baseline in both phoneme segmentation fluency and NWF were found for all participants.  
-Effect sizes of the intervention on NWF were from 0.9 to 1.9 with only one student falling below 1.0.  
-CBM showed that all participants improved with mean percentage gains of 1.9, 2.2, and 0.5 correct responses for groups 1, 2, 3 respectively.  
-Pretest/posttest scores on Woodcock Johnson showed that group 1 experienced larger gains on the LWID than groups 2 and 3 with group mean raw score gains of 4.7, 3 respectively.  
-WA subtests showed a similar achievement pattern with a mean group gain of raw score points of 12.3, 2.2, and 0.2 for groups 1, 2, 3 respectively.  
-Lower performing students showed greater gains than their higher performing counterparts. |
| Kamps et al. (2007) | 6 elementary schools | GED | -ELLS in the experimental condition experienced the greatest improvements in early literacy skills.  
-Significant differences were found in NWF between the experimental and control groups (df, 1; F = 10.800, p = .001).  
-NWF means calculated in the fall, winter, and spring for the ELL experimental group were 22.7, 47.8, and 60.5 versus 17.5, 29.2, and 35.2 for the comparison group.  
-Effect sizes using Cohen's d formula for the ELL group were 0.70. |
| Leafstelt et al. (2004) | Elementary school | GED | -Students in the intervention condition made greater gains than those in the control condition in both PA and word reading skills.  
-Pre-intervention skills appeared correlated with post intervention outcomes.  
-The lowest achieving students in the intervention condition made fewer gains in late PA skills. |

*Note. GED = group experimental design; PA = phonological awareness; AP = alphabetical principle; Comp. = comprehension; Vocab. = vocabulary; NWF = nonsense word fluency; MBAPD = multiple baseline across participants design; CBM = curriculum-based measurement; LWID = letter-word identification; WA = word attack; ELL = English language learner; WRMT = Woodcock Reading Master Test; PPVT-III = Peabody Picture Vocabulary Test, 3rd Edition; WPSSI = Wechsler Primary Scale of Intelligence; CTOPP = Comprehensive Test of Phonological Processing; WLPB-R = Woodcock Language Proficiency Battery Revised*
### Study Features and Results

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Setting</th>
<th>Design</th>
<th>Results</th>
</tr>
</thead>
</table>
| Linan-Thompson et al. (2006)     | 4 elementary schools             | GED    | - More participants in the intervention condition responded to instruction than did those in the control condition.  
                                  |                                  |        | - At the end of first grade, 59% of control students remained at risk of reading failure.  
                                  |                                  |        | - Of the four who were available for assessment at the end of the second grade, three remained at-risk.  
                                  |                                  |        | - Conversely, 91% of the first-grade students in the intervention condition responded to instruction and were no longer considered at risk at the end of grade one.  
                                  |                                  |        | - Of those students who remained available for assessment in second grade, only one student was found to have returned to an at-risk status |
| Lovett et al. (2008)             | 16 elementary schools in a large urban school district in Canada | GED    | - Significant fixed effects were found for the intervention condition on all measures except for one (WRMT Passage Comprehension).  
                                  |                                  |        | - Participants showed greater rates of growth in their reading and reading related skills and scored higher on all measures than those in the control condition.  
                                  |                                  |        | - On Test of Transfer, the participants in the intervention condition scored higher on posttest (M- 69.5/100) than control (M = 55.2/100).  
                                  |                                  |        | - Intervention participants increased their accuracy by 12.1 words at each testing point compared to the control students who increased their accuracy by 5.0. |
| O’Connor et al. (2010)           | 8 school                         | GED    | - Significant effects for PA.  
                                  |                                  |        | - Pretest = Rapid Letter Naming mean = 1.95; Initial Sound Fluency mean = 2.38; PPVT-111 mean = 68.14; WPSSI Picture Naming mean = 4.57  
                                  |                                  |        | - Posttest = Rapid Letter Naming mean = 32.05; Phoneme Segmentation Fluency mean = 33.71; Nonword Fluency mean = 16.48 |
| Solari & Gerber (2008)           | One school                       | GED    | - Students who received the most instruction in LC performed best on measures of LC.  
                                  |                                  |        | - Students in the LC Concentration intervention groups performed as well as students in the PA Concentration and PA only control groups on measures of PA |

*Note. GED = group experimental design; PA = phonological awareness; AP = alphabetical principle; Comp. = comprehension; Vocab. = vocabulary; NWF = nonsense word fluency; MBAPD = multiple baseline across participants design; CBM = curriculum-based measurement; LWID = letter-word identification; WA = word attack; ELL = English language learner; WRMT = Woodcock Reading Master Test; PPVT-III = Peabody Picture Vocabulary Test, 3rd Edition; WPSSI = Wechsler Primary Scale of Intelligence; CTOPP = Comprehensive Test of Phonological Processing; WLPB-R = Woodcock Language Proficiency Battery Revised*
### Table 3 Continued

#### Study Features and Results

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Setting</th>
<th>Design</th>
<th>Results</th>
</tr>
</thead>
</table>
| Tam et al. (2006) | Elementary school     | MBA PD | - All participants increased their correct reading rate by 11.1, 0.9, 23.4, 14.5, and 19.8 words per minute in new passage sessions and 35.7, 0.3, 44.7, 39.5, and 34.2 correct words in same passage to criterion sessions.  
- Progress was most distinct during same passage to criterion interventions.  
- All participants had a decrease in number of incorrect reading rates with an average decrease across learners of 11.7 fewer incorrect words in the new passage sessions and 13.5 fewer incorrect words during the same passage to criterion sessions.  
- The average number of correct answers given by learners to literal comprehension questions was 1.0 correct answers during baseline, 4.1 correct answers during new passage each session, 4.8 correct answers during same passage to criterion, and 4.3 correct answers during maintenance.  
- All participants had an increase in their comprehension scores on taught passages in the two intervention conditions.  
- 100% of participants improved their oral reading rates and reading comprehension. |
| Vadaszy & Sanders (2010) | 10 schools | GED | - Treatment groups outperformed control groups on all measures of alphabeticics, word reading, spelling, passage reading, fluency, and comprehension.  
- ELLs were less responsive to treatment on word reading than non-ELLs in treatment group. |
| Vaughn et al. (2006a) | 4 schools | GED | - Intervention students had slightly higher gains in English letter naming and English letter naming fluency than control.  
- In terms of English phonological processing, intervention students correctly answered an average of 52% of items across the subtests administered, relative to 47% for comparison students.  
- Effect size for the between-groups difference was moderate.  
- In terms of phonological memory, no difference between groups.  
- Intervention and comparison groups did not differ on the WLPB-R English oral language composite.  
- Intervention students performed better than control students on the WA subtest.  
- Intervention students outperformed comparison students on (a) a measure of isolated word reading fluency and (b) an experimental measure of spelling. |
| Vaughn et al. (2006b) | 4 schools | GED | - Students in the treatment condition achieved higher outcomes than their peers in the control condition on all measures.  
- The lowest achieving participants in all reading skills on pretest in the treatment group scored (a) above average on WA and comprehension, (b) average on dictation, and (c) near average on fluency and language composite following intervention. |

Note: GED = group experimental design; PA = phonological awareness; AP = alphabetical principle; Comp. = comprehension; Vocab. = vocabulary; NWF = nonsense word fluency; MBAPD = multiple baseline across participants design; CBM = curriculum-based measurement; LVID = letter-word identification; WA = word attack; ELL = English language learner; WRMT = Woodcock Reading Master Test; PPVT-III = Peabody Picture Vocabulary Test, 3rd Edition; WPSSI = Wechsler Primary Scale of Intelligence; CTOPP = Comprehensive Test of Phonological Processing; WLPB-R = Woodcock Language Proficiency Battery Revised
Table 4

Coding of Quality Indicators of Included Single Case Design Studies

<table>
<thead>
<tr>
<th>Quality Indicators</th>
<th>Gyovai et al. (2009)</th>
<th>Tam et al. (2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Describing participants and setting</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant description</td>
<td>Yes (1.00)</td>
<td>No (0.33)</td>
</tr>
<tr>
<td>Participant selection</td>
<td>Yes (0.33)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Setting description</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description</td>
<td>Yes (0.20)</td>
<td>Yes (0.20)</td>
</tr>
<tr>
<td>Quantifiable measurement</td>
<td>Yes (0.20)</td>
<td>Yes (0.20)</td>
</tr>
<tr>
<td>Valid and well-described measurement</td>
<td>Yes (0.20)</td>
<td>Yes (0.20)</td>
</tr>
<tr>
<td>Measured repeatedly</td>
<td>Yes (0.20)</td>
<td>Yes (0.20)</td>
</tr>
<tr>
<td>Interobserver agreement</td>
<td>Yes (0.20)</td>
<td>Yes (0.20)</td>
</tr>
<tr>
<td><strong>Independent variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent variable description</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Systematically manipulated</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Fidelity of implementation</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeated measurement, established pattern</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Description</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td><strong>Experimental control/ internal validity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three demonstrations of experimental effect</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Internal validity</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Pattern of results</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td><strong>External validity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social validity</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>DV is socially important</td>
<td>Yes (0.25)</td>
<td>Yes (0.25)</td>
</tr>
<tr>
<td>Change in DV is socially important</td>
<td>Yes (0.25)</td>
<td>Yes (0.25)</td>
</tr>
<tr>
<td>IV is practical and cost effective</td>
<td>Yes (0.25)</td>
<td>Yes (0.25)</td>
</tr>
<tr>
<td>Use in typical contexts</td>
<td>Yes (0.25)</td>
<td>Yes (0.25)</td>
</tr>
<tr>
<td><strong>Number of indicators: Absolute coding</strong></td>
<td>7.00</td>
<td>6.00</td>
</tr>
<tr>
<td><strong>Number of Indicators: Weighted coding</strong></td>
<td>7.00</td>
<td>6.33</td>
</tr>
</tbody>
</table>

Note. Absolute coding = the total number of quality indicators fully met; weighted coding = the number of components met for each indicator, divided by the total number of components for each indicator; bolded text = quality indicator; non-bolded text = component for each indicator.
Table 5

**Coding of Essential and Desirable Quality Indicators of Included Group Design Studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Describing participants</td>
<td>No (0.33)</td>
<td>No (0.33)</td>
<td>Yes (1.00)</td>
<td>No (0.66)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Participant disability/difficulty</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Comparable participants across groups</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Comparable interventionists across groups</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>No (0.00)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Implementation of the intervention</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>No (0.33)</td>
<td>No (0.00)</td>
<td>No (0.66)</td>
</tr>
<tr>
<td>Intervention described and specific</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>No (0.00)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Fidelity described and assessed</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Description of practices in comparison group</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Outcome measures</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Multiple measures</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Measures completed at appropriate times</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>No (0.50)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Linked to research question and appropriate unit of analysis</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Effect size calculations</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>No (0.00)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Number of indicators: Absolute coding</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>1.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Number of indicators: Weighted coding</td>
<td>3.33</td>
<td>3.33</td>
<td>3.33</td>
<td>2.16</td>
<td>3.66</td>
</tr>
</tbody>
</table>

Desirable Quality Indicators

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Reliability of estimates</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Outcomes beyond immediate posttest</td>
<td>No (0.00)</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Validity</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Fidelity assessed beyond surface features</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Documentation of comparison conditions</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Audio or video exerts of interventions</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
<td>Yes (1.00)</td>
<td>No (0.00)</td>
<td>No (0.00)</td>
</tr>
<tr>
<td>Clarity of results</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Number of indicators: Absolute coding</td>
<td>6.00</td>
<td>5.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Note. Absolute coding = the total number of quality indicators fully met; weighted coding = the number of components met for each indicator, divided by the total number of components for each indicator; bolded text = quality indicator; non-bolded text = component for each indicator.
Table 5 Continued

**Coding of Essential and Desirable Quality Indicators of Included Group Design Studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Describing participants</strong></td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Participant disability/difficulty</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Comparable participants across groups</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Comparable interventionists across groups</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td><strong>Implementation of the intervention</strong></td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Intervention described and specific</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Fidelity described and assessed</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td>Description of practices in comparison group</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
<td>Yes (0.33)</td>
</tr>
<tr>
<td><strong>Outcome measures</strong></td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Multiple measures</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Measures completed at appropriate times</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td><strong>Data analysis</strong></td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
<td>Yes (1.00)</td>
</tr>
<tr>
<td>Linked to research question and appropriate unit of analysis</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td>Effect size calculations</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
<td>Yes (0.50)</td>
</tr>
<tr>
<td><strong>Number of indicators: Absolute coding</strong></td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Number of indicators: Weighted coding</strong></td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
</tbody>
</table>

**Desirable Quality Indicators**

| Attraction                                                      | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Reliability of estimates                                       | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Outcomes beyond immediate posttest                            | No (0.00)               | No (0.00)              | No (0.00)               | No (0.00)             | No (0.00)             |
| Validity                                                       | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Fidelity assessed beyond surface features                      | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Documentations of comparison conditions                        | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Audio or video excerpts of interventions                       | No (0.00)               | No (0.00)              | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| Clarity of results                                             | Yes (1.00)              | Yes (1.00)             | Yes (1.00)              | Yes (1.00)            | Yes (1.00)            |
| **Number of indicators: Absolute coding**                      | 6.00                    | 6.00                   | 7.00                    | 7.00                  | 7.00                  |

**Note**: Absolute coding = the total number of quality indicators fully met; weighted coding = the number of components met for each indicator, divided by the total number of components for each indicator; bolded text = quality indicator; non-bolded text = component for each indicator.
<table>
<thead>
<tr>
<th>Grade</th>
<th>Study</th>
<th>Skill</th>
<th>Test</th>
<th>Subtest</th>
<th>Effect Size Estimate</th>
<th>Mean Effect Size Estimate Within Studies</th>
<th>Weighted Mean Effect Size Estimates Across Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hedge’s g (CI) 0.05 SE&lt;sub&gt;g&lt;/sub&gt; Hedge’s g SE&lt;sub&gt;g&lt;/sub&gt; 0.697</td>
<td>Hedge’s g (CI) 0.05 SE&lt;sub&gt;g&lt;/sub&gt; Hedge’s g SE&lt;sub&gt;g&lt;/sub&gt; 0.697</td>
<td>Hedge’s g (CI) 0.05 SE&lt;sub&gt;g&lt;/sub&gt; Hedge’s g SE&lt;sub&gt;g&lt;/sub&gt; 0.697</td>
</tr>
<tr>
<td>K</td>
<td>Leafstedt et al. (2004)</td>
<td>PA</td>
<td>Woodcock Johnson 111 Tests of Achievement 111 (2000)</td>
<td>Word Identification</td>
<td>ES = 1.288 (0.678, 1.898)</td>
<td>0.311 1.197 0.308</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word Attack</td>
<td>ES = 1.106 (0.504, 1.708)</td>
<td>0.304</td>
</tr>
<tr>
<td></td>
<td>Solari &amp; Gerber (2008)</td>
<td>AP/PA</td>
<td>Woodcock Johnson (2001)</td>
<td>Word identification</td>
<td>ES = PA = 0.563 (-0.311, 1.437)</td>
<td>0.433 0.184 0.429</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ES = LC = -0.187 (-1.05, 0.676)</td>
<td>0.427</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word Attack</td>
<td>ES = PA = -0.069 (-0.93, 0.792)</td>
<td>0.426</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ES = LC = 0.427 (-0.441, 1.295)</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vadasy &amp; Sanders (2010)</td>
<td>AP/PA</td>
<td>WRMT-R/NU (Woodcock, 1998)</td>
<td>Word Reading*</td>
<td>ES = 0.601 (0.208, 1.002)</td>
<td>0.198 0.601 0.198</td>
<td></td>
</tr>
</tbody>
</table>

Note: CI = confidence interval; PA = phonological awareness; AP = alphabetical principle; WRMT-R/NU = Woodcock Reading Mastery Tests – Revised-Normative Update; Word Reading* = mean of Word Identification and Word Attack subtest standard scores.
**Table 6 Continued**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Study</th>
<th>Skill</th>
<th>Test</th>
<th>Subtest</th>
<th>Effect Size</th>
<th>Weighted Mean Effect Size Estimates Within Studies</th>
<th>Weighted Mean Effect Size Estimates Across Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hedge’s g (CI) 0.05</td>
<td>SE&lt;sub&gt;ma&lt;/sub&gt;</td>
<td>Hedge’s g</td>
<td>SE&lt;sub&gt;ma&lt;/sub&gt;</td>
</tr>
<tr>
<td>Grade One</td>
<td>Linan-Thompson et al. (2006)</td>
<td>AP/PA/FL/COMP</td>
<td>WLPB-R (1991)</td>
<td>Letter-Word Identification</td>
<td>ES = 0.00*</td>
<td>0.316</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word Attack</td>
<td>ES = 0.00*</td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Passage Comprehension</td>
<td>ES = 0.00*</td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td>Vaughn et al. (2006a)</td>
<td>AP/PA/FL/COMP/Vocab</td>
<td>WLPB-R (1991)</td>
<td>Letter-Word Identification</td>
<td>ES = 0.353 (-0.063, 0.769)</td>
<td>0.208</td>
<td>0.317</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word Attack</td>
<td>ES = 0.467 (0.018, 0.916)</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Passage Comprehension</td>
<td>ES = 0.13 (-0.284, 0.544)</td>
<td>0.210</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word Attack</td>
<td>ES = 1.541 (0.793, 2.289)</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Passage Comprehension</td>
<td>ES = 1.325 (0.545, 2.105)</td>
<td>0.39</td>
<td></td>
</tr>
</tbody>
</table>

*Note: CI = confidence interval; * = data unavailable, set to equal zero; AP = alphabetical principle; PA = phonological awareness; FL = fluency; COMP = comprehension; Vocab = vocabulary; WLPB-R = Woodcock Language Proficiency Battery-Revised; WRMT-R = Woodcock Reading Mastery Test-Revised*
Figure 1. PRISMA flowchart (Moher, Liberati, Tetzlaff, & Altman, 2009) depicting systematic search for journal articles for inclusion in literature review. Only articles meeting inclusion criteria were considered.
Appendix B

Is a Phonologically Based Reading Intervention Functionally Related to Improved Reading Comprehension for At-Risk Undergraduate English Language Learners?

Michelle Cook

The Pennsylvania State University
Abstract

Although research has shown that at-risk ELLs in elementary and middle school can benefit from a phonologically based reading intervention to improve reading comprehension skills, there is evidence that learners become more resistant to intervention as they age. This study examined whether there is a functional relation between a phonologically based reading intervention and reading comprehension for at-risk undergraduate ELLs. This quasi-experimental group design study involved 9 participants (treatment = 6) and (counterfactual = 3) from various L1 backgrounds including Chinese, Spanish, and Arabic. Effect sizes were calculated for measures of letter-word identification, passage comprehension, and word attack. Although a significant effect size was calculated for the measure of word attack, no effect was found for either the measure of letter-word identification or passage comprehension. Implications related to intervention dosage and the addition of a more targeted vocabulary component are discussed in relation to future iterations of this study.

*Keywords:* phonologically based reading program, reading comprehension, ELLs, undergraduate
Is a Phonologically Based Reading Intervention Functionally Related to Improved Reading Comprehension for At-Risk Undergraduate English Language Learners?

Reading ability is a pivotal skill for all students. Students who become competent readers have access to content knowledge (Carnine & Carnine, 2004) and are better positioned to succeed academically (Hammer et al., 2014). Indeed, studies have demonstrated a direct and significant relation between reading proficiency and postsecondary academic achievement (Cox, Friesner, & Khayum, 2014). Conversely, students who struggle to become capable readers are not only at-risk academically, but face precarious futures beyond the classroom, extending into their ability to function in society at large (Alexander, 2005).

Research on Reading Interventions for English Language Learners

Given the significant consequences related to the successful acquisition of reading skills, educational research has focused on determining the best pedagogical practices to teach literacy. The majority of early literacy research has focused on teaching reading to native English speakers (August, Shanahan, & Escamilla, 2009; Gyovai, Cartledge, Kourea, Yurick, & Gibson, 2009). The findings of such studies were synthesized by the National Reading Panel (NRP, 2000) and the following five areas of knowledge were determined to be fundamental to reading success: (a) phonological awareness (PA), (b) alphabetical principle (AP), (c) fluency, (d) text comprehension strategies, and (e) vocabulary knowledge (Ehri et al., 2001). Additionally, extant literature indicates that explicit instruction represents a favored pedagogical approach when working with students who are struggling to read (Tam, Heward, & Heng, 2006).

Although much is known about teaching reading to native English speakers, there is a dearth of knowledge relating to the instructional components that contribute to positive reading outcomes for English Language Learners (ELLs) (Hammer et al., 2014; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzakanani, 2003). This insufficiency exists despite the fact that ELLs represent the fastest growing subgroup within the student population in North America.
(Wagner, Francis, & Morris, 2005) and that it is predicted that by 2030, up to 40% of the student population in the United States may be ELLs (U.S. Department of Education, 2003). According to the U.S. Department of Education National Center for Education Statistics (2017), in 2014-2015 approximately 665,000 ELLs were identified as students with disabilities, amounting to 13.8% of the total ELL population enrolled in U.S. public elementary and secondary schools. Given that 56% of ELLs considered for special education services seek remediation for reading problems (Gyovai et al., 2009), there is a need to extend the research to address the learning of ELLs who are at risk of reading failure.

**Learning to Read: Developmental Trajectory**

Prior to examining research relating to reading interventions for ELLs, it is important to establish whether the developmental trajectory for learning to read is the same for ELLs in their L2 (second or additional language) when compared with their native English peers who are learning to read in their L1 (first language). According to the research literature, early reading skills including word and text reading competence are equally attainable for both ELLs and native English speaking populations and achievement levels are similar between the two groups in response to effective instruction (Lovett et al., 2008). As such, Lipka et al. (2005) asserted that the mere fact that ELLs are learning to read in their non-native language does not represent a risk factor for reading disability.

Some divergence in the findings does exist with regard to the development of early phonological processing skills when comparing ELLs to their native English speaking peers (Lovett et al., 2008). According to Lovett et al. (2008), some studies have shown that ELLs in the early elementary grades achieve comparable levels of phonological skills when compared to their native English speaking peers (Chiappe, Siegel, & Wade-Wolley, 2002; Lipka et al., 2005). Whereas Lovett et al. (2008) reported that other studies have demonstrated a delay in the emergence of phonological processing skills for ELLs (Lipka & Siegel, 2007). However, despite
a possible lag in phonological and linguistic processing, a parallel developmental process in terms of the acquisition of basic literacy skills was found (Chiappe et al., 2002). Notably, Chiappe et al. (2002) asserted that for both ELLs and native English speakers, alphabetic knowledge and phonological processing remained salient early literacy skills and concluded that literacy acquisition is likely to follow a similar developmental trajectory for both groups of students. One exception for ELLs may be that alphabetic knowledge may emerge before phonological awareness and in turn facilitate the development of the latter (Chiappe et al., 2002).

Generally, the research indicates that there is parity in terms of the development of early reading skills in ELLs and native English speakers (Lesaux & Siegel, 2003) and English proficiency has not been shown to be a crucial factor in terms of developing early reading skills (Lesaux & Siegel, 2003). In terms of instructional approaches, Tam et al. (2006) noted that there is evidence to suggest that pedagogical practices that have been shown to be efficacious in terms of teaching reading to native English speakers are also effective for ELLs. However, these authors caution that some modifications might be necessary to address the unique learning needs of ELLs (Tam et al., 2006).

Reading Interventions: Language of Instruction

It is also important to address the issue of language of instruction when delivering reading interventions. Although some schools whose ELL populations are homogeneous (e.g. Spanish speaking) have employed bilingual instructional models, schools where the students identify as speaking a vast array of native languages, have primarily provided instruction in English (Calderon, Slavin, & Sanchez, 2011). In fact, Kamps et al. (2007) reviewed the extant research and found evidence to support the use of English interventions. They asserted that instruction of reading skills delivered in English is associated with positive reading outcomes for ELLs, irrespective of their L1. For the purpose of this study, the intervention will be delivered in English.
A Review of the Extant Literature

In reviewing the literature to determine what researchers have reported about the components of reading interventions delivered in English associated with positive outcomes for ELLs identified as at risk of reading failure in elementary and middle school, the findings suggest that a comprehensive instructional package that reflects the essential reading skills identified by the NRP (2000) may be effective. According to the extant literature, components of a comprehensive instructional package for ELLs at risk of failure should include (a) PA, (b) AP, (c) comprehension strategies, (d) fluency, and (e) vocabulary instruction. This finding suggests that intervention packages that have been found to be successful with native English readers are also promising practices for use with ELLs at risk of reading failure. This also applies to a delivery of instruction that is explicit. However, modifications to address the specific needs of ELLs (e.g., vocabulary instruction that includes concrete examples, visuals, and facial expressions) may be advisable in order to differentiate instruction for this population of students.

Pertinent to the research proposed herein, it is important to note that post intervention reading performance outcomes for older ELLs (i.e. middle school and higher) are mixed. For example, two studies involving middle school students (Lovett et al., 2008; Tam et al., 2006) reported positive effects, whereas the results from a study by Denton et al. (2008) suggested that ELLs at risk of reading failure may become less responsive to intervention in the later grades. A discussion of these divergent findings is warranted given that Foorman, Breier, & Fletcher (2003) synthesized the research and found evidence to support the assertion that early intervention is most effective and that older students may become less responsive to intervention. Accordingly, Denton et al. (2008) offered an important glimpse at the resistance to treatment that may develop for ELLs who receive reading interventions in the later grades. Although the study’s data for ELLs and non-ELL participants were only disaggregated for pre-test measures and not post-test measures, it is relevant to this discussion as it emphasizes the importance of providing timely
interventions for ELLs at risk of reading failure. In their study, the authors obtained disappointing results in terms of response to intervention for struggling eighth-grade ELLs. In fact, despite employing an intervention that included all of the components identified by the NRP (2000) for ELLs at risk of reading failure, outcomes for participants in the treatment condition for word recognition, comprehension, and fluency were not significantly greater than the those achieved by the control group. Moreover, they reported that over the course of the study, neither the treatment nor the control group experienced significant growth. The authors emphasized that on all reading measures, students in the treatment group did not show significantly higher results than their peers in the control condition. These findings suggest that ELLs with risk indicators who fail to receive effective early remediation are likely to become more impervious to remediation in the later grades (Gyovai et al., 2009).

Given that postsecondary institutions are recruiting a greater number of students who are considered to be academically underprepared (Mulvey, 2008) including immigrant and refugee ELL students (Kanno & Varghese, 2010), concerns relating to how to effectively address their learning needs are mounting. For example, Perin (2013) reviewed the extant literature relating to the literacy skills of postsecondary academically underprepared students and concluded that there remains a need to learn more about how to address the literacy skills of these students. Additionally, Perin (2013) posited that in terms of reading comprehension, both decoding and linguistic comprehension strategies needed to be explicitly taught to postsecondary students who are academically underprepared such as those with low English language proficiency.

As such, a pilot investigation into the effects of a reading intervention for undergraduate ELLs who self-identify as experiencing difficulty in the area of reading comprehension represents an initial exploratory examination. Given that Lovett et al. (2008) reported that a phonologically based reading program was found to improve reading performance for ELLs with reading
difficulty between grades 2 and 8, this study is designed to examine whether these findings can be extended to older ELLs with reading difficulty. In this study, a phonologically based reading intervention was implemented to address reduced reading comprehension for first year postsecondary ELLs when reading in their L2. The objective was to determine if there is a functional relation between a phonologically based reading intervention and reading fluency for adult ELLs.

Methods

Participants

Participants for this study were chosen following a three step selection process. First, participant recruitment was undertaken at three different levels of student contact. The university’s ELL Program Director sent a recruitment email with the necessary contact information to all students who were deemed suitable candidates for participation in the study. Additionally, instructors who work directly with freshmen ELLs were informed of the study and encouraged to notify suitable study participants of the opportunity. In both cases, the onus was placed on the students to contact the lead researcher via email to communicate their interest in participating in the study. Finally, the lead researcher visited four university level introductory composition classes for ELLs and described the project to the students. A participant sign-up sheet was distributed and interested students provided their contact information. The second step for locating eligible study participants involved the completion of a participant questionnaire (see figure 1). A total of two students contacted by the ELL program director expressed interest in participating in the study and completed the participant questionnaire. Of the approximately 60 students enrolled in the classes visited by lead researcher, 25 initially expressed interest in participating in the study, and five followed through and completed the participant questionnaire. Finally, information obtained using the participant questionnaire was organized into a table to allow for a point-by-point evaluation using inclusion criteria. Inclusion criteria were as follows:
7. Speak a language other than English as L1.

8. Be in the U.S. for two years or more and/or have been enrolled in English classes for more than four years.


Six of the seven students were identified as meeting the inclusion criteria. The student who was excluded from the study did not self-report as having difficulty with reading comprehension in English.

Tables 1 and 2 provide detailed information regarding the participants in both the treatment and counterfactual groups. The treatment participants for this study were six undergraduate university students. The majority of the students (66.7%) were from China and spoke Chinese as their L1. The other participants were from either Bolivia (16.7%) and spoke Spanish as their L1 or Syria (16.7%) and spoke Arabic as their L1. All of the included participants reported experiencing difficulties with reading comprehension in their L2. Time spent in the U.S. ranged from 7 months to five years, however, all participants met the inclusionary criteria when exposure to English through school courses was considered. In terms of gender, 66.7% of the participants were male and 33.3% were female. Their ages at the time of the study were ranged between 18 and 22 years old.

Given the small number of participants and time restrictions due to the end of the school year, the counterfactual group was comprised of three students who had initially volunteered for participation in the study but had not followed through on their intentions. These students were enrolled in the ELL composition classes and expressed a willingness to take the pre and posttests without an opportunity to participate in the intervention. Counterfactual participants had similar demographic profiles when compared with the treatment group. However, all counterfactual participants spoke Chinese as their L1.

**Setting**
This study was conducted at a university campus in the northeastern part of the United States. The campus is located just on the periphery of a medium-sized urban center in a suburban setting. The university campus offers both undergraduate and master level courses. Assessments and intervention sessions for the treatment group (administered by the experimenter) took place in a small office in the English Language Study Centre. The room itself was approximately 7’ by 8’. There were three chairs in the room, a small desk, and the materials needed for the intervention (see materials section below). Only the experimenter and the participant were present during each intervention session. However, an undergraduate student who collected treatment integrity data was present for 20% of the sessions. The pre and posttests for the counterfactual group were administered by the experimenter and data were collected at the university campus library in a study room. The room itself was approximately 8’ by 20’. There were two chairs in the room, a table, and the materials needed for the assessments (see materials section below).

**Materials**

**Measures.** Pre and posttest materials consisted of the *Woodcock Johnson IV Tests of Achievement* (Schrank, Mather, & McGrew, 2014). Specifically, the pretest involved *Form B Standard* and the posttest involved *Form A Standard*. In both cases, the following subtests were utilized: (a) *Test 1 Letter-Word Identification*, (b) *Test 4 Passage Comprehension*, and (c) *Test 7 Word Attack*. These subtests were selected because they are well standardized and extensively used in educational settings (Lovett et al., 2008).

- *Test 1 Letter-Word Identification* was used to assess word reading accuracy. The test required that participants read isolated words aloud. Words were presented in order of difficulty beginning with easier words. The words were presented in printed form. The
participants’ scores were the number of words read correctly. A cut-off rule of six consecutive mistakes was applied.

- **Test 4 Passage Comprehension** was used as a measure of reading comprehension. Participants were asked to read sentences in printed form to themselves and provide the missing words. A cut-off rule of six consecutive errors was applied. The participants’ scores were the number of correct items.

- **Test 7 Word Attack** was used as a measure of decoding. Participants were asked to read nonwords presented in printed form aloud. A cut-off rule of six consecutive errors was applied. The participants’ scores were the number of correct items.

**Reading Intervention.** Intervention materials consisted of lessons drawn from the *Corrective Reading: Decoding Strand* (Engelmann et al., 2008). Each lesson involved the use of a leveled Teacher Presentation Book and a corresponding non-consumable Student Book. Data collection materials included a pencil, a MacBook Air (used to time the participant’s reading trials), and accompanying data-tracking sheets such as fluency charts. A small whiteboard and whiteboard marker with eraser were used for some word attack activities.

**Effect size calculations**

In order to calculate effect size for each subtest, I replicated the procedures employed by Hwang and Riccomini (2016). Hedges’ *g* was calculated for these studies to estimate the effect size related to specific measures of reading skill: letter-word identification, passage comprehension, and word attack. Given that these studies included a control group, effect size was calculated as 

\[ d = \frac{(X_{\text{post.trt}} - X_{\text{pre.trt}}) - (X_{\text{post.ctr}} - X_{\text{pre.ctr}})}{\sqrt{\frac{(n_{\text{trt}}-1)s_{\text{trt}}^2 + (n_{\text{ctr}}-1)s_{\text{ctr}}^2}{n_{\text{trt}}+n_{\text{ctr}}-2}}} \]

For the respective treatment and control groups, \(X_{\text{post.trt}}\) and \(X_{\text{post.ctr}}\) were unadjusted post-test means, \(X_{\text{pre.trt}}\) and \(X_{\text{pre.ctr}}\) were unadjusted pretest means, \(n_{\text{trt}}\) and \(n_{\text{ctr}}\) were sample sizes, and \(s_{\text{trt}}\) and \(s_{\text{ctr}}\) were unadjusted posttest standard deviations. Following the recommendations of Hedges and Olkin
(1985), an unbiased estimator was obtained by multiplying all of the effect sizes by \(1 - \frac{3}{4(n_{tr} + n_{ct})^9}\). Standard errors and confidence intervals were calculated and reported with the corresponding effect sizes.

**Procedures**

**Pretest.** Students in both the treatment and the counterfactual groups completed the pretest. The pretest consisted of three subtests from the *Woodcock Johnson IV Tests of Achievement Form B Standard* (Schrank et al., 2014): (a) *Test 1 Letter-Word Identification*, (b) *Test 4 Passage Comprehension*, and (c) *Test 7 Word Attack*. The experimenter met individually with each participant in an office in the ELL Study Center (treatment group) or in a study room in the library (counterfactual group) with all materials ready including the *Woodcock Johnson IV Tests of Achievement Form B Standard* (Schrank et al., 2014) and the corresponding scoring sheets. The testing book was positioned upright on the table directly between the participant and the experimenter with the pages to be read by the participant facing the participant and the pages to be read by the experimenter facing the experimenter. The scoring sheets were arranged behind the testing book so that they were not visible to the participant. The experimenter followed the script provided in the standard test book to guide all verbal prompts/feedback. The students completed the activities while the experimenter scored each response according to the guidelines indicated by the *Woodcock Johnson IV Tests of Achievement Form B Standard* (Schrank et al., 2014).

**Pre-Intervention.** As required by the *Corrective Reading: Decoding Strand* (Engelmann et al., 2008), a decoding placement test was used to determine the appropriate leveled placement for each participant in the treatment condition to start the program. The decoding placement test required that participants read four passages orally. A single two-sided sheet with the four passages was provided to the participants. The experimenter retained an identical copy of the
single two-sided sheet with the four passages and noted errors directly on this form. The form was positioned so that the participant could not see what the experimenter was writing. The experimenter followed the verbal prompts as directed in the guidelines for administering the test. The following participant responses were noted as errors: (a) adding a word that did not appear in the story; (b) words that were misidentified; (c) any word that was not identified within 3 seconds; (d) a mistake that was subsequently self-corrected by the student; and (e) a word that was sounded out at a rate that was not equivalent to a normal speaking rate that could not be correctly identified by the participant following the experimenter’s prompt “What word?”.

**Intervention.** Intervention sessions took place two times per week for four weeks. Participants began the intervention using the appropriate decoding strand level (see figure 2) as indicated by the decoding placement test. For each intervention session the participant and the experimenter met under the same conditions as described during the pretest. The experimenter followed the scripted lessons in the leveled Teacher Presentation Book. Each lesson, regardless of level, followed a similar structure. The lessons began with exercises targeting word-attack skills using a whiteboard. For example, the experimenter printed a word on the board such as ‘never’ and asked the participant “What word?” Then the experimenter erased the letter ‘n’ and asked the student “What word now?” Then the experimenter added a ‘y’ and asked the student “What word now?” The experimenter followed the following procedure to correct word identification errors:

a) The word is _______.

b) What word? (signal) _______.

c) Spell _______.

d) Go back to the first word in the row/column.

The next exercises involved the introduction of various sound combinations. For example, participants were told that the letters A-I go together and make the sound åâå. The
participants were then presented with a series of words containing the sound combination A-I which was underlined. Participants were asked to say the sound for the underlined part and then read the word. Next, participants read a series of words containing the sound combination A-I without first saying the sound. This process was repeated for other sound combinations, including sound combinations that make the same sound such as I-R, U-R, and E-R.

Exercises involving the presentation of vocabulary and corresponding definitions were presented next. Participants were both presented with the meaning of words as well as asked to use the vocabulary word in a sentence. For example, after learning that the word ‘stout’ is another way of saying ‘strong and heavy,’ participants were asked “What’s another way of saying ‘a strong and heavy man?’” Following the introduction of vocabulary, participants were provided with a list of words that were to be practiced by reading the words orally. Error correction procedures were the same as those described above.

Finally, participants engaged in a story reading. First, they were asked to read the title and infer the plot presented in the passage. Next, participants read the passage. The experimenter would stop the participants and ask comprehension questions throughout the reading of the text as indicated in the teacher presentation book. To correct word reading errors, the following procedures were used:

a) The word is _______.

b) Touch under that word.

c) What word?

d) Go back to the beginning of the sentence and read that sentence again.

Lastly, a fluency assessment was conducted using the same passage. Participants were provided 2-minutes and were asked to “read like they talk.” Errors were scored by the experimenter and the same scoring procedures as those described for the decoding placement test (described above) were employed. At the end of the 2-minute fluency assessment, the
participants used an individual reading progress chart to graph the number of words read during the 2-minutes and the number of errors.

**Posttest.** Students in both the treatment and the counterfactual groups completed the posttest. The posttest consisted of three subtests from the *Woodcock Johnson IV Tests of Achievement Form A Standard* (Schrank et al., 2014): (a) *Test 1 Letter-Word Identification*, (b) *Test 4 Passage Comprehension*, and (c) *Test 7 Word Attack*. The experimenter met individually with each participant in an office in the ELL Study Center (treatment group) or in a study room in the library (counterfactual group) with all materials ready including the *Woodcock Johnson IV Tests of Achievement Form A Standard* (Schrank et al., 2014) and the corresponding scoring sheets. The testing book was positioned upright on the table directly between the participant and the experimenter with the pages to be read by the participant facing the participant and the pages to be read by the experimenter facing the experimenter. The scoring sheets were arranged behind the testing booklet so that they were not visible to the participant. The experimenter followed the script provided in the standard test book to guide all verbal prompts/feedback. The student completed all activities while the experimenter scored each response according to the guidelines indicated by the *Woodcock Johnson IV Tests of Achievement Form A Standard* (Schrank et al., 2014).

**Dependent Variable**

The dependent variable (DV) for this study was reading comprehension. Reading comprehension was operationally defined as the ability to extract meaning from English text. Reading comprehension was measured using the *Woodcock Johnson IV Tests of Achievement Form B Standard* (Schrank et al., 2014) *Test 4 Passage Comprehension*. However, given that efficient word identification and word-attack skills are hypothesized to reduce demands on working memory and subsequently support reading comprehension (Cain, Oakhill, & Bryant, 2004), measures of these two skills were also taken.
Design

A group quasi-experimental research design was employed in order to determine if there exists a functional relation between a phonologically based reading intervention and reading comprehension for at-risk undergraduate ELLs. Participants in the treatment condition received 30 minute intervention sessions twice per week for four weeks. Pretesting occurred before treatment sessions began with all participants and post-testing occurred following approximately a total of 320 minutes of intervention per student in the treatment group. All pretests, post-tests, and intervention sessions were administered by the lead researcher. Additionally, treatment integrity data were collected for 20% of sessions.

Treatment Integrity

A treatment integrity checklist was developed by the experimenter (see figure 2). Treatment integrity data was obtained for 20% of intervention sessions. Treatment integrity was measured using a point by point format where the independent observer scored treatment integrity on 14 items. Items could be scored as either ‘not observed’ or receive a score of one (inappropriate), two (somewhat appropriate), and three (appropriate). The total number of possible points per session was 42. To obtain a treatment integrity score, the independent observer divided the total number of points by the total number of items. An independent scorer was trained in data collection prior to the commencement of the intervention. Practice data collection sessions were scheduled to allow the observer the opportunity to request clarifications. The independent scorer was a senior undergraduate student who had been accepted into a Master of Clinical Psychology program for the fall semester. Average treatment integrity was 98.81% for all observed sessions.

Results

Table 3 contains a comparison of the means, standard deviations, sample sizes, effect sizes, and standard errors for the treatment and counterfactual groups.
Decoding. The Word Attack mean at pretest of 21.8 for the treatment group corresponds roughly to a grade equivalent (GE) of 4-9. The Word Attack mean at posttest of 26 for the treatment group corresponds roughly to a GE of 9-2. Thus indicating a change in GE of 4-3. All treatment participants scored below their current grade equivalent level at pretest, with three participants scoring at or below a GE of 3-6 at pretest indicating initial difficulties in decoding. The Word Attack mean at pretest of 23 for the control group corresponds roughly to a grade equivalent (GE) of 5-7. The Word Attack mean at posttest of 25.3 for the control group corresponds roughly to a GE of 7-6. Thus indicating a change in GE of 1-9. All control participants scored at or below a GE of 6.7 at pretest, indicating initial difficulties in decoding.

Word Reading. The Letter-Word Identification mean at pretest of 66.3 for the treatment group corresponds roughly to a GE of 7-7. The Letter-Word Identification mean at posttest of 66.5 for the treatment group corresponds roughly to a GE of 9-6. Thus indicating a change in GE of 1-9. All treatment participants scored below their current grade equivalent level at pretest indicating initial difficulties in word reading. The Letter-Word Identification mean at pretest of 66.3 for the control group corresponds roughly to a GE of 8-9. The Letter-Word Identification mean at posttest of 64 for the control group corresponds roughly to a GE of 7-6. Thus indicating a change in GE of negative 1-3. All control participants scored below their current grade equivalent level at pretest indicating initial difficulties in word reading.

Reading Comprehension. The Passage Comprehension mean at pretest of 33.8 for the treatment group corresponds roughly to a GE of 4-5. The Passage Comprehension mean at posttest of 38.8 for the treatment group corresponds roughly to a GE of 7-7. Thus indicating a change in GE of 3-2. No participant in the treatment group scored above a GE of 6-3 at pretest indicating initial difficulties in reading comprehension. The Passage Comprehension mean at pretest of 34.7 for the control group corresponds roughly to a GE of 4-5. The Passage Comprehension mean at posttest of 39.3 for the control group corresponds roughly to a GE of 7-
Thus indicating a change in GE of 3-2. No participant in the treatment group scored above a GE of 5-7 at pretest indicating initial difficulties in reading comprehension.

**Effect size.** The result for the Word Identification subtest indicates that the sample estimate for the effect of the reading intervention on study participants is 0.372 SD, but the effect of the reading intervention on study participants in the population (from which the sample was drawn) has a range (95% CI = -0.826, 1.664). The result for the Passage Comprehension subtest indicates that the sample estimate for the effect of the reading intervention on study participants is 0.043 SD, but the effect of the reading intervention on study participants in the population (from which the sample was drawn) has a range (95% CI = -1.185, 1.281). Finally, the result for the Word Attack subtest indicates that the sample estimate of the effect of the reading intervention on study participants is 0.288 SD, but the effect of the reading intervention on study participants in the population (from which the sample was drawn) has a range (95% CI = -0.921, 1.583).

Anecdotally, the experimenter observed an increase in the confidence level demonstrated by the participants in the treatment group when reading aloud. The students were noted to read with a louder voice and hesitate less frequently as they progressed through the intervention. Participants demonstrated generalization of the learning from previous reading sessions to new cold readings. For instance, participants frequently experienced difficulty with novel vocabulary introduced in the lessons, however, when these words were repeated in the reading for the following sessions, the participants were able to read these words accurately.

**Social Validity**

The participants were asked to complete a short Social Validity Questionnaire following the final intervention session (see Figure 3). The participants rated eight statements between one (strongly disagree) and five (strongly agree). All of the participants reported that they enjoyed participating in the reading study (average score = 5). Participants reported that the reading activities were interesting (average score = 4.8) and that the individual reading sessions lasted an
appropriate amount of time (average score = 4.8). In terms of their perceptions of the impact of the reading intervention, participants reported: (a) that they felt that their reading fluency had improved over the course of the program (average score = 4.5), (b) that their reading comprehension had improved (average score = 4.5), and that overall they felt more confident about their reading skills in English (average score = 4.7). Participants indicated that they would like to continue with more sessions of the reading intervention (average score = 4.8). Anecdotally, one student asked the experimenter if the reading intervention sessions could be continued during the upcoming fall semester. Finally, all of the participants indicated that they would recommend the reading intervention to other adult ELLs who would like to improve their reading comprehension (average score = 4.7). Participants were also given the opportunity to make additional comments. One student stated that “The program was really good and it is pay off to your effort.” It is worth noting that two students did state that they would like to see more opportunity to practice vocabulary and the meaning of words. Furthermore, one student noted that worksheets that could support supplementary practice at home would be helpful. Such feedback will be taken into consideration by the experimenter to inform further iterations of the reading study. The results of the participant questionnaire support the social validity of the phonologically based reading intervention to improve the reading comprehension skills of at-risk adult ELLs within the university context.

Discussion

The objective of this study was to determine if there is a functional relation between a phonologically based reading intervention and reading comprehension for at-risk undergraduate ELLs. The current results cannot support a conclusion of a functional relation between the independent and dependent variables, however, there may be reason for cautious optimism.

Given that the present study employed a phonologically based reading intervention, the lessons that were used during the treatment sessions placed a significant emphasis on decoding
skills. This translates directly into the findings described above where a small effect size was found for both the Word Identification and Word Attack subtests. According to Cain et al. (2004), comprehension is impacted by word reading that is either slow or inaccurate. These authors posited that text comprehension is curtailed due to the demands placed on the working memory which limit the functioning of other necessary processes such as integration and inference (Cain et al., 2004). Additionally, Burt, Peyton, and Adams (2003) noted that adult ELLs should be taught to decode. Therefore, considering that the participants in the treatment condition demonstrated some improvement in their decoding skills, amelioration in reading comprehension performance might have been expected. However, improvements in decoding did not result in a discernable improvement in reading comprehension. It could be posited that the lack of improvement seen in reading comprehension performance may be due to the truncated treatment time of four weeks. In order to observe an impact of improved decoding skills on passage comprehension, more time in treatment may be necessary given that many reading intervention studies in the extant literature describe a treatment dosage of approximately 4-6 months. For example, Lovett et al. (2008) employed three phonologically based reading intervention programs (including Corrective Reading: Decoding Strand (Engelmann et al., 2008)) with ELLs in grades 2 through 8 who had been identified as experiencing reading difficulty. Although these researchers found that the use of these programs resulted in significant improvements in reading achievement for the participants, it is important to note that the participants received approximately 105 intervention hours.

Notably, a correlation between working memory and performance on standardized assessments of comprehension skill for postsecondary students was noted by Daneman and Carpenter (1983). This finding is echoed by Nouwens, Groen, and Verhoeven (2017) who concluded that working memory is a reliable predictor of differences in reading comprehension for adults as well as children. As such, it is possible that as student performance improves in the
areas of letter-word identification and word attack with the provision of greater treatment dosage, reading comprehension may increase as a result of reduced demands on working memory.

One caveat to this hypothesis (and an opportunity to adapt future iterations of this study) is that Cain et al. (2004) noted that for L1 struggling readers, as they aged and saw improvements in word reading performance, performance in reading comprehension was not necessarily concomitant with this skill improvement. One important variable identified by these authors was vocabulary knowledge. Similarly, Burt et al. (2003) explained that ELLs who are able to decode efficiently, may still experience reduced performance in reading comprehension if they experience difficulty identifying the meaning of the vocabulary contained within the text. As such, a greater emphasis on vocabulary development may be warranted in addition to the phonologically based reading intervention.

**Participants.** The recruitment of participants for this study relied on student self-reports of reading difficulty. Such methodology may elicit concerns related to the subjectivity of the assessment of the initial reading skill level of participants. However, Parrila, Georgiou, and Corkett (2007) investigated the accuracy of using student self-reports to identify university students with a history of reading difficulty and determined that employing student self-reports is a viable method for accurately locating individuals with a history of reading problems.

The results of this present study and others demonstrate that there is a need for further investigation regarding interventions to build reading comprehension in at-risk undergraduate ELLs. Although the present data does not provide evidence to indicate a possible functional relation between a phonologically based reading program and reading comprehension for undergraduate at-risk ELLs, there was a small effect on word identification and word attack performance. Research that increases treatment dosage is required to examine whether improved performance in areas such as decoding and word identification will support reading
comprehension performance. Additionally, the inclusion of a vocabulary development component may improve reading comprehension performance.

**Limitations of this study**

A disadvantage related to the use of a group design for this particular study is associated with the reduced number of participants. As such, the limited number of participants in this present study will affect the statistical power of the findings. Additionally, the treatment group consisted of individuals with divergent L1 backgrounds, whereas the counterfactual group consisted of individuals from a singular L1 background. The duration of the overall study is also a limitation as more intervention sessions might have led to stronger improvements reflected in the data. Although speculative, it could be hypothesized that the short duration of this study may explain why the intervention was found to only have had an effect on decoding skills and not reading comprehension. Perhaps greater gains might have been observed if the participants would have benefited from a longer duration in intervention. An uncontrollable threat to the internal validity of the study was that the participants were concurrently participating in a variety of English university courses. As it was not possible to require students to be truant from class for the duration of the study, it is not feasible to eliminate learning from other classes as a confounding variable. For example, a student in the treatment group was enrolled in an introductory psychology class during the study. This course involved extensive reading assignments and the writing of response activities which may have supported the student’s reading skills.

**Conclusion**

This study represents an initial investigation into the possible functional relation that may exist between a phonologically based reading intervention and reading comprehension for at-risk undergraduate ELLs with self-reported reading difficulty. Although more research into this relationship is required, initial findings support the use of a phonologically based reading
program to improve students’ decoding skills which is a contributing variable to reading comprehension. However, more research over an extended period of time is needed to examine whether this improvement will generalize into improved reading comprehension results. If so, educators and those working with adult ELLs might be advised to use a phonologically based reading program build reading comprehension in their students.
References


Nouwens, S., Groen, M. A., & Verhoeven, L. (2017). How working memory relates to
children’s reading comprehension: The importance of domain-specificity in storage and processing. *Reading and Writing, 30*, 105-120.


## Table 1

*Treatment Group Demographic and Reading Information*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Country of Origin</th>
<th>L1</th>
<th>Time in U.S.</th>
<th>Self-Reported Concerns Related to Reading</th>
<th>CR</th>
<th>Decoding program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19</td>
<td>Bolivia</td>
<td>Spanish</td>
<td>7 months (However 80% of high school day spent in English)</td>
<td>Slow pace Difficulty with reading comprehension</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>China</td>
<td>Chinese</td>
<td>2 years</td>
<td>Slow pace Reading comprehension Forgetting what was read in the previous sentence after reading the following sentence.</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>China</td>
<td>Chinese</td>
<td>5 years</td>
<td>Slow pace Reading comprehension</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>Syria</td>
<td>Arabic</td>
<td>2.5 years</td>
<td>Reading comprehension</td>
<td>B2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>China</td>
<td>Chinese</td>
<td>7 months (has taken English classes since elementary school)</td>
<td>Reading comprehension</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>China</td>
<td>Chinese</td>
<td>7 months (has taken English classes since elementary school)</td>
<td>Reading comprehension Difficulty remembering what was read. Need to reread text to aid with comprehension.</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

*Note. CR = Corrective Reading: Decoding Strand (Engelmann et al., 2008)*
Table 2

Counterfactual Group Demographic and Reading Information

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Country of Origin</th>
<th>L1</th>
<th>Time in U.S.</th>
<th>Self-Reported Concerns Related to Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
<td>China/Hungary</td>
<td>Chinese</td>
<td>More than 2  years</td>
<td>Slow pace Difficulty with reading comprehension</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>China</td>
<td>Chinese</td>
<td>7 months     (courses in English for more than 4 years)</td>
<td>Difficulty with reading comprehension (especially nonfiction texts such as research papers and history textbooks)</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>China</td>
<td>Chinese</td>
<td>4 years</td>
<td>Difficulty with reading comprehension Tried using a translator when reading, but still experienced difficulty with comprehension.</td>
</tr>
</tbody>
</table>

*Note.* L1 = first language
Table 3

*Descriptive and Inferential Statistics*

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Comparison</th>
<th>Treatment</th>
<th></th>
<th></th>
<th>Control</th>
<th></th>
<th></th>
<th>Effect Size Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>√s²</td>
<td>n</td>
<td>x</td>
<td>√s²</td>
<td>n</td>
<td>g_H</td>
<td>SE_g</td>
</tr>
<tr>
<td>Word Identification</td>
<td>Treatment vs. Control</td>
<td>66.5</td>
<td>13.91</td>
<td>6</td>
<td>64</td>
<td>3.742</td>
<td>3</td>
<td>0.372</td>
</tr>
<tr>
<td>Passage Comprehension</td>
<td>Treatment vs. Control</td>
<td>38.8</td>
<td>9.739</td>
<td>6</td>
<td>39.3</td>
<td>2.161</td>
<td>3</td>
<td>0.043</td>
</tr>
<tr>
<td>Word Attack</td>
<td>Treatment vs. Control</td>
<td>26</td>
<td>6.928</td>
<td>5</td>
<td>25.3</td>
<td>1.634</td>
<td>3</td>
<td>0.288</td>
</tr>
</tbody>
</table>
Figure 1. Participant checklist. This figure provides an exemplar of the checklist.

PARTICIPANT CHECKLIST

Title of Project: Multicomponent Reading Intervention: Effects on the Reading Comprehension of At-Risk ELLs

Your participation in this questionnaire and willingness to provide the requested information is strictly voluntary. Refusal to complete this form will involve no penalty or loss of benefits you would receive otherwise.

Please circle the appropriate response.

1. Are you currently enrolled as an undergraduate student at Penn State Erie? Yes No

2. Are you currently 18 years of age or older? Yes No
   - If you circled “yes” for question 2, please indicate your age:________

3. Was the primary language spoken at home with your parent(s) when you were first learning to speak a language other than English? Yes No
   - If you circled “yes” for question 3, please indicate what primary language you spoke at home with your parent(s) when you were first learning to speak:________

4. Have you lived in the United States for at least 2 years? Yes No
   - If you circled “yes” for question 4, please indicate the number of years you have lived in the United States:________

5. Do you experience difficulty reading in English? Yes No
   - If you circled “yes” for question 5, please provide more specific information (E.g. “When I read text in English, I have difficulty understanding the meaning of what I have read” or “I read slowly in English and this can cause me difficulty making sense of what I have read”).

6. Are you able to understand what is being asked in order for you to participate in this study? Yes No
Figure 2. Treatment Fidelity Observation Form. This figure provides an exemplar of the form.
Figure 3. Social Validity Questionnaire. This figure provides an exemplar of the questionnaire.

**Participant Feedback Questionnaire**

Thank you for participating in this reading study. In order to fully ascertain the value of this reading project, I would like to understand your personal impressions of the reading activities. Your feedback will remain anonymous and will be used to inform how future reading interventions similar to this one should be conceived and carried out. Your completion of this questionnaire is completely voluntary.

Please read the following statements and circle the appropriate number.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
<td>Disagree</td>
<td>Neither agree nor disagree</td>
<td>Agree</td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

a) I enjoyed participating in this reading project.  

b) Overall, the reading activities were interesting.  

c) The individual reading sessions lasted an appropriate amount of time.  

d) I feel that my reading fluency (the ability to read words quickly and accurately) improved over the course of the sessions.  

e) I feel that my ability to understand what I read in English improved over the course of the sessions.  

f) Overall, I feel more confident about my reading skills in English.  

g) Given the opportunity, I would like to continue with more sessions of this reading program.  

h) I would recommend this reading program to others.  

Is there any other feedback that you would like to provide regarding the reading program?
APPENDIX C

Social Validity Questionnaire

Name:____________________________________________________
Date:_____________________________________________________

I would very much like to hear from the participants in this reading study. It is helpful for me to understand your impressions of the reading interventions.

Please answer the following questions:

1. How many times did you view each intervention video? _________

2. Did you refer back to any parts of the video lessons when you were reading an article?

   Yes    No

3. If yes, what part(s)?

   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________

4. Do you think that the video lessons helped you to better understand the related reading?

   Yes    No

5. If yes, what specifically helped?

   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________

6. What parts of the video lessons did you find to be helpful? Why?

   __________________________________________________________
7. What parts of the video lesson did you find to be less helpful? Why?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

8. Is there anything about the video lessons that you would change?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

9. Is there anything else that you would like to say about the video lessons?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
APPENDIX D

Timeline and Schedules for the Intervention

Study Timeline

The following is the timeline of events for the study:

**Participant Recruitment:**

Week 1: Recruitment via Zoom during regular class times.

Weeks 2: First one-on-one meeting between potential participants and the interventionist via Zoom. Implied consent is obtained and screening questions are asked to ensure that potential participants meet the criteria for inclusion in the study.

**Pretest**

Weeks 3: Pretesting involving the Woodcock Johnson IV Subtest 7 Word Attack (WJ-IV subtest 7), Test of Knowledge of English Morphology, Test of Silent Contextual Reading Fluency – 2 (TOSCRF-2).

**Reading Intervention:**

Week 5: **Intervention 1: Introduction to Morphology and Phonological Awareness**
- Participants view the Intervention Video 1
- Participants complete the Intervention Video Quiz 1 to verify understanding of video.

Week 6: **Intervention 2: Review of Morphology and Phonological Awareness and Application to the Essay by Author Anzaldua**
- Participants view the Intervention Video 2.
- Participants complete the Intervention Video Quiz 2 to verify understanding of video.

Week 7: **Intervention 3: Review of Morphology and Phonological Awareness and Application to the Essay by Author Lippi-Green**
- Participants view Intervention Video 3
- Participants complete the Intervention Video Quiz 3 to verify understanding of video.

Week 8: **Intervention 4: Review of Morphology and Phonological Awareness and Application to the Essay by Author Fillmore**
- Participants view Intervention Video 4
- Participants complete the Intervention Video Quiz 4 to verify understanding of video.
Posttest

Participants complete the Social Validity Questionnaire

*All pretests and post-tests will be individually administered by the interventionist using Zoom and a document camera. Intervention videos will be pre-recorded using Kaltura and will be emailed to participants according to the schedule outlined above.
APPENDIX E

Treatment Fidelity Checklists

Treatment Fidelity: Pretest and Posttest

*Please circle the correct response as observed in the intervention videos.

(Note: 1 = yes and 2 = no)

Completed by:______________________________

Date:_____________________________________

Test of Knowledge of English Morphology

1. Did the interventionist provide the following directions:

   “You will hear a derived word (e.g., discussion) followed by a lean sentence context (e.g., What did he want to _______?). Please provide the word that fits in the sentence (i.e., discuss) by writing the word in the space provided and saying the word orally”.

   1   0

2. Did the interventionist read each derived word and each lean sentence aloud?

   1   0

3. Did the interventionist provide time for the participant to print their answer and say their answer aloud?

   1   0
Treatment Fidelity: Pretest and Posttest

*Please circle the correct response as observed in the intervention videos.

(Note: 1 = yes and 2 = no)

Completed by: ________________________________

Date: ______________________________________

TOSCRF-2

1. Students were provided with the opportunity to practice placing lines between words using the annotation feature prior to viewing the TOSCRF-2 Student Record Form via Zoom.

1 0

2. Students were provided with the TOSCRF-2 Student Record Form via Zoom.

1 0

3. The following instructions were provided by the interventionist: “You will have 3 minutes to find all the words in as many sentences as you can. When I say, ‘Begin’, work as fast as you can without making mistakes. When you get to the bottom of the page, go to the next page and keep working until I say, “Stop,” or you come to the stop sign at the end of the test. Begin.”

1 0

4. Students were provided with 3 minutes to complete the activity.

1 0
**Treatment Fidelity: Pretest and Posttest**

*Please circle the correct response as observed in the intervention videos. (Note: 1 = yes and 2 = no)*

Completed by: __________________________

Date: __________________________

**WJ-IV – Subtest 7 Word Attack**

1. Students were provided with the following instructions by the interventionist: “Read the following nonsense words aloud”.

   | 1 | 0 |

2. Students were provided with the WJ-IV- Subtest 7 Word Attack Student Record Form via Zoom.

   | 1 | 0 |

3. A cut-off rule of 6 consecutive errors was applied.

   | 1 | 0 |
Treatment Fidelity: Interventions

Applicable to all 4 intervention videos.

*Please circle the correct response as observed in the intervention videos.

(Note: 1 = yes and 2 = no)

Completed by:______________________________
Date:_____________________________________

1. Did the interventionist introduce/or review how to conduct a morphological analysis?
   1  0

2. Did the interventionist demonstrate how morphological analysis can be applied to vocabulary in general/or from a reading?
   1  0

3. Did the interventionist introduce/or review word-attack skills?
   1  0

4. Did the interventionist demonstrate how word-attack skills can be applied to general vocabulary/or specific vocabulary from the essay?
   1  0
TABLE 1-1

Summary of Descriptive Statistics

Proximal DVs: Decoding and Knowledge of English Morphology

<table>
<thead>
<tr>
<th>Participant</th>
<th>WJ-IV Subtest 7</th>
<th>Test of English Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td></td>
<td>Raw score</td>
<td>AE</td>
</tr>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td></td>
<td>Raw score</td>
<td>AE</td>
</tr>
<tr>
<td>A</td>
<td>17</td>
<td>8-0</td>
</tr>
<tr>
<td>B</td>
<td>19</td>
<td>8-7</td>
</tr>
<tr>
<td>C</td>
<td>17</td>
<td>8-0</td>
</tr>
</tbody>
</table>

Note. AE = age equivalent; GE = grade equivalent

Distal DV: Reading Comprehension

<table>
<thead>
<tr>
<th>Participant</th>
<th>TOSCRF-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
</tr>
<tr>
<td></td>
<td>Raw score</td>
</tr>
<tr>
<td></td>
<td>Score</td>
</tr>
<tr>
<td>A</td>
<td>16</td>
</tr>
<tr>
<td>B</td>
<td>78</td>
</tr>
<tr>
<td>C</td>
<td>79</td>
</tr>
</tbody>
</table>

Note. AE = age equivalent; GE = grade equivalent
CURRICULUM VITA

MICHELLE J. COOK
mbc5376@psu.edu

ACADEMIC HISTORY
2017 - present  PhD Candidate, Special Education, Penn State University
2015 - 2016  Master in Special Education, Pennsylvania State University
2014 - 2016  Graduate Certificate in Applied Behavior Analysis, Pennsylvania State University
1996 - 2000  Bachelor of Education, York University, Toronto, Canada
1995 - 2000  Honors Bachelor of Arts with Specialization (Major: French Language Arts), York University, Toronto, Canada

SELECTED HONORS AND AWARDS
2020  Student Dissertation Research Initiation Grant, Penn State College of Education
2019 – 2020  Graduate Assistantship, Department of Educational Psychology, Counseling, and Special Education, Penn State University
2017 - 2019  Dean's Graduate Assistantship for Engaged Scholarship and Research in Education, College of Education, Penn State University
2017  Graham Fellowship Endowment, Penn State University
1995 - 1999  Member of the Principal’s Honor Roll, York University
1998  York University Continuing Student Scholarship
1996  York University Continuing Student Scholarship
1996  Glendon College Discretionary Scholarship, York University
1995  York University Entrance Scholarship

RESEARCH INTERESTS
The development of evidence-based practices related to interventions for ELLs with disabilities or at-risk of reading and writing difficulties.

ADDITIONAL RELATED EXPERIENCE