**Distinguishing Difficulty From Disability: Next Steps in Improving the Identification of Reading Disabilities in English Learners**

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Mesite, Laura. 2020. Distinguishing Difficulty From Disability: Next Steps in Improving the Identification of Reading Disabilities in English Learners. Doctoral dissertation, Harvard Graduate School of Education.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364527">https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364527</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Distinguishing Difficulty from Disability: Next Steps in Improving the Identification of Reading Disabilities in English Learners

Laura Mesite

Dr. Meredith Rowe
Dr. Gigi Luk
Dr. Dana McCoy
Dr. Joanna Christodoulou

A Thesis Presented to the Faculty of the Graduate School of Education of Harvard University in Partial Fulfillment of the Requirements for the Degree of Doctor of Education

2020
DEDICATION

To my parents, Richard and Diane Mesite, for instilling in me a strong value system, sparking my passion for social justice, and sacrificing so much to support my education. I promise to use my position of incredible privilege for the betterment of society.
ACKNOWLEDGEMENTS

I’d like to thank the many people who supported me throughout this doctoral journey. First and foremost, my dissertation committee. Dr. Gigi Luk, I’m so grateful for the mentorship you’ve provided over the past six years. Our shared academic backgrounds and similar value system provided for a fruitful scholarly and personal relationship that I will continue to cherish. Thank you for your financial and intellectual contributions to this project. Dr. Meredith Rowe, thank you for always supporting my work. I’m incredibly appreciative of the role you’ve played in building my confidence as a scholar. Dr. Joanna Christodoulou, I’ve enjoyed the many opportunities that you have provided me, and I am thankful for your mentorship throughout my doctoral career. Thank you for collaborating with me on survey development and dissemination for Study 1. Dr. Dana Charles McCoy, thank you for your ample methods advice and willingness to listen to me when I was struggling. I am also thankful for the guidance of other faculty members, including Dr. James Kim and Dr. Pamela Mason. I’d also like to acknowledge the Mind Brain Behavior Initiative, Dean Ryan, and the family of Cheng Yu Tung for providing funding for this research as well as the Center for Brain Science and FAS Research Computing staff for their access to neuroimaging resources.

As for my friends, I am especially appreciative of my research partner, Sibylla Leon Guerrero, who conducted all of the behavioral testing in English and Spanish, programmed the fMRI task, and collected the fMRI data with me for Study 2. Sibylla has been an invaluable colleague and friend throughout my time at HGSE. I’d also like to thank Dr. Luk’s former research assistants, Valerie Woxholdt and Rachel Hantman for their help with data collection and double-scoring. I would also like to thank my former Haskins colleagues, Dr. Jeffrey G. Malins and Dr. Peter Molfese, for their fMRI analysis advice. I am so thankful for all of my wonderful HGSE colleagues who have supported me through struggles and triumphs, especially Dr. Mary Burkhauser, Dr. Courtney Pollack, Dr. April Choi, Dr. Hadas Eidelman, Sarah Surrain, So Yeon Shin, Mariam Dahbi, Gladys Aguilar, Catherine Armstrong, and my amazing cohort members. I would also like to acknowledge my friends from home who have provided moral support, including Dale Baldwin, Nicole Ferri, as well as Jean and John Razzano. I’d also like to thank various HGSE staff members including Anne Blevins, Joelle Mottola, Andrena Mason, Dinesh Thapa, Pasquale DiCenzo, Pete Shapiro, Mirabel Batista, and Nikki Burns for keeping my spirits up throughout this process. I am also incredibly grateful for the clinicians, educators, families, and students who took the time out to participate in these studies. Lastly, I’d like to thank my family, including my parents, twin brother, in-laws, late grandparents, and especially my amazing wife, Lacey, who relocated to MA; took on more than her fair share of the chores; and developed a strange affinity for watching football during the nights and weekends that I was busy working. I don’t know if I could have done it without any of these folks, but I’m glad that I didn’t have to try!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>V</td>
</tr>
<tr>
<td>GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>STUDY 1:</td>
<td>13</td>
</tr>
<tr>
<td>Abstract</td>
<td>14</td>
</tr>
<tr>
<td>Introduction</td>
<td>16</td>
</tr>
<tr>
<td>The Present Study</td>
<td>30</td>
</tr>
<tr>
<td>Methods</td>
<td>31</td>
</tr>
<tr>
<td>Results</td>
<td>41</td>
</tr>
<tr>
<td>Discussion</td>
<td>82</td>
</tr>
<tr>
<td>STUDY 2:</td>
<td>91</td>
</tr>
<tr>
<td>Abstract</td>
<td>92</td>
</tr>
<tr>
<td>Introduction</td>
<td>94</td>
</tr>
<tr>
<td>The Present Study</td>
<td>107</td>
</tr>
<tr>
<td>Methods</td>
<td>111</td>
</tr>
<tr>
<td>Results</td>
<td>134</td>
</tr>
<tr>
<td>Discussion</td>
<td>150</td>
</tr>
<tr>
<td>GENERAL DISCUSSION</td>
<td>167</td>
</tr>
<tr>
<td>APPENDIX A: READING DIAGNOSTICS SURVEY</td>
<td>184</td>
</tr>
<tr>
<td>APPENDIX B: BRAIN-BEHAVIOR CORRELATIONS</td>
<td>236</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>239</td>
</tr>
</tbody>
</table>
ABSTRACT

English Learners (ELs) comprise approximately 10% of the U.S. public school population and face the daily challenge of acquiring skills and content in a developing second language. These students often experience difficulties with reading comprehension, even after acquiring a high level of conversational English and adequate word-level reading skills. Given that a majority of disciplinary content is learned through reading, ELs are at risk of school failure since English is the major instructional language in public schools. Those adequately supported through second language services have the potential to achieve academic success. However, like their English-proficient peers, a portion of ELs need additional support when they struggle with learning due to a reading disability (RD). Nevertheless, it is especially challenging to identify RDs in this population, as practitioners must rule out second language acquisition, cultural misunderstandings, inadequate instruction, and socioeconomic disadvantage as primary sources of reading difficulties.

This thesis aims to contribute to the body of literature that addresses this challenge. The first study utilizes an online survey of 598 school- and clinic-based practitioners from across the U.S. to provide a descriptive investigation of the current roles of professionals in the identification process; range of assessments and methods used; and reported barriers to accurate and timely identification. Findings suggest that there is extensive variability in the methods
involved, but insufficient training, bilingual practitioners, and valid and reliable assessments for ELs are major professional barriers.

The second study compares behavioral ($n=76$) and neural ($n=46$) measures of reading processes among adolescents who are Spanish-speaking ELs, Spanish-English bilinguals, typically-developing monolingual English speakers, or monolingual English speakers with an RD. While typically-developing monolingual and bilingual adolescents generally demonstrated adequate reading skills in English, ELs and monolinguals with an RD demonstrated difficulties across a variety of English word reading measures. ELs also demonstrated substantially lower performance than all other groups on measures of English vocabulary and reading comprehension despite adequate performance on these measures in Spanish. Generally, participants shared similar neural correlates of component word reading processes in English. Taken together, this dissertation aims to provide a foundation for future research on this important topic.
GENERAL INTRODUCTION

English Learners¹ (ELs) are one of the fastest growing subgroups of students in the U.S., comprising approximately 10% of public school students (McFarland et al., 2019). The federal government defines ELs as elementary or secondary school students who speak a language other than English as the first or native language and whose difficulties with speaking, reading, writing, or understanding English, may hinder their ability to succeed in school (ESEA Section 8101(20)). While an estimated 22% of children in the United States speak a language other than English at home (U.S. Census Bureau, 2019), ELs are the subset of these children who demonstrate limited English proficiency as a result of less exposure to English. Although approximately 77% of ELs in the U.S. school context are Spanish-speakers (McFarland et al., 2019), they are a fairly heterogeneous group comprised of various racial and cultural backgrounds, different circumstances for limited English proficiency, and varying levels of prior schooling and/or literacy in their non-English language(s). Over three-quarters of ELs were born in the U.S. (U.S. Census Bureau, 2013), yet this group also includes immigrants who may enter U.S. schools at any grade level. One commonality is that these students are often instructed in English-only mainstream classrooms and are therefore tasked with the challenge of learning skills and content vital for academic success in a language that they are simultaneously trying to learn.

¹ Other terms commonly used to denote this group of students include English Language Learner (ELL), Limited English Proficient (LEP), and English as a Second Language (ESL)
It can take 6 months to 2 years for ELs to learn conversational English; however, it typically takes these students 5 to 7 years to achieve proficiency in the decontextualized, specialized English used in the classroom (Cummins, 1979; Thomas & Collier, 2002). Reading is one crucial way that students access academic content, with increasing demand in the upper elementary grades and beyond. Due to the dependence of reading on language processes, ELs often struggle with reading as compared to their English proficient peers (e.g., Geva, 2000; Lesaux, Crosson, Kieffer, & Pierce, 2010; Lesaux, Rupp, & Siegel, 2007; Nakamoto, Lindsey, & Manis, 2007; Polat, Zarecky-Hodge, Schreiber, 2017). Such difficulties are fairly widespread, as 65% of fourth grade ELs across the U.S. scored below the Basic level\(^2\) on the 2019 National Assessment for Education Progress reading assessment (NAEP; U.S. Department of Education [DoE], 2019).

A reading profile is the pattern of strengths and weakness in reading, language, and cognitive skills typically experienced by an individual reader or subgroup of readers. The general consensus is that, provided adequate literacy instruction, the reading profile of ELs is characterized by reading comprehension difficulties due to a lack of language proficiency in English, particularly an underdeveloped vocabulary, rather than due to deficits in word-level reading or reading-related cognitive skills (e.g., Lesaux, Rupp, & Siegel, 2007; Geva, 2000;)

\(^2\) The NAEP Basic level indicates that students demonstrate partial proficiency in grade-appropriate reading skills. At fourth grade, these skills include locating textual information, making simple inferences, identifying main ideas and relevant details, and determining the meaning of a word based on context.
Nakamoto, Lindsey, & Manis, 2007; Lesaux, Crosson, Kieffer, & Pierce, 2010; Harrison et al., 2015). Reading can be promoted among ELs through: multifaceted, scaffolded literacy instruction that builds students’ background knowledge, academic vocabulary, and oral language; capitalizing on first language strengths; and including explicit phonics and comprehension strategy instruction (August, McCardle, & Shanahan, 2014; Rivera, Moughamian, Lesaux, & Francis, 2009; Snyder, Witmer, & Schmitt, 2017). Nevertheless, such approaches may not be sufficient for ELs whose reading struggles are compounded due to an underlying reading disability (RD).

**Reading Disabilities**

Many children in the U.S. experience difficulties with reading. Among fourth grade students who were not ELs, 29 percent scored below the Basic level on NAEP reading assessment (DoE, 2019). A majority of children struggle with reading due primarily to what Vellutino and colleagues (1996) refer to as “experientially limiting factors” (p. 601), which include inadequate classroom literacy instruction, limited English proficiency, socioeconomic disadvantage (i.e., impoverished exposure to print and/or decontextualized language in the home), low reading interest or motivation, and sensory (i.e. hearing or visual) or intellectual deficits. While such children may exhibit reading difficulties, they do not have a reading disability (RD), which is defined as reading challenges due primarily to deficits in the cognitive processes underlying reading that are not explained by the aforementioned experiential factors.
In school settings, RDs are typically classified as a type of specific learning disability under the Individuals with Disabilities Education Improvement Act (IDEA), which is defined as “a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations” (IDEA, 2004, 20 U.S.C. §1401 [30]). School-based practitioners determine that a child has a specific learning disability if he or she fails to meet age or grade-level standards in one or more of the following areas when provided with adequate instruction: oral expression, listening comprehension, basic reading skills, reading comprehension, mathematics calculation, or mathematical reasoning, and if these learning challenges do not primarily stem from another disability (intellectual, visual, hearing, or motor), emotional disturbance, or environmental, cultural, or economic disadvantage. An estimated 80% of students with a specific learning disability primarily struggle with reading (Berdine, 2003). Students identified as having a specific learning disability are entitled to receive personalized special education services in school, which are laid out in an Individualized Education Program (IEP) in public schools or a service plan for students attending private schools. Such plans may include testing or classroom accommodations, curriculum modifications, a list of academic and functional goals, and a detailed plan to meet those goals including a list of services that will be provided. By contrast, students determined to have reading difficulties due to
experiential factors are not entitled to receive such supports, as their needs should be accommodated through other offerings (e.g., English language services in the case of ELs).

In clinical settings, such as hospitals or private practices, RDs are classified under the specific learning disorder diagnosis in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), which requires having 1) a difficulty for at least six months despite targeted help in word reading, reading comprehension, spelling, written expression, numerical cognition, or mathematical reasoning; 2) academic deficits compared to same-age peers that cause problems in school, work, and/or everyday activities; 3) difficulties that begin in childhood; and 4) that are not due to an intellectual disability, hearing or visual problems, a neurological condition, economic disadvantage, lack of instruction, or language background (APA, 2013). Children who are diagnosed with a specific learning disorder in a clinical setting are legally entitled to receive accommodations in public schools and private schools that obtain public funding under Section 504 of the Rehabilitation Act of 1973. These children may also qualify for special education services under IDEA, however, to do so they need to undergo evaluation by school-based practitioners to ensure that they meet qualifications for a specific learning disability.

Regardless of the label, there are two main subtypes of RD. The most well-known, an RD in word-level reading skills, is often referred to as developmental dyslexia. People with dyslexia have difficulties with decoding
and/or reading fluency, which are often attributed to a core phonological processing deficit (i.e., poor phonological awareness, phonological memory, phonological retrieval, and/or processing speed) (Ramus, 2003; Saksida et al., 2016; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Wright, Bowen, & Zecker, 2000). The other main RD subtype, a specific reading comprehension disorder, is characterized by severe reading comprehension difficulties, independent of word-level reading abilities, that are not better explained by another condition (Cain & Oakhill, 2006; Cutting et al., 2013; Landi & Ryherd, 2017). Associated characteristics include deficits in vocabulary; higher-order language skills, such as inferencing and comprehension monitoring; and/or executive functions (Cain & Oakhill, 2006; Landi & Ryherd, 2017; Locascio, Mahone, Eason, & Cutting, 2010).

**Reading Disabilities in English Learners**

Researchers have suggested that word-level RDs in ELs can be identified prior to the establishment of English language proficiency and can be reliably differentiated from reading difficulties due to underdeveloped English (Geva, 2000; Lesaux et al., 2007). They suggest that, similar to monolinguals, students who are substantially behind on measures of word-level reading skills (in English and in their native language, if available) and have much lower levels of reading as opposed to listening comprehension in English, may have an RD (Geva, 2000; Lesaux et al., 2007). Furthermore, they hold that if a child has an RD, it will impact all of the languages he or she is literate in, so if an RD can be reliably
identified in the child’s native language, it is likely that the child’s reading difficulties in English are due to an RD (Geva, 2000).

Nevertheless, longitudinal studies of reading development comparing ELs and their monolingual peers have typically studied ELs with optimal learning conditions: those with higher SES backgrounds, literate in their first language (e.g., attending bilingual schools), and/or instructed in English during their early elementary school years, when literacy instruction focuses on code-based skills (e.g., Lesaux and Siegel, 2003; Manis, Lindsey, & Bailey, 2004; Nakamoto, Lindsey, & Manis, 2007; Swanson, Orosco, & Lussier, 2011). The issue of identification becomes more complex in practice, given that ELs in the U.S. often receive inadequate language and literacy instruction in English and their native language, and therefore may not have developed adequate word-level reading skills due to a lack of educational opportunity (August and Hakuta, 1997; Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Snow, Burns, & Griffin, 1998). In cases where ELs are literate in their native language, testing in that language may be unavailable and determinations must be made solely off of the child’s performance on English assessments, which may not be valid or reliable for use with ELs (Luk & Christodoulou, 2016; McCardle et al., 2005). Furthermore, the procedures used to identify students with a reading-related specific learning disability in schools vary across states, districts, and schools, and are underspecified regarding their application to ELs (Hauerwas, Brown, & Scott, 2013; Reschly and Hosp, 2004; Youman & Mather, 2013). School-based
practitioners have reported that they struggle to properly identify and evaluate
disabilities in ELs (United States Government Accountability Office [GAO], 2019),
yet there is insufficient research regarding the roles of practitioners involved,
current materials and methods used, and challenges faced when evaluating ELs
for RDs in practice.

Additionally, particularly in the case of suspected RDs in reading
comprehension, it is implausible that the distinction between disability and
difficulty can be made based on behavior alone, given that ELs generally tend to
have a similar profile of reading comprehension difficulties despite adequate
word-level abilities (Abedi, 2005; Klinger, Artiles, & Barletta, 2006). Researchers
have called for behavioral and neuroimaging studies to tackle this challenge of
distinguishing reading difficulties due to second language (L2) learning from
those resulting from an underlying RD (McCardle, Mele-McCarthy, & Leos, 2005;
Pugh et al., 2005; Simos, Billingsley-Marshall, Sarkari, Patakaia, & Papanicolaou,
2005). These researchers hypothesize that RDs in ELs might share a common
neurobiological etiology as those in monolinguals, however there has been scant
neuroimaging research on L2 reading among proficient bilingual children
generally, let alone those who are not yet proficient in their L2.

**Project Overview**

The current thesis attempts to address these gaps in the literature to
advance our understanding of how to improve identification of RDs in ELs
through two studies. The first study utilizes a comprehensive online survey to
examine a) the roles and characteristics of practitioners with experience identifying RDs in ELs, b) the materials and procedures that they use in this process, and c) the reported challenges faced in this endeavor. The survey was developed in conjunction with Dr. Christodoulou and her lab who were interested in the processes used to identify RDs in the general school-age population rather than those specific to ELs. Participants with roles in identifying or diagnosing RDs in school-age children in the U.S. were invited to participate in an online survey through relevant electronic mailing lists, Facebook pages, and emails sent to administrators of school districts and reading clinics. Study 1 draws on the responses from the subset of participants with experience identifying RDs in ELs (n=598) on the questions I developed to address the aforementioned research aims.

Although there have been a number of studies on the differences in state guidelines and criteria for the school-based identification of RDs in ELs and how these may relate to challenges experienced by practitioners (GAO, 2019; Hauerwas, Brown, & Scott, 2013; Reschly and Hosp, 2004; Scott, Hauerwas, & Brown, 2014; Youman & Mather, 2013), no studies to date have surveyed practitioners across the U.S. to explore the range of training, experience, methods, materials, and reported challenges faced by those who directly conduct this process. An in-depth understanding of the ways in which ELs are currently identified in practice can be used to inform school leaders and policymakers, among others, as to the challenges that can be immediately addressed through
training, clearer guidelines, and/or additional resources. Moreover, such an investigation can provide a nuanced understanding of the more complex challenges that cannot be easily addressed by drawing on current research and resources alone, and therefore can provide insight into future research of high utility to practitioners.

In the second study of my thesis, I explore behavioral and neural profiles of reading processes in English among adolescents from diverse language and literacy backgrounds. Specifically, I compare behavioral and neural profiles of English reading processes and brain-behavior correlations of these reading measures among adolescents, ages 11 to 15, who are Spanish-speaking English Learners (EL), Spanish-English bilinguals (BL), typically-developing monolingual English speakers (ML), and monolingual English speakers with a diagnosed reading disability (RD). Participants were drawn from a larger fMRI study conducted by Dr. Luk on language and learning in middle school students (n=96). In this study, participants completed a battery of cognitive, language, and reading assessments in English (and Spanish, for the adolescents in the EL and BL groups); an eye tracking paradigm; a demographic questionnaire; and an MRI session, consisting of a variety of anatomical and functional neuroimaging scans. My analyses draw on the behavioral and demographic data (n=76) as well as data from an fMRI “fast localizer” word-reading task in English (n=46) among participants who met group eligibility and data quality criteria. Using behavioral and neural measures of reading processes, and their correlations, I explore
similarities and differences in the reading profiles of these groups, discuss the possible implications of this work, and provide suggestions for future cognitive neuroscience research to build on these findings.

Prominent scholars in the cognitive neuroscience community suggest that neuroimaging could potentially help distinguish difficulty from disability among ELs in cases where behavior alone is insufficient (McCardle, Mele-McCarthy, & Leos, 2005; Pugh et al., 2005; Simos, Billingsley-Marshall, Sarkari, Pataaraia, & Papanicolaou, 2005). Cognitive neuroscience research has led to advancements in the identification and remediation of RD among monolinguals (see Gabrieli, 2016; Eden et al., 2016). Even so, there is currently a dearth of neuroimaging research on second language (L2) reading in both bilinguals and ELs. Given that almost a quarter of school-age children in the U.S. speak a non-English language at home (U.S. Census Bureau, 2019), extension of this research to bilingual populations is necessary to address the reading difficulties and disabilities experienced by all children in U.S. classrooms. In Study 2, I initiate this process by comparing behavioral and neural measures of reading among adolescents from diverse language and literacy backgrounds. By directly comparing bilinguals and ELs with each other as well as with monolingual readers with and without reading disabilities, I establish a baseline understanding of the similarities and differences in reading profiles among these groups that can inform both current educational practice, as well as future research aimed at distinguishing reading difficulty from disability among ELs. Much of the previous behavioral research on
this topic has focused on ELs in elementary school, while this study focuses on
the behavioral and neural reading profile of adolescent Spanish-speaking ELs.
For logistical reasons, this study does not include ELs with diagnosed RDs, as a
fairly small number of children would have been eligible to participate in that
group due to the occasional practice of assigning children exclusively to either EL
or special education status. Therefore, this study cannot directly meet the aim of
distinguishing reading difficulties from disabilities among ELs using cognitive
neuroscience methods. Nevertheless, it provides the foundation necessary to
inform such research by establishing normative behavioral and neural reading
profiles among typically-developing adolescent Spanish-English bilinguals and
ELs.

Together, these two studies use novel approaches to advance the
research on the identification of RDs among ELs. In Study 1, I utilize a national
survey to uncover the practical challenges faced when tasked with this question
under real-life circumstances. Further, in Study 2, I initiate cognitive
neuroscience research comparing English reading processes among adolescent
ELs, bilinguals, and monolinguals with and without RD to build a foundation for
future research aiming to distinguish reading difficulty from disability through this
approach. These findings will provide novel data useful for informing both future
research and current educational practice and policy on this topic.
STUDY 1:

Investigating Current Practices Used to Identify Reading Disabilities in English Learners
Abstract

As the number of English Learners (ELs) in U.S. schools continues to rise, practitioners are increasingly tasked with the challenge of distinguishing reading disabilities (RDs) from difficulties with reading due to normative second language acquisition. Disproportional representation of ELs with such disabilities evident at the state-, district-, and grade-levels, however, suggests that they are not being identified in an accurate and timely manner. Prior research indicates that a lack of bilingual practitioners, non-English assessments, and sufficient training may explain the high prevalence of inaccurate eligibility decisions for ELs. A majority of these studies were conducted 10 to 20 years ago and/or focus on a subset of practitioners involved in this process, however. The present study seeks to explore current practices among all practitioners with roles in identifying RDs in ELs using an online survey of 598 school- and clinic-based practitioners across the U.S. This survey explores practitioners’ roles in the identification process, the specific assessments and practices that they typically use with ELs, the criteria that they use for decision-making, and their perceived greatest barriers to accurate identification of RDs in ELs. Findings suggest that a variety of professionals participate in this process, including school psychologists, special educators, speech language pathologists, reading specialists, and that these practitioners using a wide range of standardized assessments, informal measures, and evaluation criteria. These practitioners are largely monolingual or do not use their non-English language to assess ELs. Despite federal regulations
that stipulate that ELs should be assessed in all of their languages in virtually all cases using valid, reliable, and culturally-unbiased assessments, a majority of participants report that ELs in their settings are assessed in English only most of the time or always. Additionally, many of the standardized assessments used with ELs, particularly in the reading and writing domains, are only available in English and do not include ELs or bilinguals in their norming samples. Reported challenges to accurate and timely identification for ELs are largely consistent with previous research, and include a shortage of bilingual practitioners, dearth of appropriate assessments for ELs in English and non-English languages, difficulty involving families in the evaluation process, and insufficient training on best identification practices with ELs. Suggestions to improve this process through research, policy, and practice are discussed.
Introduction

English Learners (ELs) comprise an estimated 9.6 percent (4.9 million) of students in U.S. public schools, a figure that has risen from an estimated 8.1 percent (3.8 million) in 2000 (McFarland et al., 2019). Approximately 14 percent of these students receive special education services due to an identified disability, which is equivalent to the percentage of students in the general population who are identified with a disability. While identification rates for most disability categories (e.g., intellectual disability, autism) are also comparable for ELs and the general student population at the national level, the identification rate for specific learning disabilities (SLDs) in the U.S. is higher among ELs (50%) than in the general population (39%) (IDEA Data Center, 2015). This suggests that ELs are over-identified with SLDs at the national level.

When broken down by state averages, however, ELs were both over- and under-identified with SLDs, ranging from a low of 27% of ELs with disabilities identified as having an SLD in Kentucky to a high of 71% in Nevada. Given that SLDs are thought to have a neurological basis (MacMillan & Siperstein, 2002), their prevalence would be expected to be consistent across states. A number of studies have also found both over- and under-identification of ELs with SLD at the district-level (De Valenzuela, Copeland, Qi, & Park, 2006; Linn & Hemmer, 2011; Sullivan, 2011; Zehler et al., 2003). Furthermore, ELs are typically under-identified in the early elementary grades and over-identified in the middle school years and beyond (Artiles et al., 2005; Hibel & Jasper, 2012; Yamasaki & Luk,
Taken together, this disproportionality at the national-, district-, and grade-levels suggests that some ELs are receiving unnecessary special education services when they may just need additional instruction appropriate for ELs, while others are missing out on the personalized resources provided through special education services to support students with an SLD.

While the SLD category includes reading, writing, and/or math disabilities, an estimated 80% of students with this classification have a reading disability (RD). It can be especially challenging for practitioners to distinguish between reading difficulties due to a lack of English proficiency and a neurologically-based RD among ELs because reading processes are heavily dependent on language abilities (Abedi, 2006; Klingner, Artilés, & Barletta, 2006; McCardle, Mele-McCarthy, Cutting, Leos, & Emilio, 2005; Shore & Sabatini, 2009; Wagner, Francis, & Morris, 2005). It is important that ELs’ learning needs are properly identified. For children whose primary challenge is underexposure to English, reading interventions are less effective than appropriate English language instruction at improving academic performance, and vice versa for readers struggling due to RD (Callahan, 2006). Additionally, ELs receiving special education services are legally entitled to receive specialized English language instruction through programs such as English Language Development (ELD), English as a Second Language (ESL), or Structured English Emersion (SEI). However, in practice, they may lose access to such instruction due to logistical constraints (Zehler et al., 2003; Sánchez, Parker, Akbayin, & McTigue, 2010),
such as scheduling difficulties or budget limitations. In such cases, if students’ reading difficulties were actually attributable to underdeveloped English, they would lack the appropriate instruction necessary to address them. It is imperative to improve the process of identifying RDs among ELs to ensure that those with reading struggles receive the suitable support to succeed in school and beyond.

**General Diagnostic Guidelines and Procedures**

**Identifying RDs in School Settings.** Identification of RDs among school-age children in the U.S. can be complex due to differences in policies, guidelines, definitions, eligibility criteria, and diagnostic processes used across states and school districts (Maki, Floyd, & Roberson, 2015; Reschly & Hosp, 2004). In school settings, struggling readers are typically first referred by their teacher to a team (often referred to as the child study team or student support team) consisting of a general education teacher, special educator, school psychologist, school administrator, and/or other practitioners who meet to determine interim measures and monitor the child for follow up during later meetings (Ysseldyke, 2001; Klingner & Harry, 2006). The team may decide to refer the child for a formal evaluation at which point a variety of standardized assessments, curriculum-based assessments, achievement tests, classroom observations, parent and teacher interviews, and other measures may be conducted and/or retrieved for review during a placement meeting. This meeting, typically referred to as an IEP meeting, staffing, or multidisciplinary team meeting, is usually attended by a school psychologist, special education teacher, general education
teacher, school-level administrator, district-level administrator, parent, or others, and sometimes the student under review (Klingner & Harry, 2006). This process is more complex for ELs, often involving an evaluation to determine English proficiency level, if possible; testing in the child’s native language, if deemed warranted; and additional practitioners involved in the process, including a bilingual diagnostician and/or the child’s ESL teacher.

Federal regulations stipulate that eligibility for an SLD in reading is contingent on the child’s underachievement on reading relative to age, the state’s grade-level academic standards, and/or IQ when using a scientifically-backed process such as the child’s failure to respond to research-based intervention or a pattern of strengths and weaknesses (IDEA, 2004, 20 U.S.C. §1401 [30]). Additionally, practitioners must provide evidence that the child has received appropriate scientifically-based literacy instruction delivered by qualified practitioners and that the reading difficulty does not primarily stem from another disability (e.g., visual, hearing, motor, or intellectual), emotional disturbance, or environmental, cultural, or economic disadvantage.

While few studies have explored the range of current identification practices used across U.S. schools, those that have been conducted have found that the assessments, cut points on standard scores, pre-referral and/or progress monitoring approaches, magnitude of discrepancies between scores (when applicable), definition of adequate progress, use of professional judgments, among other factors, vary widely at the state-, district-, and practitioner-levels.
(Mellard, McKnight, & Woods, 2009; Scruggs & Mastropieri, 2002). The specific criteria used to make the determination are further specified by state- and district-level policies and guidelines; all U.S. states allow for the use of a Response to Intervention (RTI) approach, 67% of states allow for use of the IQ/Achievement discrepancy criteria, and 28% of states allow for the use of the Pattern of Strengths and Weaknesses (PSW) criteria (Maki et al., 2015).

**IQ/Achievement Discrepancy.** Prior to 2004, all states used the IQ/Achievement discrepancy criteria, which requires that a child’s academic performance falls significantly below their IQ score (e.g., by a standard deviation), however the validity of this model was called into question by a number of studies that found similar skill profiles and instructional response for poor readers with and without such a discrepancy (Fletcher et al., 1994; Stuebing et al., 2002; Vellutino, Scanlon, Lyon, 2000). The IQ/Achievement discrepancy was also found to be biased against ELs, as they typically score lower on IQ measures conducted in English due to increased cultural and linguistic demands on verbal portions of IQ tests in particular (Huang, Clark, Milczarski, & Raby, 2011; Geva & Wiener, 2015).

**Response to Intervention.** With the reauthorization of IDEA in 2004, the policy was changed such that states must not require the use of the IQ/Achievement discrepancy and should allow for other evidence-based approaches. Many states took up the RTI model, which involves offering high quality, evidence-based reading instruction to all students, screening all children
for reading difficulties, providing evidence-based small group interventions to struggling readers while monitoring their performance, and referring children for assessment if they fail to make adequate progress (Fletcher & Vaughn, 2009; Fuchs & Fuchs, 2006). This approach is more promising for ELs, but may result in misidentification or delay in identification for this population due to inappropriate or ineffective literacy instruction, an overreliance on standardized measures conducted in English, a lack of knowledge as to what constitutes adequate progress, and other factors (Orosco & Klingner, 2010; Thorius & Sullivan, 2013).

**Pattern of Strengths and Weaknesses (PSW).** Some states also allow for the use of a PSW approach to identification, which involves identifying a cognitive deficit (e.g., phonological awareness, rapid naming, working memory) that theoretically underlies the child’s academic difficulty (e.g., decoding, reading fluency). Common examples of this approach include the concordance/discordance method (Hale & Fiorello, 2004) and the dual discrepancy/consistency method (Flanagan, Ortíz, & Alfonso, 2013). The concordance/discordance method involves finding a cognitive processing weakness (e.g., processing speed) that is not significantly different from an academic weakness (e.g., reading fluency), while also finding a cognitive processing strength (e.g., working memory) that is significantly higher than both the cognitive processing and academic weaknesses. The dual discrepancy/consistency method requires a weakness in at least one academic
area, a low IQ score in at least one broad ability as defined by Cattel-Horn-Carroll theory, a theoretical relationship between the low IQ score and the area of academic weakness, and scores in the average range or above on all of IQ measures. Empirical evaluations, suggest that PSW methods are flawed because there is low agreement among the different PSW approaches; poor readers who meet criteria using PSW approaches share similar skill profiles as those who do not; the use of different standardized measures results in different qualification decisions across practitioners; and cognitive test results cannot be meaningfully used to inform reading remediation (Kranzler, Floyd, Benson, Zaboski, & Thibodaux, 2016; Miciak, Fletcher, Stuebing, Vaughn, & Tolar, 2014; Miciak, Taylor, Denton, & Fletcher, 2015).

**Diagnosing RDs in Clinical Settings.** Even less has been published on diagnostic practices in clinical settings, such as private clinical practices, hospitals, and university-based clinics. In such settings, clinical psychologists, school psychologists, pediatric neurologists, neuropsychologists, speech-language pathologists, and/or other licensed clinicians can diagnose an RD in children, although insurance companies selectively cover such sessions depending on practitioner and settings (Kamara, 2015). Therefore, it may be challenging for ELs to access such services, as approximately 60 percent of these students are estimated to be socioeconomically disadvantaged, with family incomes below 185 percent of the federal poverty line (Grantmakers for Education, 2013). Furthermore, RDs in clinical settings are classified under the
specific learning disorder diagnosis in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), which requires having 1) a difficulty for at least six months despite targeted help in word reading, reading comprehension, spelling, written expression, numerical cognition, or mathematical reasoning; 2) academic deficits compared to same-age peers that cause problems in school, work, and/or everyday activities; 3) difficulties that begin in childhood; and 4) that are not due an intellectual disability, hearing or visual problems, a neurological condition, economic disadvantage, lack of instruction, or language background (APA, 2013). The two types of RDs included in the specific learning disorder diagnosis are dyslexia, which is characterized by difficulties with word reading (decoding and/or fluency), and a reading comprehension impairment, which is characterized by difficulties understanding the meaning of text despite having adequate word reading skills.

Under Section 504 of the Rehabilitation Act of 1973, children who are diagnosed with a specific learning disorder in a clinical setting are legally entitled to receive accommodations in schools that receive public funding. To receive special education services, which include an individualized educational program of supports in addition to the kinds of accommodations provided under Section 504, children often need to undergo additional testing by school-based practitioners to ensure that they meet qualifications for a specific learning disability in accordance with IDEA. School-based evaluators may take the external evaluation results into consideration when making an eligibility
determination (Mather & Wendling, 2011). Therefore, a thorough understanding of the range of practices involved in identifying RDs among ELs should also include those implemented in clinical settings, as they may affect school-based qualification decisions.

**Considerations for Evaluating ELs**

**Policies and Guidelines.** While the general policies and guidelines for identification of SLDs across states, districts, and practitioners are highly variable, there is even less consistency and clarity in those specific to ELs (DeMatthews, Edwards, & Nelson, 2014; McCardle, Keller-Allen, & Shuy, 2008; Sánchez et al., 2010). For example, of the 20 states with the largest EL populations, only three (Illinois, Minnesota, and Virginia) have published detailed manuals on the process of identifying and remediating SLD in ELs while other states have provided little or no guidance on this topic (Burr et al., 2015).

Furthermore, DeMatthews and colleagues (2014) found a similar lack of guidance at the district-level in their qualitative multi-case study of SLD identification processes for ELs within districts on the U.S.-Mexico border.

**Practitioner training.** A lack of clear and consistent guidelines may be sufficient in cases where practitioners are adequately trained; however, a number of qualitative studies have found that both the classroom teachers who work with ELs and school-based practitioners involved in the SLD identification process for ELs report that they lack adequate training regarding cultural differences, language development, SLDs, and their intersection (Klingner & Harry, 2006;
Sánchez et al., 2010; Zehler et al., 2003). Poor quality classroom instruction and/or small group interventions for ELs due to inadequate teacher training can result in an insufficient opportunity for ELs to learn and therefore lead to inappropriate identification decisions (Orosco & Klingner, 2010). Researchers (García & Ortiz, 2006; Ortiz & Yates, 2001) suggest that the prevalence of inappropriate special education referrals for ELs can be reduced by training educators to engage in high-quality, culturally and linguistically responsive instruction and implement a wide variety of prereferral interventions for students struggling in the general education setting (e.g., Title I services, services for immigrant students, community-based programs).

Furthermore, IDEA (2004) requires that trained practitioners evaluate ELs for SLDs in their native language, in addition to English, when possible, using culturally, linguistically, and racially unbiased measures in a valid and reliable manner (§300.304). It appears that this has not typically been accomplished in practice, in the past, as in a survey of 859 school psychologists across the U.S. with prior experiences conducting bilingual assessments, Ochoa, Rivera, and Powell (1997) found that only 6% reported asking for their students’ home language, and only 1% investigated whether the child met eligibility criteria in both English and the child’s native language. This is problematic because research suggests that if a child truly has a reading disability, it will be present in

3 They define culturally responsive instruction as instruction that builds on students’ prior cultural and linguistic knowledge and experiences and distinguish this from an additive approach that involves simply adding culturally relevant examples to supplement existing curricula (García & Ortiz, 2006, p. 5).
both the child’s native language(s) and English (Olvera & Gomez-Cerillo, 2011). It may not be possible to assess literacy skills in the child’s native language in the case that he or she did not receive literacy instruction in that language, which is often the case for ELs who were born in the U.S. (Spinelli, 2008). Bilingual assessment may still prove useful for such students, as language and non-verbal IQ standardized assessments can be administered, and all assessment instructions can be relayed, in the child’s native language. Nevertheless, in a survey of nationally-representative schools across the U.S., Zehler and colleagues (2003) found a reported lack of staff members with expertise in both special education and working with ELs, and that only 14% of teachers who worked with ELs receiving special education were reportedly fluent in the non-English languages spoken by their students. Similarly, in her interviews of state education agency staff members across seven states, Keller-Allen (2006) found a lack of bilingual practitioners and those trained in multicultural and multilingual assessment.

**Assessment Practices.** An additional major challenge is the lack of valid and reliable academic and cognitive assessments available in non-English languages, and the limited availability of non-English measures in languages other than Spanish (Keller-Allen, 2006; Sánchez et al. 2010). Without trained bilingual assessors and appropriate non-English measures, schools may rely on translators to facilitate testing in the child’s native language using English measures; however, many of the standardized assessments typically used in
English are not culturally appropriate and/or do not include ELs in their norming sample (Abedi, 2006; Luk & Christodoulou, 2016; McCardle et al., 2005), and poor translations of such assessments further reduce their validity (Artiles & Ortiz, 2002; Popham, 2011; Wagner et al., 2005). One alternative is to assess the child solely in English after determining that the student has achieved adequate English language proficiency. However, the normed scores would still be invalid in this case and Zehler and colleagues (2003) found that this approach resulted in delayed identification, as ELs in many districts across the U.S. were required to be exited from second language services prior to evaluation for an SLD. School-based practitioners in two out of the three districts interviewed by Sánchez and colleagues (2010) held this belief that ELs could not be reliably identified prior to achieving English proficiency, which stands in direct contrast to the research on this topic (Geva, 2000; Lesaux et al., 2007).

Conversely, in their three-year ethnographic study of schools within a U.S. urban district, Klingner and Harry (2006) found that practitioners prematurely concluded that ELs with conversational English proficiency could be assessed in English only, therefore resulting in misidentification due to limited academic language proficiency as opposed to a genuine SLD.

Klingner and Harry (2006) also found that practitioners relied mainly on standardized test scores and students’ English proficiency levels when determining whether or not ELs qualify for an SLD. By contrast, researchers suggest that such decisions for ELs should be made based on multiple types of...
data, including informal measures, such as classroom observations (Klingner & Harry, 2006; Rinaldi & Samson, 2008; Spinelli, 2008); curriculum-based dynamic assessment (Barrera, 2006; Barrera & Liu, 2010; Huang et al., 2011; Jitenrda, Rohena-Diaz, & Nolet, 1998; Spinelli, 2008); performance-based assessment (Spinelli, 2008), interviews of teachers and parents (Rinaldi & Samson, 2008; Spinelli, 2008; Wilkinson et al., 2006); and comparisons to students with similar circumstances (Geva & Herbert, 2013; Wilkinson et al., 2006). Informal measures are especially useful in assessing ELs who are not literate in their native language because practitioners are unable to glean insight into their literacy abilities through native language assessment. For instance, dynamic assessment is a test-teach-retest procedure that allows practitioners to explore a child’s ability to learn new information and transfer that knowledge to novel items (Roseberry-McKibbin & O’Hanlon, 2005). The practitioner identifies an area of weakness on a curriculum-based measure administered in English, provides individualized instruction, and then reassesses the child’s understanding of the concept using novel items. If the practitioner needs to use more effort than usual to teach the child and/or the instructional strategies needed to teach the child deviate from those commonly used to teach the child’s peers, this suggests that the child may be struggling due to an RD rather than inadequate instruction. Given the lack of valid and reliable standardized assessments for ELs, particularly those who do not speak Spanish and/or are not literate in their native language, practitioners are encouraged to collect as much informal assessment data as possible to allow
for data triangulation in the special education decision-making process (Spinelli, 2008).

**Relationships with Family Members.** Poor relationships between ELs’ family members and school personnel have been identified as another key hindrance to accurate eligibility decisions (Klingner & Harry, 2006; Orosco & Klingner, 2010; Zehler et al., 2003). Whether due to ignorance, such as a lack of consideration of parents’ work schedules; logistical factors, such as difficulties obtaining qualified translators; or prejudicial beliefs, such as making negative comments in front of family members, these studies describe factors that prevented family members from fully engaging in the decision-making process for ELs. Family members are important in the identification process, as they can provide information regarding their child’s previous educational experiences; home language and literacy environment; language and learning development; socioemotional, behavioral, and/or physical health considerations; and available learning opportunities, among other factors (Spinelli, 2008). Kalyanpur, Harry, and Skrtic (2000) argue that family involvement in the special education process can be improved through the establishment of parent-professional collaborative relationships by which parents’ schedules, language and literacy abilities, economic circumstances, and cultural values are considered and respected by practitioners.
The Present Study

Taken together, these prior studies suggest that a lack of: clear and consistent guidelines; uniform identification methods; bilingual and/or knowledgeable practitioners; valid and reliable standardized assessments; informal assessment strategies in use; and family involvement have impeded the accurate and timely identification of ELs with SLDs in the past. Such information can be used to inform policy changes, practitioner training, and assessment development to ultimately improve this process. Even so, a majority of these studies were conducted 10 to 20 years ago, and it's unclear whether these challenges still hold, or whether great strides have been made to improve the identification process for ELs since then. Furthermore, no studies on this topic to date have explored the typical roles of various practitioners in the assessment process, the specific standardized assessments commonly used, or practitioners’ perceptions of the greatest barriers to accurate and timely identification, and all prior studies have been limited to school settings. The present study was designed to address these gaps.

I developed and disseminated an online survey to clinic- and school-based practitioners from across the U.S. to address the following research questions:
1. Who are the practitioners that identify or diagnose RDs in ELs? What are their roles in this process, proficiencies with non-English languages, and training experiences?

2. What are the current measures and processes used by practitioners in schools and clinical settings to identify RDs in ELs?

3. What are practitioners’ perceptions of a) the primary causes of reading difficulties facing ELs referred for assessment and b) the greatest challenges to accurate and timely identification of RDs in ELs?

This investigation provides an in-depth exploration of the current range of practices utilized in the U.S. to identify RDs in ELs that can be used to inform future efforts to improve this process across research, policy, and practice.

Methods

Survey

In collaboration with Dr. Joanna Christodoulou and her lab, I developed a survey for online dissemination to explore the range of practices used to identify RDs among school-age children in the U.S. The current paper focuses solely on the subset of questions specific to identifying RDs in ELs. Remaining survey items are discussed in another publication. The complete survey is included in Appendix A with questions relevant to this study indicated with bold font.

Qualtrics software (May 2018 version) was used to implement the survey in an

---

4 In school settings, children are “identified” in accordance with federal law (IDEA) as opposed to in clinical settings where they are “diagnosed” using the DSM-5. I will use these terms interchangeably throughout the paper to avoid redundancy, however “identification” is the proper term in school settings versus “diagnosis” in clinical settings.
easy-to-use, visually-appealing environment compatible for use across device types. The entire survey consisted of 56 questions, although participants were branched to fewer questions depending on their responses. No identifying information was collected from participants. The first section consisted of questions on the practitioners’ role in identifying RDs. This was followed by a section on participants’ demographic information including occupation, educational and work experiences, occupational setting, language backgrounds, and caseload composition. A set of questions on general assessment practices was then counterbalanced in presentation order with a set of questions on practices specific to identifying RDs in ELs. The section specific to ELs included questions adapted from Arias and Friberg’s (2017) survey on bilingual language assessment as well as questions developed to address my research aims. The final section asked participants to rate their confidence in evaluating children for RDs as well as their willingness to learn more about this process.

In addition to the specific response options provided, most questions also included an “other” category and text entry box so that practitioners could input additional information. This option was included to ensure that the survey would be inclusive of additional experiences of practitioners that were relevant and could capture participants’ perspectives as closely as possible within a survey framework. The survey was designed to take approximately 15 to 20 minutes to complete, which was fairly consistent with the median survey duration (19 minutes) for the analytic sample. Only the first few questions regarding the
participant’s role were required; otherwise, participants were able to skip questions and still advance through the survey. The survey content and procedure were piloted by three practitioners in advance, and revisions were made based on feedback to improve clarity and content relevance prior to survey distribution.

**Participant Recruitment**

Practitioners with roles in identifying or diagnosing RDs in school-age children in the U.S. were invited to participate in the online survey through multiple approaches. A recruitment email containing a link to the survey was sent to a variety of relevant organizations’ electronic mailing lists including SpellTalk and the Society for the Scientific Study of Reading. Study information was also posted to the American Speech-Language-Hearing Association (ASHA) Community Page and relayed to members of the Language Learning and Education, Cultural and Linguistic Diversity, and School-Based Issues ASHA special interest groups (SIGs). The survey link was also posted to various Facebook pages along with the brief recruitment message, “Have you identified reading disabilities in school-age children? Volunteer 15 minutes to complete this online survey of current practices conducted by researchers at Harvard and MGH Institute.” These Facebook pages included our lab websites, the Massachusetts Branch of the International Dyslexia Association (MABIDA), Decoding Dyslexia-MA, Decoding Dyslexia-TX, Decoding Dyslexia-CA, Wilson Language Training, the Stern Center for Language and Learning, as well as groups for professionals
in relevant fields: Reading Specialists & Literacy Coaches, SLPs for Evidence Based Practice, Bilingual School Psychologists, Bilingual Speech Language Pathologists, Bilingual and Special Education Teachers Network, Get Psyched About School Psychology, Private Practice SLPs, and Clinical Research for SLPs. In addition, personalized recruitment emails were sent to public school district administrators across Massachusetts containing both the link to this study as well as information regarding a qualitative interview study that I was simultaneously conducting on the practices used to identify RDs in ELs attending Massachusetts public schools. A generic recruitment email was also sent to administrators with publicly accessible email addresses from across all public-school districts in the United States. Recruitment emails were also sent to reading clinics across the U.S. with publicly accessible email addresses (e.g., Children’s Dyslexia Centers, university clinics, those listed on the IDA website). Lastly, Dr. Christodoulou and I reached out to members of our professional networks for help disseminating the recruitment email to the practitioners in their networks. Upon clicking on the study link, practitioners were routed to an online consent page where they could read about the study and decide whether or not to participate by clicking “Agree” or “Disagree” (see Appendix A). This procedure was approved by both the Partners Institutional Review Board and Harvard University Committee on the Use of Human Participants (CUHS).
Sample

A total of 1,360 survey responses were collected over a nine-month data collection period. Respondents who indicated that they hold no role in the diagnostic process (22.9%, \( n=312 \)) were immediately routed to the completion page, and no subsequent data were collected. Blank responses, responses with less than 25% of the survey completed, and participants that did not meet the study criteria (e.g., parents, practitioners outside of the U.S., those not currently working in the field) were excluded from analyses. This resulted in a full sample of 965 practitioners, 62.0% of whom reported experience identifying or diagnosing RDs in ELs. Given that this paper focuses on diagnostic practices specific to ELs, all analyses presented below are restricted to the analytic sample of practitioners (\( n=598 \)) who indicated prior experience identifying RDs in ELs.

The demographic characteristics of the sample are presented in Table 1.1 and indicate that participants have a range of professions, experiences, and settings. A variety of occupations were represented; however, the majority of respondents were in the target demographic, including school psychologists (24.9%), special educators (19.2%), reading specialists (12.2%), and speech-language pathologists (11.0%). Most participants (92.3%) reported that they had earned a Master’s degree or higher and had 10 or more years of experience in their profession (62.5%). In addition, 88.9% of participants indicated that they work primarily in schools, while 11.1% reported that they work mainly in clinical settings, such as a hospital or private practice. Approximately half of participants
reported that they work with students in suburban settings (50.9%, \( n=298 \)) as compared to 25.4% (\( n=149 \)) in urban settings and 23.7% (\( n=139 \)) in rural settings. Practitioners from all U.S. states except for North Dakota, West Virginia, and Wyoming were represented in the analytic sample, with 41.5% located in the Northeast, 26.4% in the South, 14.2% in the Midwest, 17.6% in the West, and 0.03% from other U.S. territories (\( n=2 \)). Some states were only represented by a single participant (Vermont, Delaware), but the median number per state was 8 participants. California (5.4%, \( n=32 \)), Texas (9.4%, \( n=56 \)), and Massachusetts (26.3%, \( n=157 \)) were the states represented by the largest number of participants. Given that California and Texas are the two states with the highest percentage of ELs (20.2% and 17.2%, respectively, McFarland et al., 2019), it is not surprising that these states were among those with the highest number of participants in the analytic sample. By contrast, nine percent of students in Massachusetts are ELs, however, the large number of participants from this state likely resulted from the more personalized email approach taken to reach these practitioners, the fact that my personal and professional network is largely based in Massachusetts, and the additional word-of-mouth from those who participated in the interview study. A majority of practitioners (88.2%, \( n=493 \)) indicated that they work with elementary school students, while 62.6 percent reported working with middle school students (\( n=350 \)), and 38.6 percent reported working with high school students (\( n=216 \)). Few participants reported working exclusively with middle and/or high school students (11.8%, \( n=66 \)), and the survey did not ask
practitioners how their identification methods differ, if at all, depending on the age of the ELs evaluated. Additional research is necessary to explore the role of age, as well as other student-level characteristics (e.g., English proficiency level, acculturation, instructional history), in the RD identification process for ELs.
Table 1.1.

Demographic characteristics of the analytic sample

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>% (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occupation (n=598)</strong></td>
<td></td>
</tr>
<tr>
<td>School Psychologist</td>
<td>24.9% (n=149)</td>
</tr>
<tr>
<td>Special Educator</td>
<td>19.2% (n=115)</td>
</tr>
<tr>
<td>Reading Specialist</td>
<td>12.2% (n=73)</td>
</tr>
<tr>
<td>Speech-Language Pathologist</td>
<td>11.0% (n=66)</td>
</tr>
<tr>
<td>More than one</td>
<td>12.0% (n=72)</td>
</tr>
<tr>
<td>Other</td>
<td>20.6% (n=123)</td>
</tr>
<tr>
<td><strong>Highest degree (n=595)</strong></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>7.7% (n=46)</td>
</tr>
<tr>
<td>Master’s</td>
<td>64.9% (n=386)</td>
</tr>
<tr>
<td>Specialist</td>
<td>9.9% (n=59)</td>
</tr>
<tr>
<td>Certificate of Advanced Study</td>
<td>3.0% (n=18)</td>
</tr>
<tr>
<td>Doctorate</td>
<td>11.8% (n=70)</td>
</tr>
<tr>
<td>Medical Doctor</td>
<td>1.5% (n=9)</td>
</tr>
<tr>
<td>Other</td>
<td>1.2% (n=7)</td>
</tr>
<tr>
<td><strong>Years of Experience in Occupation (n=595)</strong></td>
<td></td>
</tr>
<tr>
<td>Less than 1 year</td>
<td>1.5% (n=9)</td>
</tr>
<tr>
<td>1 to 3 years</td>
<td>10.4% (n=62)</td>
</tr>
<tr>
<td>4 to 6 years</td>
<td>12.6% (n=75)</td>
</tr>
<tr>
<td>7 to 9 years</td>
<td>12.9% (n=77)</td>
</tr>
<tr>
<td>10 to 12 years</td>
<td>12.8% (n=76)</td>
</tr>
<tr>
<td>13 to 15 years</td>
<td>10.4% (n=62)</td>
</tr>
<tr>
<td>16 years or more</td>
<td>39.3% (n=234)</td>
</tr>
<tr>
<td><strong>Primary Work Setting (n=596)</strong></td>
<td></td>
</tr>
<tr>
<td>Public school</td>
<td>81.0% (n=483)</td>
</tr>
<tr>
<td>Charter school</td>
<td>4.7% (n=28)</td>
</tr>
<tr>
<td>Private school</td>
<td>2.2% (n=13)</td>
</tr>
<tr>
<td>Private special education school</td>
<td>1.0% (n=6)</td>
</tr>
<tr>
<td>Hospital</td>
<td>1.7% (n=10)</td>
</tr>
<tr>
<td>Private Practice</td>
<td>9.4% (n=56)</td>
</tr>
<tr>
<td><strong>Urbanicity (n=586)</strong></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>25.4% (n=149)</td>
</tr>
<tr>
<td>Suburban</td>
<td>50.9% (n=298)</td>
</tr>
<tr>
<td>Urban</td>
<td>23.7% (n=139)</td>
</tr>
<tr>
<td><strong>Geographic Region (n=598)</strong></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>41.5% (n=248)</td>
</tr>
<tr>
<td>South</td>
<td>26.4% (n=158)</td>
</tr>
<tr>
<td>Midwest</td>
<td>14.2% (n=85)</td>
</tr>
<tr>
<td>West</td>
<td>17.6% (n=105)</td>
</tr>
<tr>
<td>Other U.S. territories</td>
<td>0.03% (n=2)</td>
</tr>
<tr>
<td><strong>Student population (n=559)</strong></td>
<td></td>
</tr>
<tr>
<td>Elementary school only (K-5)</td>
<td>33.3% (n=186)</td>
</tr>
<tr>
<td>Elementary, middle, &amp; high school (K-12)</td>
<td>31.5% (n=176)</td>
</tr>
<tr>
<td>Elementary &amp; middle school only (K-8)</td>
<td>22.5% (n=125)</td>
</tr>
<tr>
<td>Middle school only (6-8)</td>
<td>5.7% (n=32)</td>
</tr>
<tr>
<td>High school only (9-12)</td>
<td>3.0% (n=17)</td>
</tr>
<tr>
<td>Middle and high school only (6-12)</td>
<td>3.0% (n=17)</td>
</tr>
<tr>
<td>Elementary and high school only (K-5, 9-12)</td>
<td>1.1% (n=6)</td>
</tr>
</tbody>
</table>

Note: Participants were assigned to the “More than one” occupation category if they currently work in more than one capacity in addressing reading difficulties. Participants in the “Other” occupation category included school administrators, educational specialists, dyslexia therapists, ESL teachers, general education teachers, educational diagnosticians, neuropsychologists, clinical psychologists, and literacy coaches.
Analysis

All analyses were conducted in Stata 15 (StataCorp, 2017). I tabulated the responses to each survey question using the total number of responses for that question as the denominator. I coded all open-text responses by theme and generated additional response categories using the codes. When appropriate, I also performed chi square tests of independence to explore whether the proportion of responses to each question differs by occupation, workplace urbanicity, primary workplace setting (clinic vs. school), or participant bilingual status. I only reported the results of such tests when they were statistically significant.

The total number of responses differed per question due to branching participants to relevant questions and the ability of participants to skip questions. For instance, although the full sample consisted of 598 practitioners, only 469 of them reported directly conducting some \( n=391 \) or all \( n=78 \) of the assessments used in the evaluation process, and therefore those with no direct role in assessment \( n=128 \) were not directed to questions regarding specific assessment practices. Questions related to the first research question about participants’ characteristics and roles were always presented first, so these had the highest completion rate, which ranged from 88.0 to 100 percent. Questions related to the second research question regarding the specific measures, assessment practices, and eligibility criteria that practitioners use came later during the survey, and therefore had a lower completion rate, which ranged from
67.4 to 82.9 percent of the participants who were presented with them.

Participants who skipped at least one question related to assessments \((n=259)\) were more likely to report working primarily in private practice \(\left(\chi^2(1)=97.90, p<0.001\right)\) and to be speech-language pathologists \(\left(\chi^2(5)=19.21, p=0.002\right)\) than those who completed this section \((n=339)\). The number of participants who work primarily in clinical practices and responded to each question regarding assessment practices ranged from 0 to 55, so additional research is necessary to generalize findings regarding assessment practices used with ELs to practitioners in clinical settings. Questions related to the final research question about participants’ perceptions and confidence regarding the identification process had a completion rate that ranged from 74.3 to 81.8 percent of participants, as these were always presented last in the survey. Participants who skipped at least one question in this last section \((n=184)\) were more likely to have a Master’s degree as opposed to a Doctoral degree \(\left(\chi^2(7)=16.75, p<0.019\right)\), more likely to work primarily in suburban settings \(\left(\chi^2(2)=6.40, p<0.041\right)\), and less likely to be from the Western region of the U.S. \(\left(\chi^2(3)=8.42, p<0.038\right)\) than those who completed all questions in this section \((n=414)\).

Given the non-random pattern of skipped questions, oversampling of participants from Massachusetts, and underrepresentation of participants who work primarily in clinical settings, these results cannot be expected to generalize to all practitioners across school and clinical settings in the U.S. with experience identifying RDs in ELs. The findings do provide a good initial indication of the
kinds of processes used in this endeavor by practitioners in public schools in Massachusetts, and to a lesser extent, across the U.S. Additionally, survey respondents who only reported having identified RDs in students who are not ELs \((n=367)\) were less likely to be school psychologists \((\chi^2(3)=16.55, p=0.001)\), more likely to work primarily in clinical settings \((\chi^2(1)=41.84, p<0.001)\), and less likely to be from the Midwest \((\chi^2(1)=41.84, p<0.001)\) than those in the analytic sample \((n=598)\). Therefore, additional research is necessary to investigate whether the findings hold in a more representative sample and explore whether such differences are due to non-random sampling or whether they are truly reflective of the characteristics of practitioners who identify RDs with and without experience working with ELs.

**Results**

In this study, I utilized an open online survey of practitioners with reported experience identifying RDs in ELs to explore 1) the roles and characteristics of practitioners involved in this process; 2) the measures and practices used; and 3) their perceptions of a) the primary causes of reading challenges among ELs referred for evaluation and b) the challenges they face in the identification process. Below, I present the results to each question, accordingly.

**Question 1: Who Identifies RDs in ELs?**

**Independent vs. Team Approach.** A majority of participants \((82.9\%, n=496)\) reported working with a team of practitioners to identify RDs in ELs,
however 17.1% (n=102) responded that they do so independently. This proportion differed by participants’ primary work setting ($\chi^2(1)=56.58, p<0.001$), such that only 13.0% of participants in school settings reported identifying RDs in ELs independently (n=69), versus 50.0% of practitioners in clinical settings (n=33). Furthermore, it differed by occupation ($\chi^2(1)=26.04, p<0.001$), such that special educators were less likely to report that they independently identify RDs in ELs than those in other professions. Nevertheless, among all participants who reported identifying RDs in ELs independently, only 45.1% (n=46) reported conducting all assessments used to make the determination. This suggests that the others who reported identifying RDs in ELs independently (n=56) make their determination alone despite only conducting some (51.0%, n=52) or none (3.9%, n=4) of the assessments. Furthermore, this indicates that only 7.7% of the analytic sample (n=46) reported conducting all assessments and independently making the determination for ELs. Perhaps the percentage of practitioners responsible for the entire testing and identification process for ELs would be higher if more practitioners from clinical settings participated, however this finding suggests that the identification process, particularly in school settings, often involves multiple practitioners.

Participants who reported working on a team to identify RDs in ELs were asked which practitioners in their setting typically comprise that team. Among participants who responded to this question (n=430), the team members most frequently reported were general educators (78.8%, n=339), school psychologists
(77.0%, n=331), special educators (76.5%, n=329), speech-language pathologists (71.2%, n=306), ESL teachers (68.1%, n=293), and reading specialists (56.7%, n=244). Other reported team members included educational specialists (24.0%, n=103), literacy coaches (17.2%, n=74), clinical psychologists (10.0%, n=43), school administrators (6.5%, n=28), social workers or guidance counselors (4.0%, n=17), neuropsychologists (2.6%, n=11), pediatricians (2.3%, n=10), occupational therapists (1.4%, n=6), and neurologists (1.2%, n=5). This suggests that a variety of occupational roles are often represented on teams tasked with identifying RDs in ELs, including general educators and ESL teachers.

**Practitioners’ Roles.** To delve deeper into the roles of various practitioners involved in the identification process, participants were asked which team members or stand-alone practitioners conduct assessments to identify RDs in ELs in their setting. Of the participants who responded to this question (n=516), 71.5% selected school psychologists (n=369), 56.2% selected speech-language pathologists (n=290), 51.4% selected special educators (n=265), 46.1% selected ESL teachers (n=238), 36.4% selected reading specialists (n=188), and 21.5% selected general educators (n=111). Other selected professionals included educational specialists (14.7%, n=76), clinical psychologists (11.8%, n=61), literacy coaches (5.2%, n=27), neuropsychologists (5.4%, n=28), medical professionals (1.6%, n=8), and school administrators (1.4%, n=7). Additionally, 37 participants (7.2%) wrote in that a bilingual
professional assesses EL students for RDs in their setting using terms such as “Bilingual Evaluator,” or “Bilingual team member.”

Although participants were not explicitly asked which kinds of assessments each of these professionals conduct in their setting, those who reported only conducting some of the assessments \((n=391)\) were asked which categories of assessments (e.g., cognitive, reading, language) they personally conduct when assessing ELs for a suspected RD, which may provide insight into the types of assessments typically carried out by each professional. Table 1.2 presents the percentage of these participants who reported conducting each category of assessments in English and in the child’s native language by occupational role. A majority of the school psychologists who responded to this question (90.5\%, \(n=95\)) indicated that they are responsible for conducting cognitive assessments in English and 22.9\% (\(n=24\)) reported that they also do so in the child’s native language. Similarly, all but one of the speech-language pathologists in the analytic sample reported conducting language assessments in English (98.0\%, \(n=48\)); and 40.8\% of these practitioners reported assessing proficiency in the child’s native language as well (\(n=20\)). All of the reading specialists (\(n=48\)), 96.4\% of the special educators (\(n=80\)), 90.0\% of practitioners with more than one occupational role (\(n=45\)), 64.8\% of school psychologists (\(n=68\)), and 38.8\% of speech-language pathologists (\(n=19\)) indicated typically conducting reading assessments in English with ELs. Likewise, 83.1\% of the special educators (\(n=69\)), 74.0\% of those with more than one occupational role
(n=37), 54.3% of the school psychologists (n=57), 41.7% of the reading specialists (n=20) and 36.7% of the speech-language pathologists (n=18) reported conducting writing assessments in English with ELs. Overall, relatively few participants indicated that they assess children in their native language, with 6.9% of respondents (n=27) reporting that they do so in the cognitive domain, 10.7% (n=42) reporting that they do so when assessing language, and 9.7% (n=38) and 7.9% (n=31) reporting that they do so in the reading and writing domains, respectively.

Table 1.2.
Percentage of participants who conduct each category of assessments for RDs in English (Eng) and the child’s native language (NL) with ELs by occupational role for those who reported only conducting a subset of the assessments (n=391)

<table>
<thead>
<tr>
<th>Occupational Role</th>
<th>Cognitive</th>
<th>Language</th>
<th>Reading</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Psychologists (n=105)</td>
<td>Eng: 90.5% NL: 22.9%</td>
<td>Eng: 16.2% NL: 9.5%</td>
<td>Eng: 64.8% NL: 12.4%</td>
<td>Eng: 54.3% NL: 10.5%</td>
</tr>
<tr>
<td>Special Educators (n=83)</td>
<td>Eng: 6.0% NL: 0.0%</td>
<td>Eng: 21.7% NL: 1.2%</td>
<td>Eng: 96.4% NL: 6.0%</td>
<td>Eng: 83.1% NL: 3.7%</td>
</tr>
<tr>
<td>More than one occupational role (n=50)</td>
<td>Eng: 34.0% NL: 4.0%</td>
<td>Eng: 48.0% NL: 14.0%</td>
<td>Eng: 90.0% NL: 16.0%</td>
<td>Eng: 74.0% NL: 16.0%</td>
</tr>
<tr>
<td>Speech-language pathologists (n=49)</td>
<td>Eng: 6.1% NL: 0.0%</td>
<td>Eng: 100.0% NL: 40.8%</td>
<td>Eng: 38.8% NL: 8.2%</td>
<td>Eng: 36.7% NL: 10.2%</td>
</tr>
<tr>
<td>Reading Specialist (n=48)</td>
<td>Eng: 6.3% NL: 0.0%</td>
<td>Eng: 20.8% NL: 0.0%</td>
<td>Eng: 100.0% NL: 2.1%</td>
<td>Eng: 41.7% NL: 0.0%</td>
</tr>
<tr>
<td>Other Occupation (n=56)</td>
<td>Eng: 28.6% NL: 1.8%</td>
<td>Eng: 50.0% NL: 7.1%</td>
<td>Eng: 85.7% NL: 12.5%</td>
<td>Eng: 71.4% NL: 7.1%</td>
</tr>
</tbody>
</table>

Languages Spoken by Practitioners and ELs. Only 27.7% of the analytic sample (n=151) reported speaking another language in addition to English, and this percentage did not differ significantly by primary work setting.
\( \chi^2(1)=1.07, \ p=0.301 \) or participant occupation \( \chi^2(5)=2.98, \ p=0.703 \). It did differ by the urbanicity of participants’ primary work setting, however, with 41.3% of practitioners in urban settings \( (n=52) \) reporting that they speak another language versus 24.7% in suburban \( (n=66) \) and 22.0% in rural \( (n=31) \) settings. Participants who reported working primarily in urban settings also reported a higher estimated percentage of ELs with RDs on their caseloads (30.1%), as opposed to those in suburban and rural settings (16.5% and 16.1%, respectively). This suggests that urban settings have a higher percentage of both bilingual practitioners in this role, as well as ELs with RDs on practitioners’ caseloads, as compared to suburban and rural settings. This finding confirms the national estimates of a higher proportion of ELs in urban districts (14.0%) as compared to suburban (9.3%) or rural (3.8%) districts (McFarland et al., 2019).

As illustrated in the bar graph of languages reportedly spoken by practitioners and their students Figure 1.1, substantially fewer practitioners reported speaking each language than those who reported evaluating EL speakers of those languages. This held for all languages except for French, German, and Italian, which were reportedly spoken by 5.5% \( (n=30) \), 2.8% \( (n=15) \), and 1.7% \( (n=9) \) of participants, yet spoken by 5.7% \( (n=30) \), 0.9% \( (n=5) \), and 1.3% \( (n=7) \) of their EL students, respectively. Spanish was overwhelmingly the most commonly reported non-English language spoken by practitioners with 19.8% of the total analytic sample \( (n=108) \) reporting that they speak Spanish. By contrast, 82.6% of participants \( (n=450) \) reported evaluating Spanish-speaking
ELs for a suspected RD. A substantial number of participants also reported evaluating ELs who speak the following languages: Arabic (16.9%, n=89), Portuguese (15.4%, n=81), Haitian Creole (10.3%, n=54), Mandarin (9.1%, n=48), Vietnamese (9.1%, n=48), Russian (7.4%, n=39), Hmong (6.5%, n=34), Tagalog (6.5%, n=34), Cantonese (5.1%, n=27), Cambodian (4.2%, n=22), Korean (4.0%, n=21), Cape Verdean Creole (3.4%, n=18), and Polish (3.2%, n=17). Fewer than ten participants reported evaluating speakers of Japanese (1.5%, n=8), Hindi (1.3%, n=7), Somali (1.1%, n=6), Urdu (1.1%, n=6) and 55 other languages. These results were generally consistent with national estimates of the most commonly reported home languages of ELs in U.S. public schools obtained from the Fall 2016 State Nonfiscal Survey of Public Elementary and Secondary Education (McFarland et al., 2019), in which the top languages included Spanish (76.6% of ELs), Arabic (2.6%), Chinese (2.1%), Vietnamese (1.6%), Somali (0.8%), Russian (0.7%), Hmong (0.7%), Haitian Creole (0.6%), and Portuguese (0.6%). Conversely, 2.9% of participants (n=16) reported speaking a language not commonly spoken by the ELs, such as American Sign Language (ASL), Hebrew, Norwegian, and Thai. These results suggest that many practitioners who evaluate ELs are either monolingual or do not speak the same language as the students that they are evaluating.
When asked whether participants have used their non-English language(s) to assess an EL for a suspected RD, only 41.7% of bilingual participants (n=63), or 12% of the total analytic sample, responded affirmatively. This proportion differed significantly by occupation ($\chi^2(5)=13.22$, $p=0.021$): only 20.0% of bilingual special educators (n=5) and 38.5% of bilingual reading specialists (n=5) reported using their non-English language to assess ELs for

---

5 The term “bilingual” is being used loosely here to represent the group of practitioners who report speaking one or more languages in addition to English.
RDs, while over half of bilingual school psychologists (61.1%, \(n=22\)), speech-language pathologists (75.0%, \(n=9\)), participants with more than one occupation (55.0%, \(n=11\)), and other professionals (50.0%, \(n=9\)) reported doing so. Taken together, these results further underscore the disconnect between the language profiles of the ELs and their evaluators, who are largely monolingual or speak additional languages that are not shared by their students.

**Training.** Participants who attended graduate school were asked how much training they received during their graduate studies on identifying RDs in ELs. Overall there was wide variation in training, with 28.3% \(n=149\) reporting receiving no graduate training on this topic, 14.8% \(n=78\) reporting a single lecture, 25.9% \(n=136\) reporting multiple lectures, 18.4% \(n=97\) reporting a single course, and 12.6% \(n=66\) reporting multiple courses. Participants who work primarily in schools reported significantly more graduate training on this topic than those who work primarily in clinical settings \(\chi^2(4)=9.81, p=0.044\) with 33.9% of practitioners in clinical settings \(n=21\) reporting no training on the topic as opposed to 27.4% of school-based practitioners \(n=127\). This finding cannot be explained by differences the amount of graduate training participants received on identifying RDs in the general student population \(\chi^2(4)=1.80, p=0.772\).

Additionally, school psychologists less frequently reported receiving no graduate training on the topic of identifying RDs in ELs than those in other professions \(\chi^2(20)=41.54, p=0.003\), but this may be because they more frequently reported
taking multiple courses on identifying RDs among the general student population \( (\chi^2(20)=42.99, p=0.002) \).

Participants were also asked how many professional development (PD) trainings they have participated in on this topic. Wide variation was observed in the reported training that participants have received on identifying RDs in ELs, with 14.9% \((n=84)\) reporting no PD sessions on this topic, 13.8% \((n=78)\) reporting a single PD session, 20.4% \((n=115)\) reporting multiple PD sessions, 14.2% \((n=80)\) reporting a single workshop or conference, and 36.7% \((n=207)\) reporting multiple workshops or conferences. Significant differences by occupation were observed \( (\chi^2(20)=38.69, p=0.007) \), with both school psychologists and professionals with more than one occupation more frequently reporting that they have attended multiple workshops and less frequently reporting that they have received no PD training on this topic as compared to other professionals. School psychologists and professionals with more than one occupation also more frequently reported attending multiple PD workshops on the topic of identifying RDs among the general student population as compared to the other participants \( (\chi^2(20)=40.46, p=0.004) \). These results suggest that school psychologists receive more training in graduate school and PD on identifying RDs in ELs than those in other professions. Furthermore, roughly eight percent of respondents \((n=42)\) indicated that they received no training on this process in either graduate school or professional development. These practitioners reported various
occupational roles, however the majority of them (66.7%, \( n=28 \)) reported directly conducting assessments with ELs.

**Question 2: Current Measures and Practices Used to Identify RDs in ELs**

**Measures.** Participants with experience personally conducting assessments for RDs with ELs (\( n=367 \)) indicated the measures\(^6\) that they typically administer when working with this population. Table 1.3 presents the cognitive, language, reading, and writing assessments selected by at least ten participants in response to this question, as well as whether each assessment can accommodate ELs (e.g., through inclusion of bilinguals in the norming sample and/or a bilingual or Spanish version). The *Wechsler Intelligence Scale for Children* (WISC; Wechsler, 2014) was the most commonly selected cognitive measure, indicated by 63.2% of participants (\( n=120 \)) who reported conducting a cognitive assessment. The *Test of Nonverbal Intelligence* (TONI; Brown et al., 2010; 38.4\%, \( n=73 \)) and *Kaufman Assessment Battery for Children* (KABC; Kaufman & Kaufman, 2004a; 28.4\%, \( n=54 \)) were also commonly selected by participants. The English version of the WISC does not include bilinguals in the norming sample, however there is a Spanish version that is appropriate for use with ELs. The TONI does include bilinguals in the norming sample and has test directions in a variety of languages for use with ELs. The KABC provides correct

---

\(^6\) Participants were asked to select the measure name regardless of edition, form, or version and were also able to write in any measures not listed.

\(^7\) All percentages reported in this section are calculated as the number of participants who selected the measure out of the number of participants who selected or wrote in a measure in that assessment category (e.g., cognitive, language, reading). The percentages are calculated within assessment category rather than across all participants because many of the participants indicated only administering a subset of assessments based on their occupational role.
responses and teaching text to accommodate Spanish-speaking ELs, although bilinguals are not included in the norming sample. Information on assessment versions or specific practices used with each assessment were not collected, however, so it is unclear whether these practitioners take advantage of these accommodations/different test versions provided for ELs when utilizing these measures. The other commonly selected cognitive measures listed do not provide specific accommodations for ELs, however the *Universal Nonverbal Intelligence Test* (UNIT; Bracken & McCallum, 2016; n=11) does include bilingual students in the norming sample.

The *Clinical Evaluation of Language Fundamentals* (CELF; Wiig et al., 2013) was the most commonly selected language assessment for use with ELs (54.7%, n=94). Although bilinguals are included in the norming sample, Paradis (2005) found that ELs tested using the English version may be incorrectly identified as having a language disorder. There is a Spanish version, however, to accommodate Spanish-speaking ELs. The next most frequently selected measures were vocabulary assessments, including the *Peabody Picture Vocabulary Test* (PPVT; Dunn & Dunn, 2007; 39.5%, n=68), *Expressive One-Word Picture Vocabulary Test* (EOWPVT; Martin & Brownell, 2011, 32.6%, n=56), and the *Receptive One-Word Picture Vocabulary Test* (ROWPVT; Martin & Brownell, 2010, 32.0%, n=55), followed by the *Test of Auditory Processing Skills* (TAPS; Martin & Brownell, 2005, 27.9%, n=48). All of these measures have separate versions appropriate for use with Spanish-speaking ELs. The remaining
language measures selected do not include accommodations for ELs except for the *Bilingual Verbal Ability Test* (BVAT, Muños-Sandoval et al., 1998, 18.6%, \(n=32\)), which measures overall verbal ability as well as English proficiency by assessing in English as well as the child’s native language with 16 commonly-spoken languages represented.

The most common reading and academic achievement assessments selected for use with ELs were the *Comprehensive Test of Phonological Processing* (CTOPP; Wagner et al., 1999; 52.0%, \(n=191\)), *Wechsler Individual Achievement Test* (WIAT; Wechsler, 1992; 39.0%, \(n=143\)), *Woodcock-Johnson Achievement Test* (WJ ACH; Schrank et al., 2014; 38.7%, \(n=142\)), and *Gray Oral Reading Tests* (GORT; Wiederholt & Bryant, 2012; 33.5%, \(n=123\)). There is a Spanish version of the CTOPP called the *Test of Phonological Processing in Spanish* (TOPPS; Francis et al., 2001), however participants’ open text responses suggested that they commonly use the English version with ELs. Further, one participant erroneously claimed, “The CTOPP doesn’t require prior knowledge, so it’s helpful to use to assess phonological processing and also language impairment in general.” The Spanish version of the WJ ACH, the *Woodock-Muños Language Survey* (WMLS; Schrank et al., 2010) was included as a separate survey item, and the WIAT and GORT do not provide accommodations for ELs. Therefore, the most frequently selected assessments used to measure reading skills in ELs were not necessarily designed for use with this population. Interestingly, all benchmark screening assessments selected,
including the *Dynamic Indicators of Basic Literacy Skills* (DIBELS; Kaminski et al., 2002, 26.4%, n=97), AIMSweb (Shinn & Shinn, 2002; 19.3%, n=71), *Fountas & Pinnell Benchmark Assessment System* (BAS; Fountas & Pinnell, 2008; 18.8%, n=69), and *Developmental Reading Assessment* (DRA; Beaver & Carter, 1997; 17.4%, n=64) had Spanish versions, however no informal reading inventories or standardized reading assessments selected were designed for use with ELs except for the WMLS (17.7%, n=65). Additionally, no languages other than English and Spanish were represented in the assessments that practitioners selected to assess reading and academic skills. Furthermore, none of the writing assessments selected by participants, such as the *Test of Written Language* (TOWL; Hammill & Larsen, 2009; 36.8%, n=77) and *Oral and Written Language Scales* (OWLS; Carrow-Woolfolk, 2011; 22.5%, n=47), were designed for use with ELs. This suggests that formal reading and writing assessments for ELs suspected of an RD are often performed in English only, while formal cognitive and language assessments are more likely to accommodate ELs, particularly those whose native language is Spanish. Perhaps this lack of non-English reading and writing assessments explains why fewer bilingual special educators and reading specialists reported using their non-English language when assessing ELs, given that such professionals typically reported performing reading and writing assessments rather than language or cognitive assessments. An alternative possibility is that the ELs that these practitioners have worked with were not biliterate given that the majority of ELs in U.S. public schools enter in
early elementary school (McFarland et al., 2019) and therefore may not have received literacy instruction in their first language.

Participants were then asked the main reasons for selecting these measures when assessing ELs. The majority of participants (72.8%, n=265) indicated that they selected these assessments because these were the measures that they had access to. Approximately half (51.9%, n=189) of the respondents said that these measures were selected because they were trained to use them. Ease of use (34.3%, n=125), quick administration (29.7%, n=108), and obligation (29.7%, n=108) were other common reasons identified. Only 15.6% selected affordability as the main reason and written-in responses indicated that such measures were selected because they were valid and reliable (5.8%, n=21) and because they were the best measures to use when assessing ELs (3.8%, n=14).
### Table 1.3
**Most Frequently Administered Standardized Assessments and Informal Inventories Used by Practitioners to Identify RDs in ELs (n=367)**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Assessments (n=190)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WISC</strong></td>
<td>Wechsler Intelligence Scale for Children (Wechsler, 2014)</td>
<td><strong>Measures:</strong> Full Scale IQ, Verbal Comprehension, Visual Spatial Skills, Fluid Reasoning, Working Memory, and Processing Speed</td>
<td>Not in English version</td>
<td>Spanish version appropriate for Spanish-speaking ELs</td>
<td>120</td>
<td>63.2%</td>
</tr>
<tr>
<td><strong>TONI</strong></td>
<td>Test of Nonverbal Intelligence (Brown et al., 2010)</td>
<td><strong>Measures:</strong> Intelligence, aptitude, abstract reasoning, and problem solving despite requiring minimal language demands</td>
<td>Yes</td>
<td>No, but test directions in Spanish, French, German, Chinese, Vietnamese, Korean, and Tagalog for use with ELs</td>
<td>73</td>
<td>38.4%</td>
</tr>
<tr>
<td><strong>KABC</strong></td>
<td>Kaufman Assessment Battery for Children (Kaufman &amp; Kaufman, 2004a)</td>
<td><strong>Measures:</strong> Sequential processing, simultaneous processing, learning ability, planning ability, and knowledge</td>
<td>No, however the test makers claim that the assessment is “culturally fair” for bilingual/EL students</td>
<td>No, but correct responses and teaching text are provided in Spanish for ELs</td>
<td>54</td>
<td>28.4%</td>
</tr>
<tr>
<td><strong>KBIT</strong></td>
<td>Kaufman Brief Intelligence Test (Kaufman &amp; Kaufman, 2004b)</td>
<td><strong>Measures:</strong> Verbal and Nonverbal intelligence</td>
<td>No</td>
<td>No</td>
<td>21</td>
<td>11.1%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>SBIS</td>
<td>Stanford-Binet Intelligence Scales (Roid &amp; Barram, 2004)</td>
<td><strong>Measures:</strong> Full Scale IQ, Verbal IQ, Nonverbal IQ, Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory. <strong>Age range:</strong> 2-85+ years <strong>Administration time:</strong> 50 minutes</td>
<td>No</td>
<td>No</td>
<td>16</td>
<td>8.4%</td>
</tr>
<tr>
<td>UNIT</td>
<td>Universal Nonverbal Intelligence Test (Bracken &amp; McCallum, 2016)</td>
<td><strong>Measures:</strong> Memory, Reasoning, Quantitative, and Full Scale <strong>Age range:</strong> 5-21 years <strong>Administration time:</strong> 10-45 minutes</td>
<td>Yes</td>
<td>No</td>
<td>11</td>
<td>5.8%</td>
</tr>
<tr>
<td>Other (n &lt; 10)</td>
<td>e.g., Differential Ability Scales (Dumont, Willis, &amp; Elliot, 2008), Wechsler Nonverbal Scale of Ability (Wechsler &amp; Naglieri, 2006)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>24</td>
<td>12.6%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>CELF</td>
<td>Clinical Evaluation of Language Fundamentals (Wiig et al., 2013)</td>
<td><strong>Measures:</strong> Core Language Abilities, Receptive Language, Expressive Language, Language Content, Language Structure. <strong>Age range:</strong> 5-21 years. <strong>Administration time:</strong> 30-45 minutes for core language score.</td>
<td>Yes, however Paradis (2005) found that ELs tested using the English version may be incorrectly identified as having a language disorder.</td>
<td>Spanish version available</td>
<td>94</td>
<td>54.7%</td>
</tr>
<tr>
<td>PPVT</td>
<td>Peabody Picture Vocabulary Test (Dunn &amp; Dunn, 2007)</td>
<td><strong>Measures:</strong> Receptive vocabulary. <strong>Age range:</strong> 2.5-90+ years. <strong>Administration time:</strong> 10-15 minutes.</td>
<td>No</td>
<td>Spanish version called Test de Vocabulario en Imágenes Peabody (TVIP; used by 12 participants in our sample)</td>
<td>68</td>
<td>39.5%</td>
</tr>
<tr>
<td>EOWPVT</td>
<td>Expressive One-Word Picture Vocabulary Test (Martin &amp; Brownell, 2011)</td>
<td><strong>Measures:</strong> Expressive vocabulary. <strong>Age range:</strong> 2-70+ years. <strong>Administration time:</strong> 15-25 minutes.</td>
<td>No</td>
<td>Bilingual version which assesses total expressive vocabulary (English and Spanish combined)</td>
<td>56</td>
<td>32.6%</td>
</tr>
<tr>
<td>ROWPVT</td>
<td>Receptive One-Word Picture Vocabulary Test (Martin &amp; Brownell, 2010)</td>
<td><strong>Measures:</strong> Receptive vocabulary. <strong>Age range:</strong> 2-70+ years. <strong>Administration time:</strong> 15-25 minutes.</td>
<td>No</td>
<td>Bilingual version which assesses total receptive vocabulary (English and Spanish combined)</td>
<td>55</td>
<td>32.0%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------</td>
<td>------------------------------</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td><strong>TAPS</strong></td>
<td>Test of Auditory Processing Skills (Martin &amp; Brownell, 2005)</td>
<td>Measures: Phonological processing, auditory memory, listening comprehension</td>
<td>No</td>
<td>Bilingual version for Spanish speakers</td>
<td>48</td>
<td>27.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age range: 5-21 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Administration time: 60-90 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CASL</strong></td>
<td>Comprehensive Assessment of Spoken Language (Carrow-Woolfolk, 1999)</td>
<td>Measures: Lexical/semantic, syntactic, supralinguistic, and pragmatic oral language processing skills</td>
<td>No</td>
<td>No</td>
<td>42</td>
<td>24.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age range: 3-21 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Administration time: 45 minutes for General Language Ability Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EVT</strong></td>
<td>Expressive Vocabulary Test (Williams, 1997)</td>
<td>Measures: Expressive vocabulary</td>
<td>No</td>
<td>No</td>
<td>33</td>
<td>19.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age range: 2.5-90+ years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Administration time: 10-20 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BVAT</strong></td>
<td>Bilingual Verbal Ability Tests (Muñoz-Sandoval et al., 1998)</td>
<td>Measures: Bilingual verbal ability (English and home language combined), English proficiency, Receptive Vocabulary</td>
<td>Yes</td>
<td>Items initial given in English and then incorrect items administered in native language (16 languages represented)</td>
<td>32</td>
<td>18.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age range: 5 years - adult</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Administration time: 30 minutes or less</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>------------------------------</td>
<td>---</td>
<td>----</td>
</tr>
<tr>
<td>TACL</td>
<td>Test for Auditory Comprehension of Language (Carrow-Woolfolk, 1985)</td>
<td>Measures: Receptive vocabulary, grammar, and syntax</td>
<td>No</td>
<td>No</td>
<td>16</td>
<td>9.3%</td>
</tr>
<tr>
<td>Other (n &lt; 10)</td>
<td>e.g., Ortiz Picture Vocabulary Acquisition Test (Ortiz, 2018), Test of Integrated Language and Literacy Skills (Nelson, Plante, Helm-Estabrookes, &amp; Hotz, 2016)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>21</td>
<td>12.2%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>CTOPP</td>
<td>Comprehensive Test of Phonological Processing (Wagner et al., 1999)</td>
<td>Type: Norm-referenced reading skill Measures: Phonological awareness, phonological memory, rapid naming Age range: 5-24 years Administration time: 30 minutes</td>
<td>No</td>
<td>Spanish version called Test of Phonological Processing in Spanish (TOPPS)</td>
<td>191</td>
<td>52.0%</td>
</tr>
<tr>
<td>WIAT</td>
<td>Wechsler Individual Achievement Test (Wechsler, 1992)</td>
<td>Type: Norm-referenced academic achievement Measures: Oral Language, Total Reading, Basic Reading, Reading Comprehension and Fluency, Written Expression, Mathematics, Math Fluency, Total Achievement Age range: 4-50 years Administration time: 35-104 minutes</td>
<td>No</td>
<td>No</td>
<td>143</td>
<td>39.0%</td>
</tr>
<tr>
<td>WJ ACH</td>
<td>Woodcock-Johnson Achievement (Schrank et al., 2014)</td>
<td>Type: Norm-referenced academic achievement Measures: Brief Achievement, Reading, Reading Fluency, Reading Rate, Reading Comprehension, Broad Reading, Basic Reading Skills, Written Language, Mathematics, Math Problem Solving Age range: 2-80+ years Time: Varies</td>
<td>No</td>
<td>Spanish version called Woodcock-Muños Language Survey (see below)</td>
<td>142</td>
<td>38.7%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-------------</td>
<td>------------------------------</td>
<td>-----------------------------</td>
<td>----</td>
<td>------</td>
</tr>
</tbody>
</table>
| GORT    | Gray Oral Reading Tests (Wiederholt & Bryant, 2012) | **Type:** Norm-referenced reading  
**Measures:** Rate, Accuracy, Fluency, Comprehension, Oral Reading Index  
**Age range:** 6-23 years  
**Administration time:** 30 minutes | No | No | 123 | 33.5% |
| DIBELS  | Dynamic Indicators of Basic Early Literacy Skills (Kaminski et al., 2002) | **Type:** Benchmark screening  
**Measures:** Phonological awareness, alphabetic principle and phonics, accuracy, fluency, comprehension, vocabulary, oral language  
**Grade range:** Kindergarten-8th  
**Administration time:** 30 minutes | N/A | Spanish version called Indicadores Dinámicos del Exito en la Lectura (IDEL) | 97 | 26.4% |
| KTEA    | Kaufman Test of Education Achievement (Kaufman, 2004) | **Type:** Norm-referenced academic achievement  
**Measures:** Oral expression, listening comprehension, basic reading skills, reading comprehension, reading fluency, written expression, mathematics calculation, mathematics problem solving  
**Age range:** 4-25 years  
**Administration time:** 15-85 minutes | No | No | 87 | 23.7% |
<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
</table>
| TOWRE        | Test of Word Reading Efficiency (Torgesen et al., 2012)               | **Type:** Norm-referenced word reading  
**Measures:** Sight word efficiency, phonemic decoding efficiency  
**Age range:** 6-24 years  
**Administration time:** 5-10 minutes | No                           | No                          | 73  | 19.9%   |
| AIMSweb      | AIMSweb (Shinn & Shinn, 2002)                                         | **Type:** Digital benchmark screening  
**Measures:** Early literacy, reading, spelling, writing, early numeracy, computation, concepts and applications  
**Grade range:** Kindergarten-12th  
**Administration time:** varies | N/A                          | Spanish early reading and reading assessments included | 71  | 19.3%   |
| BAS          | Fountas & Pinnell Benchmark Assessment System (Fountas & Pinnell, 2008) | **Type:** Benchmark screening  
**Measures:** Reading level, oral reading rate  
**Grade range:** Kindergarten-8th  
**Administration time:** 20-40 minutes | N/A                          | Spanish version called Sistema de evaluación de la lectura (SEL) | 69  | 18.8%   |
| WMLS         | Woodcock-Muñoz Language Survey (Schrank et al., 2010)                  | **Type:** Norm-referenced academic achievement  
**Measures:** Listening, Speaking, Reading, Writing, and Comprehension  
**Age range:** 2-80+ years  
**Administration time:** 35-45 minutes | Yes                         | Spanish version of the WJ ACH         | 65  | 17.7%   |
<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
</table>
| **DRA** | Developmental Reading Assessment (Beaver & Carter, 1997) | Type: Benchmark screening  
Measures: Reading level, oral reading rate  
Grade range: Kindergarten-8th  
Administration time: 20-40 minutes | N/A | Spanish version called Evaluación del Desarrollo de la Lectura (EDL) | 64 | 17.4% |
| **RAN/RAS** | Rapid Automated Naming and Rapid Alternating Stimulus Tests (Wolf & Denkla, 2005) | Type: Norm-referenced reading skill  
Measures: Rapid automatized naming  
Age range: 5-18 years  
Administration time: 5-10 minutes | No | No | 64 | 17.4% |
| **WRMT** | Woodcock Reading Mastery Tests (Woodcock, 1987) | Type: Norm-referenced reading  
Measures: Total reading, readiness, basic skills, reading comprehension  
Age range: 4.5-79 years  
Administration time: 15-45 minutes | No | No | 54 | 14.7% |
| **WIST** | Word Identification and Spelling Test (Wilson & Felton, 2004) | Type: Norm-referenced reading skill  
Measures: Word identification, spelling, sound-symbol knowledge  
Age range: 7-18 years  
Administration time: 40 minutes | No | No | 34 | 9.3% |
<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>QRI</td>
<td>Qualitative Reading Inventory</td>
<td>Type: Reading inventory Measures: Reading levels, informal assessment of reading skills Grade range: K-12 Administration time: Varies</td>
<td>N/A</td>
<td>No, but there is a Spanish Reading Inventory appropriate for use with ELs (Johns &amp; Daniel, 2010)</td>
<td>33</td>
<td>9.0%</td>
</tr>
<tr>
<td>WADE</td>
<td>Wilson Assessment of Decoding and Encoding</td>
<td>Type: Norm-referenced reading skill Measures: Word identification, spelling, and sound-symbol knowledge Age range: 7-18 years Administration time: 40 minutes</td>
<td>No</td>
<td>No</td>
<td>29</td>
<td>7.9%</td>
</tr>
<tr>
<td>TOSWRF</td>
<td>Test of Silent Word Reading Fluency</td>
<td>Type: Norm-referenced reading skill Measures: Sight word proficiency Age range: 6-24 years Administration time: 3 minutes per form</td>
<td>No</td>
<td>No</td>
<td>24</td>
<td>6.5%</td>
</tr>
<tr>
<td>TORC</td>
<td>Test of Reading Comprehension</td>
<td>Type: Norm-referenced reading Measures: Reading comprehension Age range: 7-17 years Administration time: 25 minutes</td>
<td>No</td>
<td>No</td>
<td>24</td>
<td>6.5%</td>
</tr>
<tr>
<td>Measure</td>
<td>Full Name and Citation</td>
<td>Description</td>
<td>Bilinguals in norming sample</td>
<td>Bilingual or Spanish version</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>----</td>
<td>----</td>
</tr>
</tbody>
</table>
| SRI     | Scholastic Reading Inventory | Type: Computer adaptive reading inventory  
Measures: Reading levels, informal assessment of reading skills  
Grade range: K-6th  
Administration time: Varies  
Age range: 7-25 years  
Administration time: 15-20 minutes | N/A | No | 23 | 6.3% |
| GSRT    | Gray Silent Reading Tests | Type: Norm-referenced reading  
Measures: Silent reading comprehension  
Age range: 7-25 years  
Administration time: 15-20 minutes | No | No | 23 | 6.3% |
| DAR     | Diagnostic Assessments of Reading | Type: Criterion-referenced reading  
Measures: Phonemic awareness, phonics, vocabulary, reading fluency, reading comprehension  
Grade range: K-12, adult  
Administration time: 40 minutes | No | No | 17 | 4.6% |
| WRAT    | Wide Range Achievement Test | Type: Norm-referenced academic achievement  
Measures: Word reading, sentence comprehension, spelling, math computation, reading composite  
Age range: 5-85+ years  
Administration time: 15-45 minutes | No | No | 15 | 4.1% |
<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
</table>
| **GRADE** | Group Reading Assessment & Diagnostic Evaluation (Pearson Education, Inc., 2014) | **Type:** Norm-referenced reading  
**Measures:** Early reading skills, vocabulary, fluency, comprehension  
**Grade range:** PreK-12, adult  
**Administration time:** 50-90 minutes | No | No | 11 | 3.0% |
| **PAL** | Process Assessment of the Learner (Berninger, 2007) | **Type:** Norm-referenced reading  
**Measures:** Basic reading and writing skills, rapid automatic naming, executive functions  
**Grade range:** K-6  
**Administration time:** Varies | No | No | 10 | 2.7% |
<p>| <strong>Other</strong> (&lt;10) | e.g., Nelson-Denny Reading Test (Brown, 1960), Gates-MacGinitie Reading Tests (Gates &amp; MacGinitie, 1964), Analytical Reading Inventory (Woods &amp; Moe, 2015) | - | - | - | 69 | 18.8% |</p>
<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Name and Citation</th>
<th>Description</th>
<th>Bilinguals in norming sample</th>
<th>Bilingual or Spanish version</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOWL</strong></td>
<td>Test of Written Language</td>
<td>Measures: Overall writing, contrived writing, and spontaneous writing</td>
<td>No</td>
<td>No</td>
<td>77</td>
<td>36.8%</td>
</tr>
<tr>
<td>(Hammill &amp; Larsen, 2009)</td>
<td>Age range: 9-18 years</td>
<td>Administration time: 60-90 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OWLS</strong></td>
<td>Oral and Written Language Scales</td>
<td>Measures: Oral, Written Language, Receptive Processing, Expressive Processing, Overall Language Processing</td>
<td>No</td>
<td>No</td>
<td>47</td>
<td>22.5%</td>
</tr>
<tr>
<td>(Carrow-Woolfolk, 2011)</td>
<td>Age range: 3-22 years</td>
<td>Administration time: 10-30 minutes per test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TEWL</strong></td>
<td>Test of Early Written Language</td>
<td>Measures: Basic writing, contextual writing, overall writing</td>
<td>No</td>
<td>No</td>
<td>36</td>
<td>17.2%</td>
</tr>
<tr>
<td>(Hresko et al., 2012)</td>
<td>Age range: 4-12 years</td>
<td>Administration time: 30-50 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TWS</strong></td>
<td>Tests of Written Spelling</td>
<td>Measures: Written spelling</td>
<td>No</td>
<td>No</td>
<td>36</td>
<td>17.2%</td>
</tr>
<tr>
<td>(Larsen et al., 2013)</td>
<td>Age range: 6-18 years</td>
<td>Administration time: 20 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOC</strong></td>
<td>Test of Orthographic Competence</td>
<td>Measures: Orthographic ability, conventions, spelling accuracy, and spelling fluency</td>
<td>No</td>
<td>No</td>
<td>13</td>
<td>6.2%</td>
</tr>
<tr>
<td>(Mather et al., 2008)</td>
<td>Age range: 6-18 years</td>
<td>Administration time: 20 minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** All assessments listed in this table are administered in English unless otherwise specified. Citations and descriptions are provided for the newest version of each assessment.
**Assessment Practices.** Table 1.4 presents the frequency with which practitioners reported engaging in a variety of practices when assessing ELs for a suspected RD. A majority of participants (80.2%, n=287) indicated that they conduct standardized assessments with ELs in English only at least some of the time, and roughly one third of participants indicated that they always assess these students in English only. Likewise, 72.0% of participants (n=232) indicated that they never or rarely conduct standardized assessments with ELs in the child’s native language only. Only 10.8% of participants (n=38) reported always conducting standardized assessments with ELs in both English and the child’s native language, while 49.4% (n=174) reported doing so some or most of the time, and 39.8% (n=140) reported doing so rarely or never. Results were strikingly similar for informal assessments: a majority of participants reported assessing in English only at least some of the time (87.3%, n=303); 70.9% reported that they never or rarely assess in the child’s native language only (n=224); and only 10.0% of participants reported always assessing ELs in both English and the child’s native language. Additional analysis revealed that the reported frequency of conducting standardized assessments only in English differed by bilingual status ($\chi^2(4)=11.58, p=0.021$), setting ($\chi^2(16)=16.37, p=0.037$), and occupation ($\chi^2(20)=47.17, p=0.001$), such that bilingual participants, those who reported working in urban settings, as well as speech-language pathologists and school psychologists less frequently reported conducting standardized assessments with ELs in English only than their peers. Results for informal
assessments were consistent with the findings for standardized assessments for bilingual status ($\chi^2(20)=9.95, p=0.041$) and occupation ($\chi^2(20)=54.93, p<0.001$), but not for setting ($\chi^2(8)=8.54, p=0.383$). It makes sense that bilingual professionals would less frequently assess ELs in English only than their monolingual peers, given that they are more likely to have the capacity to use their non-English language with students from similar language backgrounds. Additionally, given that speech-language pathologists and school psychologists typically reported conducting cognitive and language assessments, many of which accommodate ELs, it makes sense that they would less frequently assess ELs in English only as compared to other professionals.

When asked why ELs are not all tested in their native language in their setting, participants most commonly reported a lack of bilingual practitioners (46.6%, $n=231$) and/or a lack of appropriate assessments for ELs (37.5%, $n=186$). Approximately nineteen percent of participants ($n=95$) reported that students who speak languages for which they have non-English assessments and practitioners are always assessed in their native language, but other ELs are only assessed in English. Only 10.9 percent of participants ($n=54$) indicated that all ELs in their settings are always assessed in their native language. Few participants (6.5%, $n=32$) reported a lack of time to test ELs in their native language. Nineteen participants (3.8%) wrote in that ELs with higher levels of English proficiency are assessed in English only in their setting (e.g., “If their English proficiency is considered to be sufficient that it is their primary academic
language, testing in native language is not required,” “Depends on level of English proficiency...if level is low then assessment is in native language...if level is adequate, then assess in English”). An additional nineteen participants (3.8%) indicated that ELs are not assessed in their native language due to low levels of academic language proficiency in their native language (e.g., “If BVAT and WIDA results show student is not proficient in native language...and student has never had instruction in native language, proceed with assessment in English,” “Some ELs only have conversational first languages...they do not read or write in their first language and they have never received academic instruction in their first language; assessing these children in academic areas in their first language often shows significantly low scores that are not reflective of their true abilities”). Five participants wrote in that they believed it was against state, district, or school policy to assess ELs in their native language (e.g., “I thought it was illegal to instruct/assess EL students in their native languages so I don't”). Only one participant indicated that they did not believe it was important to assess ELs in their native language. Taken together, these findings suggest that there is wide variability in the frequency of assessing ELs in both English and their native language, and this may be explained by a lack of resources to assess ELs in an accurate and timely manner.

Participants also reported that they examined assessments used with ELs at least some of the time for cultural bias (75.9%, n=264), inclusion of bilinguals in the norming sample (73.4%, n=251), and to ensure that they are
being used in a valid and reliable manner (83.6%, n=290), with 15% or fewer of participants reporting that they never engage in these practices. The distribution of responses to these questions differed by participant occupation, however, such that reading specialists and special educators reported less frequently examining assessments for cultural bias ($\chi^2(20)=71.63, p<0.001$), bilingual representation in the norming sample ($\chi^2(20)=53.01, p<0.001$), and valid and reliable use ($\chi^2(20)=57.26, p<0.001$) than those in other professions. These are also the practitioners who most often reported conducting reading and writing assessments, which might explain why they report examining assessments for such factors less frequently. For instance, only 21.1% of reading specialists and 35.0% of special educators reported always examining assessments to ensure their valid and reliable use, while 50% or more of the participants in each of the other professions reported doing so. Furthermore, some participants noted in the open response fields that they use standardized reading and writing assessments with ELs but do not consider the scores to be valid (e.g., “These tests are used to gain information, but scores are not reported”) or that they rely mainly on informal measures due to the lack of such assessments available for this population (e.g., “I primarily rely on non-standardized and dynamic assessment—also, I work exclusively with ELs”).

The use of interpreters to conduct assessments with ELs is not considered to be best practice (Abedi, 2006), however participants reported in the open response fields that they sometimes use interpreters in cases where there is no
trained bilingual assessor and/or appropriate instruments available in the child’s native language. Participants were asked the frequency with which they have used interpreters to translate or conduct assessments and found that 22.0% reported never doing so ($n=79$), 45.7% reported doing so rarely or some of the time ($n=164$), and 32.3% reported doing so most of the time or always ($n=116$). Practitioners were also asked whether they engage in dynamic assessment, an informal assessment practice considered to be ideal for use with ELs (Roseberry-McKibben & Hanlon, 2005), and found that 24.9% of participants reported never engaging in dynamic assessment with ELs ($n=86$), while 16.8% reported using it rarely ($n=58$), and 58.4% reported using it at least some of the time ($n=202$). The proportion of responses to this question differed significantly by occupation ($\chi^2(16)=32.18$, $p=0.009$), such that school psychologists reported less frequently engaging in dynamic assessment than other professionals. A majority of participants also reported observing the student in multiple classroom contexts (70.9%, $n=261$), gathering information about the child from their teachers (96.6%, $n=364$), and obtaining information from their parents (83.9%, $n=313$) at least most of the time, if not always.
Table 1.4.
Percent of participants who utilize each of the following assessment practices or methods when identifying RDs in ELs

<table>
<thead>
<tr>
<th>Practice</th>
<th>Total n</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conducts standardized assessments...</td>
<td>358</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in English only</td>
<td>7.0%</td>
<td>12.9%</td>
<td>20.1%</td>
<td>27.9%</td>
<td>32.1%</td>
<td></td>
</tr>
<tr>
<td>in child’s native language only</td>
<td>44.4%</td>
<td>27.6%</td>
<td>18.6%</td>
<td>6.2%</td>
<td>3.1%</td>
<td></td>
</tr>
<tr>
<td>in both English and the child’s native language</td>
<td>24.7%</td>
<td>15.1%</td>
<td>23.3%</td>
<td>26.1%</td>
<td>10.8%</td>
<td></td>
</tr>
<tr>
<td>Conducts informal assessments...</td>
<td>347</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in English only</td>
<td>6.1%</td>
<td>6.6%</td>
<td>23.9%</td>
<td>30.8%</td>
<td>32.6%</td>
<td></td>
</tr>
<tr>
<td>in child’s native language only</td>
<td>44.9%</td>
<td>26.0%</td>
<td>21.8%</td>
<td>4.4%</td>
<td>2.9%</td>
<td></td>
</tr>
<tr>
<td>in both English and the child’s native language</td>
<td>27.4%</td>
<td>17.7%</td>
<td>23.9%</td>
<td>20.9%</td>
<td>10.0%</td>
<td></td>
</tr>
<tr>
<td>Examines assessments for...</td>
<td>348</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cultural bias</td>
<td>14.1%</td>
<td>10.1%</td>
<td>16.1%</td>
<td>22.1%</td>
<td>37.6%</td>
<td></td>
</tr>
<tr>
<td>inclusion of ELs/bilinguals in the norming sample to ensure that they</td>
<td>15.2%</td>
<td>11.4%</td>
<td>18.4%</td>
<td>20.5%</td>
<td>34.5%</td>
<td></td>
</tr>
<tr>
<td>are being used in a valid and reliable manner</td>
<td>(52)</td>
<td>(39)</td>
<td>(63)</td>
<td>(70)</td>
<td>(118)</td>
<td></td>
</tr>
<tr>
<td>Gathers information about child in multiple classroom contexts</td>
<td>347</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from teachers</td>
<td>10.1%</td>
<td>6.3%</td>
<td>13.8%</td>
<td>24.5%</td>
<td>45.2%</td>
<td></td>
</tr>
<tr>
<td>from parents</td>
<td>(35)</td>
<td>(22)</td>
<td>(48)</td>
<td>(85)</td>
<td>(157)</td>
<td></td>
</tr>
<tr>
<td>Gathers information about child from teachers</td>
<td>377</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from teachers</td>
<td>0.3%</td>
<td>0.5%</td>
<td>2.7%</td>
<td>16.5%</td>
<td>80.1%</td>
<td></td>
</tr>
<tr>
<td>Gathers information about child from parents</td>
<td>373</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from teachers</td>
<td>1.3%</td>
<td>4.3%</td>
<td>10.5%</td>
<td>19.3%</td>
<td>64.6%</td>
<td></td>
</tr>
</tbody>
</table>

Criteria Used to Identify RDs. While practitioners were not asked the criteria that they have used to identify RDs in ELs specifically, they were asked
the criteria that they have use to identify RDs generally. As illustrated in Figure 1.2, below, 61.6% of participants reported using the IQ/Achievement Discrepancy \((n=262)\). When asked specifically whether they use this criterion when evaluating ELs for a suspected RD, just under one third of participants \((31.7\%, \ n=113)\) reported never using this approach with ELs, 36.8% \((n=131)\) indicated that they use it rarely or some of the time, and 31.5% \((n=112)\) reported using it most of the time or always. The proportion of responses to this question differed by occupation \((\chi^2(20)=44.82, \ p=0.001)\) and region \((\chi^2(3)=11.55, \ p=0.009)\), such that special educators reported using the IQ/Achievement discrepancy with ELs more frequently than those in other professions and those who work in the South reported using it less frequently than those in other regions. Other commonly reported criteria used to identify RDs included a failure to respond to Tier 2 (small group) intervention using a RTI approach \((58.4\%, \ n=248)\), a standard deviation below the mean on multiple measures \((39.5\%, \ n=168)\), and failure to respond to Tier 1 (whole class) instruction using RTI \((28.0\%, \ n=119)\). External diagnosis was selected by 16.2 percent of practitioners \((n=69)\). When asked who provides the external diagnosis, written-in responses included clinical psychologists, neuropsychologists, neurologists, pediatricians, trained bilingual assessors, and diagnosticians. Other reported criteria used to identify RDs in the general student population included a standard deviation below the mean on one measure \((11.5\%, \ n=49)\), the PSW approach \((11.3\%, \ n=48)\), and a failure to respond to Tier 3 (individualized) intervention using a RTI approach \((2.6\%, \ n=11)\). A small
percentage of participants (6.1%, \(n=26\)) wrote in that they use other criteria, such as “team agreement,” “evidence of a double deficit,” and “Dyslexia profile worksheet,” and 3.5% of participants (\(n=26\)) did not know which criteria were used to make a determination for RDs in their setting. A majority of participants (65.3%, \(n=294\)) selected more than one criterion, suggesting that practitioners may use multiple criteria simultaneously or may use different criteria to establish an RD depending on the circumstances. While ninety percent of participants (\(n=405\)) reported using an RTI approach, the relatively small percentage of participants reported using a PSW approach (11.3%, \(n=48\)) is likely due to sampling error rather than being truly reflective of the prevalence of this approach, as those who reported using PSW were more likely to be from the underrepresented Midwest and Western U.S. regions (\(\chi^2(3)=26.24, p<0.001\)). While this illustrates that multiple approaches are typically used in current practice to identify RDs, additional research is necessary to explore the prevalence of the criteria used with ELs in a representative sample of practitioners.
Figure 1.2. Identification Criteria that Participants in the Analytic Sample Reportedly Use to Identify RDs (n=450)

**Question 3: Practitioners’ Perceptions**

**Primary Causes of Reading Difficulty in ELs.** Participants were asked about their perceptions of the primary causes of word reading and reading comprehension difficulties in the ELs that they have assessed for suspected RDs. Interestingly, an underlying RD was not the most commonly perceived cause of word reading or reading comprehension difficulties among ELs referred for evaluation. Instead, a lack of English proficiency was the most commonly selected explanation for word-level difficulties (68.1%, n=314), followed by inadequate reading instruction (60.5%, n=279), an underlying RD (55.8%, n=257), and trouble distinguishing between speech sounds not present in the child’s native language (49.2%, n=227). Similarly, difficulty with word-level
reading skills was the most commonly reported perceived cause of reading comprehension difficulties in ELs (69.1%, \(n=320\)), followed by a lack of oral language proficiency in English (67.7%, \(n=294\)) and inadequate reading instruction (56.4%, \(n=261\)), prior to an underlying RD, which was the fourth most commonly selected perceived cause (50.8%, \(n=235\)). Other reasons selected as primary causes of word-level difficulties in ELs were executive function deficits (34.3%, \(n=158\)), cognitive deficits (30.2%, \(n=139\)), and social-emotional or behavioral issues (17.4%, \(n=80\)). Some additional causes were selected with similar frequency for reading comprehension difficulties among ELs including: executive function deficits (31.1%, \(n=144\)), cognitive deficits (29.4%, \(n=136\)), and social-emotional or behavioral issues (16.4%, \(n=76\)).

A number of participants further explained in the open response section that ELs referred for evaluation with word reading difficulties often struggle due to insufficient reading instruction in school or at home (e.g., “inadequate reading instruction based on...word memory, context clues and guessing...complete lack of phonemic proficiency instruction or even acknowledging that sounds of language differ,” “very little print-rich environments at home.... no reading or vocabulary development before they come to school,” “may have just arrived in US as a teen and missed developmental reading instruction offered in lower grades”). In the open response section, participants attributed difficulties with reading comprehension due to a lack of academic English proficiency (e.g., “ELs with basic interpersonal communication skills are mistaken for having...”)
competence in English.” “[ELs] lack proficiency in CALP [Cognitive Academic Language Proficiency],” “Some ELs have an innate disorder, but the majority of referrals have not had adequate instruction in English and have insufficient English skills for the classroom,” “vocabulary deficits and background knowledge”). Taken together, these results suggest that ELs may often be referred for services due to a lack of appropriate language and literacy instruction rather than a true, underlying disability.

**Challenges to Identification.** Out of 481 respondents, only four indicated that they believe there are no challenges to timely and accurate identification of RDs in ELs. The most commonly reported challenge was a lack of bilingual practitioners (64.9%, n=312) followed by a lack of training on best assessment practices with ELs (59.9%, n=288) and a lack of appropriate non-English assessments (52.6%, n=253). Other responses selected by fewer than half of participants were the improper use of tiered intervention models (40.7%, n=196), a lack of appropriate assessments in English (37.6%, n=181), insufficient research on best assessment practices with ELs (32.0%, n=154), and non-cooperative families (24.7%, n=119).

Open-ended responses included further discussion of inadequate language and literacy instruction for ELs (e.g., “Focus on reading words and ‘catching up’ without a structure of foundational English sounds and symbol taught systematically,” “Lack of access to needed services outside of special education,” “Young ELs don't have a foundation in either language, due to
transiency”) as well as barriers to evaluation (e.g., “Schools assume they are behind due to language deficits and won’t test them,” “Some districts look the other way when EL families are not there to advocate for their children, for a variety of reasons,” “Teachers’ reluctance to refer ELs, thinking their poor reading is due to incomplete L2 acquisition”).

A few other practitioners specified in this section that more time is necessary for ELs to fully acquire English before evaluation, yet another practitioner noted that this “idea that a student should be proficient in English first” is a challenge to ELs’ appropriate and timely identification. Research suggests that ELs do not need to have fully developed oral language proficiency in English to be assessed for an RD (Geva, 2000), and 66.0 percent of the sample (n=323) correctly identified this when asked, while 22.1 percent were incorrect (n=108) and 11.9 percent said they didn’t know (n=58). It is alarming that 34 percent of participants, all of whom have experience identifying RDs in ELs, lack this basic knowledge regarding the ability to assess ELs prior to full English language proficiency. These findings suggest that additional training regarding best practices in evaluating ELs for RDs, training of additional bilingual practitioners, as well as the development of additional assessments in English and other languages appropriate for use with ELs may help overcome challenges in timely and accurate identification of RDs in ELs.

Furthermore, when asked to rate their confidence in identifying RDs in ELs on a scale from 1 (least confident) to 10 (most confident), participants’
responses ranged from 1 to 10 with a mean response of 5.69, as illustrated in the right-hand plot in Figure 1.3. This stands in contrast to their responses regarding their overall level of confidence in identifying RDs in the general population, which ranged from 2 to 10 with a mean response of 8.1, as illustrated in the left-hand plot in Figure 1.3. This suggests that many practitioners who are responsible for identifying RDs in ELs lack confidence in doing so. When asked how important they believe it is to learn more about the process of identifying RDs in ELs, 67.3% said that they believe it to be extremely important (n=305), 19.4% said it was very important (n=88), 9.1% said it was moderately important (n=41), and 4.2% said it was slightly or not at all important (n=19). Participants who responded that they believed it to be very or extremely important to learn more about the process of identifying RDs in ELs had a higher reported percentage of ELs on their caseloads than those who indicated that it was less important to learn about this process, on average (21.2% versus 8.9%, respectively). This is promising because it suggests that many practitioners who regularly work with ELs would be open to receiving additional training on this topic to hopefully boost their confidence and improve the identification process for their students.
Discussion

The present study investigated the range of current practices in identifying RDs among ELs through an in-depth survey of school- and clinic-based practitioners from across the U.S. Results indicated that a team of practitioners and educators are typically involved in this process, with school psychologists often conducting cognitive assessments, speech-language pathologists often conducting language assessments, and various professionals tasked with conducting reading and writing assessments. ESL teachers and general educators were frequently reported to conduct assessments involved in this process, however there were too few in the analytic sample to explore which kinds of assessments they typically conduct. Exploration of the specific assessment selected by such participants, as well as their written-in responses, suggests that ESL teachers typically perform English proficiency testing while the general educators conduct screening assessments as part of the RTI process.
Professionals involved in this process were largely reported to be monolingual English-speaking or had proficiency in languages not typically spoken by ELs. The one exception was Spanish, which was reportedly spoken by a fifth of survey respondents. Despite this prevalence, only 12% of all participants reported using a non-English language to assess ELs, which suggests that bilinguals involved in this process do not necessarily use their non-English language in this role. Findings suggested that standardized reading and writing assessments are typically conducted in English only using measures not valid for use with ELs. However, many cognitive and language assessments reported can accommodate Spanish speakers, and a handful of them (e.g., BVAT, non-verbal IQ tests) are appropriate for speakers of other languages as well.

A lack of English proficiency and insufficient literacy instruction were among the most common perceived causes of reading difficulties among ELs referred for evaluation for a suspected RD. This suggests that English language development programs and/or classroom-level literacy instruction may not be adequate at addressing ELs’ experience-related reading difficulties in some settings. Lastly, the greatest reported challenges to accurate and timely identification for ELs still included a shortage of bilingual practitioners, a scarcity of valid and reliable assessments in both English and non-English languages, a lack of parental involvement, and inadequate training, consistent with prior research (Harry & Klingner, 2006; Sánchez et al., 2010; Zehler et al., 2003).
Implications for Policy, Practice, and Research

The findings of this study suggest that many of the practical challenges to identifying RDs in ELs have remained the same over the past twenty years despite an increase of over 1 million ELs in U.S. schools over this time period (McFarland et al., 2019). Federal policy (IDEA, 2004) requires that assessments used with ELs be valid and reliable, culturally unbiased, and provided in the child’s native language except in cases where it is infeasible. The findings suggest that this policy is not always complied with, in practice, as few of the assessments that participants reported using with ELs were developed in a manner that would be appropriate and valid for their use with this population, and a majority of participants (60%) indicated that they assess ELs in English only most of the time or always. Additionally, a subset of participants reported never or rarely examining assessments used with ELs for cultural bias (24%), inclusion of ELs in the norming sample (27%), or their valid and reliable use (16%). The scarcity of valid and reliable reading and writing assessments for ELs in English and the lack of non-English assessments for ELs, particularly in languages other than Spanish, can be addressed through the development of new measures for this population.

Additional bilingual evaluators are also necessary to administer such assessments, as participants also indicated the shortage of bilingual practitioners as a barrier to accurate identification for ELs in their setting. In their survey of 276 bilingual school psychologists from across the U.S., O'bryon and Rogers (2010)
found that a majority of these practitioners engaged in best assessment practices with ELs, including using multifaceted assessment approaches across multiple languages. This stands in contrast to my findings across the broader population of practitioners and implies that such practices would greatly improve with the inclusion of additional bilingual practitioners in schools. Nevertheless, there are reportedly fewer than ten bilingual school psychology programs (Sotelo-Dynega, 2015) and fifteen bilingual special education teacher programs (Wang and Woolf, 2015) in the U.S. Additional recruitment and training efforts are necessary to fulfill this goal of increasing the number of bilingual practitioners in U.S. schools.

In addition, there is substantial research on effective English language and literacy instruction for ELs (e.g., August, McCardle, & Shanahan, 2014; Rivera, Moughamian, Lesaux, & Francis, 2009; Snyder, Witmer, & Schmitt, 2017), yet a majority of practitioners reported that ELs are referred for special education due to reading difficulties that primarily stem from inadequate instruction in these domains rather than an underlying disability. This suggests that the research on effective practices for preventing reading difficulties among ELs does not always make its way into the classroom. For instance, Orosco and Klingner (2011) describe a case where a majority of classroom teachers in one school believed that they were using evidence-based instruction, when they were actually providing unclear and culturally inappropriate instruction for ELs, leading to inappropriate referrals for special education. Therefore, one way to prevent such inappropriate referrals, and therefore potential misidentification, would be to
ensure that classroom and ESL teachers, as well as their evaluators (e.g., school administrators), receive extensive, hands-on training in culturally responsive, evidence-based language and literacy instruction for ELs as well as information on prereferral strategies and available resources outside of special education that are available to students struggling for reasons unrelated to a disability (e.g., Title I services, community services).

Furthermore, approximately 30% of participants in the sample reported receiving no training on the process of identifying RDs in ELs during graduate school, and a similar percentage reported receiving a single PD session or less on this topic, which suggests that a sizeable percentage of practitioners who have identified RDs in ELs could benefit from additional training on this process. One-third of participants did not know that ELs could be accurately identified with an RD prior to achieving English proficiency, so this information would be relevant to include in such training sessions. Trainings should also include information about appropriate evaluation criteria for ELs, as over half of participants reporting using the IQ/Achievement discrepancy to identify RDs in ELs even though this method is not appropriate for use with this population (Huang, Clark, Milczarski, & Raby, 2011; Geva & Wiener, 2015). Information about informal assessment practices, such as dynamic assessment, should also be included in such trainings, since 43% of participants reported rarely or never using this recommended method with ELs. Although practitioners reported less confidence in evaluating for RDs in EL than in the general population, a large
majority (87%) indicated that they believed it was very or extremely important to learn more about this process in ELs. This suggests that practitioners who work with ELs would be eager to receive additional training on this topic.

**Limitations and Next Steps**

Even though this study provides a wealth of information regarding the current range of practices used to identify ELs with RDs, there are a number of limitations. Importantly, due to the nature of my recruitment approach, the sample is not representative of all practitioners with roles in identifying or diagnosing RDs in ELs across the U.S. Specifically, participants from Massachusetts were oversampled and medical professionals with roles in diagnosing RDs were greatly underrepresented. Unfortunately, many hospital clinics and private practices did not provide publicly accessible contact information on their websites, and I was unable to obtain permission to recruit through professional organizations relevant to medical professionals. The personalized emails to district leaders coupled with the tapping of my personal network may explain why such a large proportion of the sample reports working in Massachusetts. Even though recruitment emails were sent to public school district leaders across the U.S., this approach yielded a relatively low response rate with 262 total responses resulting from the 12,186 emails distributed. Supplementing the low responses, upticks in sample size directly after Facebook postings and emails to electronic lists suggest that such recruitment methods also yielded a substantial portion of the full sample. Unfortunately, due to the de-identified nature of the
survey and non-probability sampling approach, there is no measure indicating response rate for each recruitment procedure. Given that compensation was not provided upon survey completion and that participants from outside of Massachusetts were largely recruited through professional organizations, the participants who took the time to complete this survey are likely more invested and knowledgeable than the population of practitioners involved with identifying RDs in ELs across the U.S. In addition, I did not ask participants how their identification practices differ depending on student-level characteristics, such as the child’s age, instructional history, or home language. Qualitative research, such as interviews or ethnography, would be better suited to obtain such nuanced information. As a first step, however, I believe that this study provides initial insight into the variation in diagnostic processes for RDs in ELs.

Future research should seek to replicate these findings in a representative sample of practitioners with roles in identification of RDs in ELs across the U.S. Such a study should also ensure that clinic-based practitioners are well-represented so that comparisons can be made between the school- and clinic-based identification approaches. It would be especially useful to also include questions on how practitioners assess ELs’ level of language proficiency, acculturation, SES, and other relevant student-level factors, and how these characteristics factor into the decision-making process. Additional qualitative research, including ethnographic study of the RD evaluation process for ELs, is also necessary to obtain a more nuanced and objective measure of what takes
place in practice, as response bias may be present in participants’ survey responses and such approaches allow for a more in-depth understanding of the influence of contextual factors on this process.

**Conclusion**

The present study provides an initial exploration of the range of practices currently implemented in schools and clinics across the U.S. to identify RDs in ELs. Results indicate that a variety of professionals participate in this process using wide range of standardized assessments, informal measures, and evaluation criteria. These practitioners are largely monolingual English-speaking or do not assess ELs in languages other than English. Furthermore, many of their reported practices, such as frequently assessing ELs in English only using inappropriate standardized measures, do not reflect research-informed best practices for assessment with ELs and/or requirements specified by federal law (IDEA, 2004). Reported challenges to accurate and timely identification for ELs are largely consistent with previous research (Klingner and Harry, 2006; Zehler et al., 2003) and include a shortage of bilingual practitioners, dearth of appropriate assessments for ELs in English and non-English languages, difficulty involving families in the evaluation process, and insufficient training on best identification practices with ELs. Additional research is needed to replicate these findings in a representative sample, however these preliminary results suggest that additional practitioner training, bilingual practitioners, appropriate assessments for ELs, and qualitative research on this topic are crucial to improve upon the RD identification
process for ELs. Such steps are vital to ensure that these children develop essential reading skills for success in U.S. schools and society.
STUDY 2:

Exploring Behavioral and Neural Profiles of Reading Processes Among Adolescents with Diverse Language Experiences and Reading Abilities
Abstract

A substantial amount of research has been conducted on the behavioral and neural profiles of reading among students in the U.S. with and without reading disabilities (RD). However, few of these studies have included bilinguals despite the increasing prevalence of American children growing up in homes where the primary language is not English. The present study seeks to expand this body of research by comparing behavioral and neural measures of reading in English among adolescents with diverse language experiences and reading abilities. Specifically, similarities and differences in behavioral reading profiles are investigated among adolescents ages 11 to 15 ($n=76$) who are typically-developing English monolinguals, Spanish-English bilinguals, Spanish-speaking English Learners (ELs), or monolingual English speakers with RD. Neural correlates of speech and component reading processes in English are also compared by group and correlated with behavioral reading measures among a subset of these participants ($n=46$) who successfully completed a functional MRI task in which they read and listened to words, non-words, and other word-like stimuli. Typically-developing monolingual English speakers and Spanish-English bilinguals who were first exposed to English in early elementary school or before performed similarly on behavioral reading measures in English, although bilinguals exhibited slightly lower vocabulary and comprehension scores. Both monolinguals with RD and ELs exhibited difficulty across a variety of word reading measures in English, and ELs also demonstrated substantially lower
performance than all other groups on measures of English vocabulary and reading comprehension. Groups demonstrated similar neural responses to speech and component reading processes, although typically-developing monolingual English speakers recruited the left supramarginal gyrus to a greater extent for speech versus vocoded speech, while those in other groups did not.

Additionally, typically developing monolinguals recruited the left fusiform gyrus to a greater extent for written versus spoken language and the right superior temporal gyrus to a lesser extent for spoken versus written language, than bilinguals and those with RD. Lastly, exploratory brain-behavior correlations were performed across and within groups. The implications of these results and next steps for research with diverse populations are discussed.
Introduction

Reading is fundamental to academic, emotional, and economic success in a globalized society (Murnane, Sawhill, & Snow, 2012). Therefore, it is crucial that children develop the skills and knowledge needed to comprehend a wide variety of texts across different contexts. A majority of children learn to read with adequate fluency and comprehension, however others struggle with reading due to cognitive, neurological, environmental, cultural, linguistic, and/or other factors (Lesaux, 2012; McArthur & Castles, 2017). For instance, in 2015, 20% of 15-year-olds from the United States performed below the baseline proficiency level deemed necessary to fully participate in society on the Program for International Assessment (PISA) reading assessment (Kastberg, Chan, & Murray, 2016). Researchers have used a variety of approaches, including neuropsychological assessment and functional magnetic resonance imaging (fMRI) to investigate reading processes among typical and struggling readers aimed at improving the identification and remediation of reading difficulties (see Gabrieli, 2016; Eden et al., 2016).

Characteristics of Dyslexia

One group of children who struggle with reading, are those identified with a reading disability (RD). Dyslexia is a subtype of RD characterized by difficulties in accurate and/or fluent word reading (Siegel, 2006). This highly heritable developmental disorder is purported to have a neurobiological etiology (Peterson & Pennington, 2012; Ramus, 2004) and is the most common learning disability,
affecting an estimated 5 to 17% of children globally (Shaywitz, 1998). Proposed causes of dyslexia include a phonological processing deficit (Snowling, 1981; Vellutino, Fletcher, Snowling, & Scanlon, 2004), a double-deficit in phonological processing and/or rapid naming (Wolf & Bowers, 1999), the improper development of magnocellular neurons (Stein, 2001), atypical cerebellar function (Nicolson, Fawcett & Dean, 2001), visual-spatial attention deficits (Facoetti, Paganoni, Turatto, Marzola, & Mascetti, 2000), low level auditory deficits (e.g., Tallal, Miller, & Fitch, 1993), and even insufficient reading experience (Huettig, Lachman, Reis, & Petersson, 2018), among others. However, to date, there is currently no consensus on the etiology of this disorder.

The phonological processing deficit theory has received wide support, given that most dyslexics exhibit impairments in phonological awareness, phonological memory, and/or rapid naming, yet only a subset demonstrate auditory, visual, and/or motor deficits (Ramus, 2003; Peterson & Pennington, 2012). Nevertheless, in a meta-analysis of 48 studies of children with and without dyslexia, Kudo, Lussier, and Swanson (2015) found that typical readers greatly outperformed those with dyslexia (ES>0.80 SDs) on measures of word reading, reading comprehension, rapid naming, verbal IQ, phonological awareness, pseudoword reading, math, vocabulary, spelling, writing, and behavioral skills (e.g., attention and hyperactivity); moderately outperformed them (ES=0.50-0.80 SDs) on verbal working memory, inhibition, short term memory, and auditory processing; and slightly outperformed them (<0.50 SDs) on problem
solving/reasoning, visual-spatial memory, and perceptual motor skills. This suggests that the learning difficulties experienced by those with dyslexia may extend beyond the reading domain.

While children with dyslexia may exhibit a variety of cognitive, academic, and/or attentional difficulties, as compared to their peers without RDs, their participation in high-quality personalized literacy interventions, particularly at a young age, can lead to improvements or remediation of such difficulties (Shaywitz, Morris, & Shaywitz, 2008). Studies suggest that intense, systematic, and explicit instruction can remediate phonological awareness and word recognition difficulties among many children with dyslexia (Fletcher et al., 2007; Swanson et al., 2003). However, deficits in reading fluency and related processes (e.g., processing speed) are typically more resistant to treatment (Chard et al., 2002).

**Characteristics of English Learners**

Although much attention has been given to the study of dyslexia in the education and cognitive neuroscience communities, students with this disorder are not the only ones who frequently struggle with reading in U.S. schools. English Learners (EL) are students with a native language other than English who are not yet proficient in English, and therefore may experience difficulties with speaking, reading, writing, and/or understanding English that impede their ability to succeed in school (ESEA Section 8101(20)). They currently make up 10% of U.S. public school students, and over three quarters of these children
speak Spanish as their native language (McFarland et al., 2019). These students are typically instructed in English-only classrooms, and therefore are expected to acquire skills and content vital for academic success in a language that they are simultaneously trying to learn. It takes ELs an estimated 6 months to two years to learn conversational English and 5 to 7 years to achieve proficiency in academic English (Cummins, 1979; Thomas & Collier, 2002). Given the dependence of reading on language, ELs consistently perform substantially lower on measures of reading achievement as compared to their English-proficient peers (Hoff, 2013; Polat, Zarecky-Hodge, Schreiber, 2016).

A majority of ELs are U.S. born and have been taught in English since Kindergarten (Fry & Gonzales, 2008; Hernandez, Denton, & Macartney, 2008). These children typically develop word-level reading and fluency skills on par with their monolingual peers, despite their lower performance on measures of reading comprehension due to a lack of proficiency in academic English (Lesaux, Rupp, & Siegel, 2007; Nakamoto, Lindsey, & Manis, 2007). They are generally reclassified as English proficient during the upper elementary grades, although a minority of these students maintain their EL status throughout their schooling, possibly due to underlying disability and/or inadequate instruction (Motamedi, 2016; Slama, Haynes, Sacks, Lee, August, 2015; Thompson, 2017). Most of the research on ELs’ reading development focuses on this population of students in the early elementary grades (e.g., Lesaux and Siegel, 2003; Manis, Lindsey, & Bailey, 2004; Swanson, Orosco, & Lussier, 2011), and therefore less is known
regarding the reading profiles of ELs who enter U.S. schools in the upper elementary or middle school grades. Pasquarella, Gottardo, and Grant (2012) found that adolescent ELs who had moved to a Canadian city within the past two and a half years, performed one to three standard deviations below monolingual norms on decoding, vocabulary, and reading comprehension measures even though they were all literate in their native language(s). Additional research is necessary to replicate these findings in a U.S. sample; however, these results suggest that adolescent ELs may struggle to develop adequate reading skills in English despite being literate in their native language.

**Neural Basis of Reading**

In addition to behavioral findings, there is a substantial body of neuroimaging research that has contributed to our understanding of reading (see meta-analyses by Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Jobard, Crivello, & Tzourio-Mazoyer, 2003; Murphy, Jogia, & Talcott, 2019; Turkeltaub, Eden, Jones, & Zeffiro, 2002; Vigneau et al., 2006, 2011). Among typical readers, word reading is thought to involve cortical networks in the left hemisphere referred to as the ventral and dorsal systems (Cohen & Dehaene, 2009; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Jobard et al., 2003; Murphy et al., 2019). Familiar words are hypothesized to be processed along the ventral stream where basic visual processing first takes place in the occipital cortex, increasingly larger word fragments are then processed in the left occipitotemporal cortex, including the visual word-form area (VWFA), and word
meanings are then directly accessed via regions in the left middle and inferior temporal gyri as well as the pars triangularis region within the left inferior frontal gyrus (Dehaene, Cohen, Sigman, Vinckier, 2005; Murphy et al., 2019). The VWFA, located in the left posterior fusiform gyrus, is thought to be specialized for letter processing independent of case, font, location, or orientation in a bottom-up fashion, prior to lexical processing (Dehaene & Cohen, 2011). This view is controversial as other theories posit that this region responds to familiar symbols and objects beyond letters and/or is involved in integrating bottom up and top down lexical information (e.g., Kronbichler et al., 2004; Price & Devlin, 2011). Regardless of the stimulus-specificity, this region is thought to be a key hub in the ventral system that connects familiar words from print directly to their meaning.

By contrast, the dorsal pathway is theorized to process unfamiliar words and non-words by connecting left occipitotemporal regions associated with orthographic processing to left temporoparietal and frontal regions involved in phonological processing, including the parts of the supramarginal gyrus, inferior parietal lobule, superior temporal gyrus, precentral gyrus, and pars opercularis and pars orbitalis portions of the inferior frontal gyrus (Jobard et al., 2003; Murphy et al., 2019; Vigneau et al., 2006). Given that this system is used to map graphemes directly to phonemes, researchers have posited that those who are still developing their reading skills (e.g., children) as well as readers of more transparent orthographies (e.g., Spanish) may rely on this pathway to a greater
extent than expert readers and/or readers of opaque orthographies (e.g., English), who may utilize the ventral system to a greater extent (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008; Paulesu et al., 2000).

Some researchers have also hypothesized that additional regions in the frontostriatal network, composed of bilateral frontal as well as subcortical structures, are recruited during word reading when prompted by increased difficulty, such as when reading low frequency words (Black et al., 2017; Cattinelli et al., 2013; Vigneau et al., 2011). These regions, associated with attentional, motor, and/or visuospatial processes, include portions of the precentral gyrus, supplementary motor area, anterior cingulate gyrus, precuneus, thalamus, putamen, cerebellum, and inferior frontal gyrus. However, it is still debatable whether these regions are specific for print processing. For example, Murphy and colleagues (2019) attribute neural responses in a subset of these regions to task demands (e.g., use of the lexical decision paradigm) rather than their involvement in word reading.

Using a naturalistic passive word reading task, Malins and colleagues (2016) were able to avoid such potential confounding factors in their study of the neural correlates of component word reading processes among typical adult readers. They found that several regions including the VWFA, bilateral superior and middle temporal gyrus, and bilateral inferior temporal gyrus, were recruited when reading words as compared to when viewing false font, and that, conversely, was associated with higher functional activity in large bilateral areas
in the fusiform gyrus. Furthermore, the VWFA and bilateral pars opercularis responded more strongly for non-words than real words, and the reverse pattern was observed in subcortical structures. Additionally, bilateral pars opercularis and supplementary motor area as well as regions in the left precentral gyrus and supramarginal gyrus showed greater response for orthographically similar but phonologically different words (e.g., dough, cough) than orthographically and phonologically similar words (e.g., make, lake), providing empirical support that such regions may be recruited in response to lexical difficulty during word reading rather than due to task effects. Lastly, the left pars triangularis demonstrated sensitivity to semantically similar as opposed to unrelated words, and a number of regions including the left pars triangularis, left pars opercularis, left precentral gyrus, and left VWFA demonstrated significant correlations between response to semantically similar words and reading comprehension measured outside of the scanner.

Malins and colleagues (2016) designed this paradigm for its intended use with children and special populations, including those with dyslexia. In addition to passive processing, this paradigm is also advantageous to conventional tasks that have been used with children, given the possibility of assessing multiple component processes of reading in a relatively short amount of imaging time. Therefore, in the current study, I utilize this paradigm to explore the neural underpinnings of such processes among adolescents from diverse language and
literacy backgrounds, including those who struggle with reading in English due to dyslexia or a lack of English language proficiency.

**Neural Correlates of Reading in Individuals with Dyslexia**

Meta-analyses of positron emission tomography (PET), magnetoencephalography (MEG), and/or fMRI studies have consistently found that individuals with dyslexia exhibit hypoactivation of left temporoparietal and occipitotemporal regions while engaging in reading-related tasks as compared to typical readers (Maisog, Einbinder, Flowers, Turkeltaub & Eden, 2008; Martin, Kronbichler, & Richlan, 2016; Paulesu, Danelli, & Berlingeri, 2014; Pollack, Luk, & Christodoulou, 2015; Richlan, Kronbichler, & Wimmer, 2009, 2011). Reduced functional connectivity among hubs in these networks have also been observed among individuals with dyslexia during reading and at rest (Aboud, Bailey, Petrill, & Cutting, 2016; Koyama et al., 2013; Schurz et al., 2014). These findings hold across readers of different writing systems, suggesting that these differences in the dorsal and ventral systems constitute a universal neural signature of dyslexia, with variations in presentation and compensatory mechanisms depending on orthography (Martin et al., 2016). Furthermore, Richlan and colleagues (2011) propose that dysfunction of the dorsal pathway that involves print to sound mapping underlies the primary difficulty among those with dyslexia, as children reliably demonstrate hypoactivation of the temporoparietal rather than the occipitotemporal cortex, yet hypoactivation in both regions is consistently observed in adults. They posit that hypoactivation of occipitotemporal regions
within the ventral pathway develops secondarily as a result of this primary
dysfunction of the dorsal system, such that decoding difficulties lead to an
inefficient mechanism for sight word recognition and spelling.

This theory is supported by the finding that typical readers demonstrate
increased activation of the VWFA during reading-related tasks with age, relative
to its right homolog, while those with dyslexia exhibit no such effect, on average
(Shaywitz et al., 2007). Furthermore, researchers have found consistent
evidence of reduced grey matter volume and white matter integrity near
temporoparietal regions among children and adults with dyslexia (Richlan,
Kronbichler, & Wimmer, 2013; Vandermosten, Boets, Wouters, & Ghesquière,
2012), suggesting that functional differences are underpinned by structural
differences within the dorsal network. Nevertheless, both functional and structural
cuiroupatemporal and temporoparietal differences in the left hemisphere have
been reliably observed for pre-readers with a family history of dyslexia as
compared to their peers (Vandermosten, Hoeft, & Norton, 2016), so it is possible
that dysfunction in both the dorsal and ventral systems underlies dyslexia.

Some studies have also reported hyperactivation of regions within the
frontostriatal network among those with dyslexia as compared to typical readers,
particularly in the bilateral inferior frontal gyrus, precentral gyrus, supplementary
motor area, and subcortical regions (Maisog et al., 2008; Pollack et al., 2015;
Richlan et al., 2009, 2011). Given that these regions are involved in motor
planning, one theory is that this reflects a compensatory mechanism whereby
people with dyslexia use an articulatory strategy to overcome phonological processing difficulties (Black et al., 2017). This means that they rely more heavily on phonemic decoding during reading, thereby recruiting motor planning regions involved in speech production when “sounding out” words. This theory is supported by intervention outcomes, showing that children with dyslexia who respond to treatment demonstrate increased activation of bilateral regions in the frontostriatal network after receiving treatment (Barquero, Davis, & Cutting, 2014). A number of studies have also indicated that remediated dyslexics exhibit increased activation in left temporoparietal and occipital regions, similar to typical readers, and occasionally their right hemisphere homologs (Davis et al., 2011; Eden et al., 2004; Simos et al., 2002; Temple et al., 2003). Taken together, these findings suggest that those with dyslexia exhibit word-level reading difficulties due to disruptions in dorsal and ventral networks involved in reading-related processes, yet they can achieve normalization of these pathways with effective intervention and the compensatory recruitment of other brain regions that are not recruited in typical readers.

**Neural Correlates of Reading Among Biliterate Bilinguals**

Despite the surging amount of research on the neural underpinnings of reading difficulty, considerably less is known about the neural processes underlying word reading in a second language. Biliterate bilinguals showed higher activation in the dorsal pathway when reading in a shallow orthography, such as Spanish, but showed higher activation in the ventral pathway when
reading in an opaque orthography, such as English (Berken et al., 2015; Boukrina, Hanson, & Hanson, 2014; Das, Padakannaya, Pugh, & Singh, 2011; Jamal, Piche, Napolielli, Perfetti, & Eden, 2012; Oliver, Carreiras, Paz-Alonso, 2017). The order of literacy acquisition and age at which literacy in each language is achieved seems to matter, as those who first learn the more opaque orthography do not typically recruit additional regions outside of the reading network when reading in a more shallow orthography (Chee, Tan, & Thiel, 1999; Tan et al., 2003), yet those who learn the more shallow writing system first often recruit additional regions outside of the traditional reading networks when reading in the deeper orthography, such as their right hemisphere homologs and/or regions associated with executive functioning, such as the dorsolateral prefrontal cortex (DLPFC) and other frontal regions (Jamal et al., 2012; Leonard et al., 2010; Liu, Dunlap, Fiez, & Perfetti, 2007; Nelson et al., 2009; Pillai et al., 2003). The level of language and reading proficiency in each language, as well as the age of acquisition of the non-native language, seems to modulate the extent to which additional regions are recruited during second language reading, although findings are mixed regarding the specifics of these relationships (Cao, 2015).

Few neuroimaging studies of reading in a second language have been conducted with children. Hernandez, Woods, and Bradley (2015) compared the neural correlates of single word reading in English and Spanish among pre-adolescent native Spanish speakers, ages 8 to 13, who had learned English before nine years of age. They found that these children showed a more left
lateralized network and recruited more subcortical regions when reading in English compared to when reading in Spanish, yet these findings did not hold when corrected for multiple comparisons (FWER-corrected at $p<0.05$). Similarly, Malins and colleagues (2018) found no group differences in neural correlates of word reading in English when comparing Spanish-English bilingual children, ages 8 to 15, and their monolingual English-speaking peers. They did, however, find increased variability in the neural response to print in the right middle frontal gyrus among bilingual versus monolingual participants. These findings suggest that the regions recruited for word reading among Spanish-English bilingual children are largely the same across both languages.

Using functional near-infrared spectroscopy (fNIRS), Jasínska and Petitto (2014) found increased activation, as well as variability, in the bilateral inferior frontal gyrus, superior temporal gyrus, and areas implicated in executive processing, when reading in English, for bilingual as compared to monolingual children, who were between the ages 6 to 10. This finding was particularly interesting because no behavioral differences in word reading (e.g., response times, accuracy) were observed between the monolingual and bilingual groups, yet the fNIRS results suggested that the neural systems that underlie such reading processes differ depending on language experience. Using fNIRS, Jasínska, Berens, Kovelman, & Petitto (2017) also found that Spanish-English bilingual children, ages 6 to 10, showed increased activation of left posterior temporal regions and decreased activation of left frontal regions during English
word reading as compared to their monolingual peers. This pattern of results stands in contrast to fMRI studies of reading in children with RD, which find that those with RD exhibit decreased activation in left temporoparietal regions and increased activation in frontal regions as compared to their TD peers (e.g., Hoeft et al., 2007). Additional research is necessary to explore whether such discrepancies in the literature have arisen due to differences in bilingual children’s prior literacy experiences (i.e., whether they have received formal literacy instruction in their native language), differences in methodological approach (i.e., fMRI vs. fNIRS), or other factors.

No study to date has explored the neural correlates of second language word reading among children who are in the earliest stages of second language learning, such as ELs. Considering that almost a quarter of school-age children in the U.S. speak a non-English language at home (U.S. Census Bureau, 2019), and 10 percent of these students are designated as ELs due to a lack of English proficiency (McFarland et al., 2019), inclusion of bilingual children in neuroimaging research is necessary to extend our understanding of reading processes to all children within U.S. classrooms.

The Present Study

In an effort to reflect the diverse body of students present in U.S. classrooms, the current study compares both behavioral and neural measures of reading, as well as their correlations, among adolescents with different language and literacy backgrounds. I chose to explore these measures among adolescents
rather than younger children because the former better tolerate MR imaging and because I could be more confident regarding RD status among older children. Specifically, this study seeks to address the following research questions:

1. How do behavioral reading profiles differ among typically-developing adolescent English monolinguals (ML), Spanish-English bilinguals (BL), Spanish-speaking English Learners (EL), and English monolinguals with reading disabilities (RD)?

2. What are the shared and distinct neural correlates of speech perception and component reading processes in English among these groups?

3. How are neural and behavioral measures of word reading in English correlated among these adolescents?

Consistent with prior research (Kudo, Lussier, & Swanson, 2015), I expect students with RD to perform significantly lower than their typically-developing monolingual peers on all behavioral measures. Given the nature of their diagnosis, I expect that a large majority of participants with RD will perform in the below average range (SS<85) on word-reading and reading-related measures, including phonological awareness and rapid naming. These students may also score in the below average range on the reading comprehension measure as a result of their word-level reading difficulties. I anticipate that the EL group will score substantially lower on all behavioral measures as compared to the ML group, as these participants are similar to those studied by Pasquarella and colleagues (2012). Consistent with those results, I would anticipate a majority of
students in the EL group to score in the “below average” range on all measures except for the WMLS-RNU, which includes bilinguals and ELs in its norming sample. Therefore, I anticipate that the RD and EL groups will demonstrate similar behavioral reading difficulties, on average. The BL group is comprised of children who entered U.S. schools as ELs, so I expect that these students will perform similarly on decoding and reading-related measures as compared to those in the ML group, but may have significantly lower vocabulary and reading comprehension scores, on average, consistent with the research on middle school performance such students (Nakamoto et al., 2007). While I anticipate that some participants in the BL group will score in the below average range on vocabulary and reading comprehension measures, I expect that the majority will score in the average range or above on such measures.

In line with prior research (Maisog, Einbinder, Flowers, Turkeltaub & Eden, 2008; Martin, Kronbichler, & Richlan, 2016; Paulesu, Danelli, & Berlingeri, 2014; Richlan, Kronbichler, & Wimmer, 2009, 2011), during the fMRI task, I hypothesize that the RD group will exhibit reduced activation in left tempoparietal (TP) and occipitoparietal (OT) regions during word and non-word reading and may also exhibit hyperactivation of the left precentral gyrus associated with spelling-sound consistency. I do not anticipate that the BL group will demonstrate neural differences in component word reading processes as compared to the ML group, based on the findings of Hernandez and colleagues (2015) and Malins and colleagues (2018), however, it is possible that widespread differences will
emerge in line with prior fNIRS research (Jasinska & Petitto, 2014; Jasinska et al., 2017). If such differences emerge, they would be expected to present in the opposite pattern as those with RD, such that the BL group would show increased activation in left posterior temporal regions and reduced activation in left frontal regions as compared to the ML group. I anticipate that the EL group may demonstrate increased bilateral activation, particularly along the dorsal pathway during word reading, and may also recruit regions of the frontostriatal network to a greater extent as compared to the ML and BL groups, due to overreliance on pathways typically used when reading in Spanish, and additional effort necessary to read in a relatively new language. Lastly, I expect to find different brain-behavior correlations among groups as the reading strategies of these children are likely to differ from each other due to the nature of their diverse language and literacy backgrounds.

While there is a substantial body of literature on the behavioral reading profile of ELs who entered U.S. schools in early elementary school (e.g., Lesaux, Rupp, & Siegel, 2007; Nakamoto, Lindsey, & Manis, 2007; Lesaux et al., 2010), few studies have explored how this may vary among ELs who enter U.S. schools during adolescence. This study advances the literature on the behavioral reading profile of ELs, while also providing comparisons between these students and monolinguals struggling with reading due to an RD. Furthermore, the assessments are selected, administered, and interpreted in a manner consistent with those used in school settings to evaluate students for suspected RDs to
explore the ways in which the reading profiles of ELs may be misinterpreted as RD due to low scores on such measures. By comparing the performance of the EL and RD groups on these measures, this study also seeks to identify potential differences in the reading profiles of these groups to provide initial suggestions as to how to distinguish between reading difficulty and RD among ELs in practice. Furthermore, this study expands investigation of the neural correlates of component reading processes among adolescents to bilingual children to establish a neural profile of second language reading among such participants and explore how this may differ from that of monolinguals with and without RD.

**Methods**

**Participants and Procedure**

**Behavioral Sample.** Participants were recruited through Facebook advertisements and flyers to participate in a larger study on language and learning in middle school students with diverse language and literacy backgrounds. Parents interested in having their child participate first completed a 15-minute screening and informational phone call in the language of their preference (English or Spanish). Those who were still interested, and met initial eligibility requirements for the larger study, were scheduled for a three-hour-long behavioral session at the Harvard University Center for Brain Science (CBS). During this session, participants’ demographic characteristics were obtained using the Language and Social Background Questionnaire (LSBQ; adapted from Luk and Bialystok, 2013) administered to parents via a semi-structured interview when their responses are entered into the Qualtrics platform (2016 version).
Additionally, participants completed a battery of standardized assessments to characterize their non-verbal cognitive abilities as well as their language and reading skills in English and in Spanish, when applicable. Lastly, participants read passages during a 45-minute-long eye tracking session. All procedures were approved by the Harvard University Committee on the Use of Human Participants, and written informed consent was obtained from parents in the language of their preference. Child assent was obtained prior to the behavioral testing session.

Ninety-six adolescents, ages 11-15 years, who currently live in Massachusetts, participated in the behavioral session. Twenty participants were excluded from the analysis due to a failure to meet eligibility criteria for this study\(^8\), thereby yielding a total sample of seventy-six adolescents (\(M_{\text{age}} = 12.8\) years, 45 female). To be included in the behavioral sample, participants needed to achieve a standard score of 80\(^9\) or higher on the nonverbal IQ measure of the \textit{Kaufman Brief Intelligence Test, Second Edition} (KBIT-2; Kaufman & Kaufman, 2004), have a recorded behavioral session\(^10\), and meet criteria for one of the study groups\(^11\). To be included in the RD group (\(n = 26, M_{\text{age}} = 12.67, 14 \text{ female}\)),

\(^8\) The eligibility used in this study differ slightly from those initially used for the larger study.
\(^9\) Studies typically use an IQ cutoff of one standard deviation below the mean, which is equivalent to a standard score of 85 on the KBIT-2, but I chose a slightly lower cutoff of 80 because even “non-verbal” measures rely to some extent on students’ language abilities and cultural background.
\(^10\) All behavioral sessions were recorded and double-scored to ensure that test scores included in the subsequent analyses were valid and reliable.
\(^11\) Of the twenty participants who did not meet eligibility criteria, two were excluded due to having incomplete behavioral recordings; four were excluded from the RD group for not having a current diagnosis or IEP; five were excluded from the TD, BL, or EL groups due to having a family history of RD; and nine were excluded due to having additional neurological or psychological diagnoses.
participants were required to be monolingual English speakers with a current dyslexia diagnosis from a neuropsychologist or a current IEP in school for a specific learning disability in decoding or reading fluency, as reported by their parents during the screening process and on the Language and Social Background Questionnaire. These children had no known neurological or psychological conditions other than ADHD (n=11) and/or language disorders (n=3), which are both highly comorbid with RD (Boada, Willcutt, & Pennington, 2012; Margari et al., 2013; Willcutt & Pennington, 2000). Children taking medication for ADHD were asked to take a “medication holiday” on the day(s) that they participated in the study (Bledsoe, Semrud-Clikeman, & Pliszka, 2013; Langer, Benjamin, Becker, & Gaab, 2019; Lim et al., 2013), as preliminary evidence suggests that certain medications may improve performance on neuropsychological assessments (Froehlich et al., 2018). Parents and children were given the choice whether or not to comply with this suggestion, although all reportedly did so, and such sessions were scheduled over the weekend so as not to interfere with children’s school performance. To be eligible for the ML, BL, or EL groups, children needed to be typically-developing, with no known neurological or psychological conditions and no personal or immediate family history of RD, as reported by their parents. All children in the ML group (n=22, $M_{age}=12.63$, 14 female) were monolingual English speakers with minimal exposure, at most, to another language. All children in the BL (n=18, $M_{age}=12.93$, 12 female) and EL groups (n=10, $M_{age}=13.47$, 5 female) spoke Spanish as their
first language, lived in households where Spanish was the dominant language spoken, and had been identified as ELs at the beginning of their schooling in the U.S.\textsuperscript{12}

Given that previous work suggests different behavioral performance between newcomer adolescent ELs (e.g., Pasquarella et al., 2012) and those who have attended schools regularly in their immigration destination since elementary school (e.g., Nakamoto et al., 2007), I divided the bilingual participants into the EL and BL groups to compare the reading profiles of those learning conversational versus academic English, respectively. I distinguished these groups based on the age at which they were first exposed to English regularly, as this variable was most predictive of Spanish-speaking participants’ English vocabulary scores ($r=-0.81$, $p<0.001$) on the \textit{Woodcock-Munoz Language Survey-Revised Normative Update} (WMLS-RNU; Schrank, Wendling, & Alvarado, 2010). Participants who were first exposed to English at school entry or prior, at age 6 or younger ($M_{age}=3.75$ years, $SD=1.85$, $n=18$), were placed in the BL group, while those who were first exposed to English at seven or older ($M_{age}=9.90$ years, $SD=1.85$, $n=10$) were assigned to the EL group. I chose this cutoff because students who first heard English regularly at or prior to school entry often began to speak it actively, every day from the beginning of their schooling ($M_{age}=5.08$ years, $SD=1.90$), as reported by their parents, while those

\textsuperscript{12} Many participants in the BL and EL groups were first exposed to English in U.S. schools, yet some received English instruction in their home countries prior to arriving in the U.S. (e.g., through attending English-only or English-Spanish bilingual schools). I did not distinguish between these sources in my assignment of participants to groups.
who first heard it afterwards did not begin speaking it actively until the end of elementary or early middle school ($M_{age}=10.80$ years, $SD=1.23$). Furthermore, all children in the EL group moved to the U.S. from a Spanish-speaking country during upper elementary or middle school ($M_{age}=11.78$ years, $SD=1.48$), while those in the BL group were either born in the U.S. ($n=8$) or moved to the U.S. during early elementary school or before ($M_{age}=6.38$ years, $SD=2.50$). In addition to the United States, participants’ countries of origin included Colombia, El Salvador, Honduras, Mexico, Peru, Spain, and Venezuela. Children in the EL group moved to the U.S. within the past two years ($M=1.44$, $SD=1.45$), while those in the BL group who were not born in the U.S. moved here approximately six years prior, on average ($M=5.83$, $SD=2.64$), with the most recent arrival coming three years prior. Therefore, Cummins’ (1979) work would suggest that those in the EL group were likely still in the process of developing basic interpersonal communication skills in English, while those in the BL group were either English proficient or still developing academic English proficiency despite conversational fluency.

Demographic characteristics of the behavioral sample are presented in Table 2.1. Age and gender distributions were similar across groups; however, groups were not matched on socioeconomic status or race/ethnicity. The ML and RD groups were comparable to each other on these measures as were the BL and EL groups. Specifically, members of the ML and RD groups overwhelmingly identified as non-Hispanic white, Asian, or biracial, while members of the BL and
EL groups all identified as Hispanic or Latino. In addition, participants in the BL
and EL groups were more likely to receive free and/or reduced-price lunch than
those in the ML and RD groups, and also had mothers with lower levels of
educational achievement.
### Demographic Characteristics of the Behavioral Sample (N=76)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ML (n=22)</th>
<th>RD (n=26)</th>
<th>BL (n=18)</th>
<th>EL (n=10)</th>
<th>Test of Difference Among Groups</th>
<th>Post hoc tests/Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>$M=12.63$</td>
<td>$M=12.67$</td>
<td>$M=12.93$</td>
<td>$M=13.48$</td>
<td>$F(3, 72)=1.86, p=0.070$</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$SD=1.03$</td>
<td>$SD=1.03$</td>
<td>$SD=1.15$</td>
<td>$SD=0.74$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age moved to the U.S. (years)</td>
<td>$M=6.38$</td>
<td>$M=11.78$</td>
<td>$t(15)=-5.49, p&lt;0.001$</td>
<td>$EL &gt; BL$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD=2.50$</td>
<td>$SD=1.48$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years since moved to the U.S.</td>
<td>$M=5.75$</td>
<td>$M=1.47$</td>
<td>$t(15)=4.32, p&lt;0.001$</td>
<td>$BL &gt; EL$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD=2.60$</td>
<td>$SD=1.36$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of time speaking in Spanish</td>
<td>$M=44.44$</td>
<td>$M=57.60$</td>
<td>$t(26)=-2.71, p=0.010$</td>
<td>$EL &gt; BL$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vs. English</td>
<td>$SD=11.14$</td>
<td>$SD=14.23$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of time reading in Spanish</td>
<td>$M=28.61$</td>
<td>$M=43.80$</td>
<td>$Welch's t(11)=1.80, p=0.010$</td>
<td>$EL &gt; BL$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vs. English</td>
<td>$SD=10.03$</td>
<td>$SD=25.57$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender: Female</td>
<td>64%</td>
<td>54%</td>
<td>67%</td>
<td>50%</td>
<td>$\chi^2(3)=1.25, p=0.740$</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(n=14)</td>
<td>(n=14)</td>
<td>(n=12)</td>
<td>(n=5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>100%</td>
<td>100%</td>
<td>44%</td>
<td>0%</td>
<td>$\chi^2(3)=53.08, p&lt;0.001$</td>
<td>ML, RD &gt; BL, EL; BL &gt; EL</td>
</tr>
<tr>
<td></td>
<td>(n=22)</td>
<td>(n=26)</td>
<td>(n=8)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receives Free or Reduced Lunch</td>
<td>5%</td>
<td>4%</td>
<td>67%</td>
<td>70%</td>
<td>$\chi^2(3)=35.91, p&lt;0.001$</td>
<td>BL, EL &gt; ML, RD</td>
</tr>
<tr>
<td></td>
<td>(n=1)</td>
<td>(n=1)</td>
<td>(n=12)</td>
<td>(n=7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s Highest Educational Milestone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0%</td>
<td>0%</td>
<td>11%</td>
<td>0%</td>
<td>$\chi^2(9)=26.79, p=0.002$</td>
<td>ML, RD &gt; BL, EL</td>
</tr>
<tr>
<td></td>
<td>(n=0)</td>
<td>(n=0)</td>
<td>(n=2)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>0%</td>
<td>15%</td>
<td>44%</td>
<td>56%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=0)</td>
<td>(n=4)</td>
<td>(n=8)</td>
<td>(n=5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>32%</td>
<td>27%</td>
<td>11%</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=7)</td>
<td>(n=7)</td>
<td>(n=2)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>68%</td>
<td>58%</td>
<td>33%</td>
<td>44%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=15)</td>
<td>(n=15)</td>
<td>(n=6)</td>
<td>(n=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** One-way ANOVAs with Bonferroni-corrected post-hoc t-tests were performed to explore group differences in child age. Independent samples t-tests were used when comparing solely the BL vs. EL groups. I reported Welch’s t rather than Student’s t when the homoscedasticity assumption was violated, as indicated by Levene’s Test for Equality of Variances. Chi square tests of independence were used to test for differences among groups for categorical variables. All measures were parent-reported.
**MRI sample.** Participants in the behavioral sample were invited to complete a two-hour-long MRI session at the Harvard University CBS a later date if they met the CBS magnetic resonance (MR) screening requirements and demonstrated the ability to stay still and calm in the mock MRI scanner during the behavioral session. During the MRI session, participants completed an MR screening form, changed into MR-safe scrubs, became familiarized with the in-scanner task in a behavioral testing room, and completed functional and structural MR imaging. Fifty-three of these children participated in the MRI session, however seven of them (4 RD, 3 ML) were excluded from the analysis due to incomplete or unusable MRI data (e.g., due to sleeping and/or excessive motion during the in-scanner task). The final MRI sample consisted of 46 children ($M_{age}$=13 years, 25 female, 3 left-handed) comprised of the following groups: 14 ML ($M_{age}$=12.64 years, 8 female), 12 RD ($M_{age}$=12.78 years, 7 female), 12 BL ($M_{age}$=13.33 years, 7 female), and 8 EL ($M_{age}$=13.49 years, 3 female). Demographic characteristics of the MRI sample are presented in Table 2.2. Comparisons among the groups in the MRI sample were similar to those among groups in the behavioral sample.
Table 2.2.

Demographic Characteristics of the fMRI Sample (N=46)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ML</th>
<th>RD</th>
<th>BL</th>
<th>EL</th>
<th>Test of Difference Among Groups</th>
<th>Post hoc tests/ Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>M=12.64</td>
<td>M=12.78</td>
<td>M=13.33</td>
<td>M=13.49</td>
<td>F (3, 42)=1.80, p=0.160</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>SD=1.01</td>
<td>SD=1.04</td>
<td>SD=1.11</td>
<td>SD=0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age moved to the U.S. (years)</td>
<td>M=6.50</td>
<td>M=11.88</td>
<td>t(12)=-5.30, p&lt;0.001</td>
<td>EL &gt; BL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD=2.26</td>
<td>SD=1.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years since moved to the U.S.</td>
<td>M=5.83</td>
<td>M=1.44</td>
<td>t(12)=4.01, p=0.002</td>
<td>BL &gt; EL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD=2.64</td>
<td>SD=1.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of time speaking in Spanish vs. English</td>
<td>M=44.33</td>
<td>M=59.00</td>
<td>t(18)=-2.41, p=0.027</td>
<td>EL &gt; BL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD=11.54</td>
<td>SD=15.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of time reading in Spanish vs. English</td>
<td>M=29.08</td>
<td>M=47.25</td>
<td>Welch’s t(10)=1.91, p=0.087</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD=11.67</td>
<td>SD=25.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender: Female</td>
<td>57%</td>
<td>58%</td>
<td>58%</td>
<td>38%</td>
<td>(\chi^2(3)=1.11, p=0.770)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(n=8)</td>
<td>(n=7)</td>
<td>(n=7)</td>
<td>(n=3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>100%</td>
<td>100%</td>
<td>50%</td>
<td>0%</td>
<td>(\chi^2(3)=13.93, p=0.001)</td>
<td>ML, RD &gt; BL, EL; BL &gt; EL</td>
</tr>
<tr>
<td></td>
<td>(n=14)</td>
<td>(n=12)</td>
<td>(n=5)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receives Free or Reduced Lunch</td>
<td>0%</td>
<td>8%</td>
<td>75%</td>
<td>63%</td>
<td>(\chi^2(3)=23.06, p&lt;0.001)</td>
<td>BL, EL &gt; ML, RD</td>
</tr>
<tr>
<td></td>
<td>(n=0)</td>
<td>(n=1)</td>
<td>(n=9)</td>
<td>(n=5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s Highest Educational Milestone</td>
<td>High School</td>
<td>0%</td>
<td>17%</td>
<td>0%</td>
<td>(\chi^2(9)=27.89, p=0.001)</td>
<td>ML, RD &gt; BL, EL</td>
</tr>
<tr>
<td></td>
<td>(n=0)</td>
<td>(n=2)</td>
<td>(n=0)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>0%</td>
<td>8%</td>
<td>50%</td>
<td>43%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=0)</td>
<td>(n=1)</td>
<td>(n=6)</td>
<td>(n=3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>14%</td>
<td>50%</td>
<td>8%</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=2)</td>
<td>(n=6)</td>
<td>(n=1)</td>
<td>(n=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>86%</td>
<td>42%</td>
<td>25%</td>
<td>57%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>(n=12)</td>
<td>(n=5)</td>
<td>(n=3)</td>
<td>(n=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. One-way ANOVAs with Bonferroni-corrected post-hoc t-tests were performed to explore group differences in child age. Independent samples t-tests were used when comparing solely the BL vs. EL groups. I reported Welch’s t rather than Student’s t when the homoscedasticity assumption was violated, as indicated by Levene’s Test for Equality of Variances. Chi square tests of independence were used to test for differences among groups for categorical variables. All measures were parent-reported.
**Behavioral Measures**

The assessments administered during the behavioral session included the Matrix Reasoning subtest of the KBIT-2 (Kaufman & Kaufman, 2004); the 3-set Letters, Numbers, and Colors subtest of the *Rapid Automatized Naming and Rapid Alternating Stimulus Tests* (RAN/RAS; Wolf & Denckla, 2005); the Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests of the *Test of Word Reading Efficiency, Second Edition* (TOWRE-2; Torgesen, Wagner, & Rashotte, 2012); the Elision, Blending Words, Phoneme Isolation, Memory for Digits, and Nonword repetition subtests from the *Comprehensive Test of Phonological Processing, Second Edition* (CTOPP-2; Wagner, Torgesen, Rashotte, & Pearson, 1999); and the Picture Vocabulary, Letter-Word Identification, and Passage Comprehension subtests in English and in Spanish, when applicable, from the WMLS-RNU (Schrank, Wendling, & Alvarado, 2010).

All standardized behavioral assessments were administered by a Spanish-English bilingual researcher with training as a reading specialist and ample experience conducting neuropsychological assessments in both school and research settings. Assessment instructions were relayed to Spanish-speaking participants in their preferred language. All assessments measured constructs solely in English except for the WMLS-RNU Spanish subtests, which participants completed in Spanish, and the RAN/RAS, which participants completed in their preferred language. Importantly, none of the measures administered were designed for use with bilinguals or included them in the norming samples except
for the WMLS-RNU, which was deemed appropriate for use with Spanish-English bilinguals and ELs (Schrank, Wendling, & Alvarado, 2010). All measures were scored in compliance with their assessment manuals with no additional accommodations made for Spanish-speaking participants, as these measures are typically used in school settings in this manner to assess all students for suspected reading difficulties regardless of their language background (Abedi, 2006).

**Behavioral Analysis**

Behavioral data were analyzed from both a clinical and statistical perspective. From a clinical perspective, data were interpreted using criteria similar to those used in school settings to determine whether participants demonstrate below average performance in any area. Scores for all measures were on the same standard scale with a mean of 100 and standard deviation of 15. In practice, different cutoff values are used by clinicians to determine whether below average performance is clinically relevant (Mellard, McKnight, & Woods, 2009; Scruggs & Mastropieri, 2002). For instance, 16 percent of questionnaire participants from Study 1 reported using a cutoff score of 1 standard deviation below the mean (SS=85), 44 percent reported using 1.5 standard deviations below (SS=78), 34 percent reported using 2 standard deviations (SS=70), and 6

---

13 The norming group of the WMLS-R assessment predominantly consisted of Spanish speakers from Mexico, Central America, and South America; and a validity study indicated that some items were more difficult for speakers of certain Spanish dialects (Cuban and Puerto Rican) on the English, but not the Spanish, subtests (Sandilos et al., 2015). In this study, however, I administered the WMLS-RNU (the normative update). No validation studies have been conducted using this version, so it is possible that Spanish dialect may a role in participants’ performance.
percent reported using a different value to determine clinically low performance when assessing a children for a suspected RD.

For the purposes of this analysis, a cutoff score below 85 (1 standard deviation below the mean) was flagged as potentially of clinical relevance since this is the least stringent cutoff reportedly used in school settings. Scores of 85 to 115 were considered to be within the average range, while those below 85 were considered to be below average, and those above 115 were considered to be above average. It is important to note that these values are based on monolingual norms for all measures except for the WMLS-RNU, so while one may interpret scores in the below average range on these measures to potentially be of clinical relevance for children in the RD group, such scores may not represent the true abilities of those in the BL and EL groups. Nevertheless, due to the lack of appropriate standardized measures available for bilingual children, some practitioners may rely on these scores and interpret them as they would with monolingual students, particularly in the case of bilingual students with strong conversational English skills (Abedi, 2006; Klingner & Harry, 2006). Therefore, all scores are interpreted using this frame of reference to explore how practitioners might interpret such students’ data in a school setting should they choose to rely on standardized measures to make eligibility decisions. In addition to exploring the percentage of students in each group who would potentially be flagged in a clinical setting on each measure due to below average performance, statistical analyses were also performed to explore group differences in average
performance on these measures. Significant differences between groups may not necessarily be of clinical relevance, particularly in cases where virtually all students are in the average and/or above average range, which is another reason why it is important to approach the data from both a clinical and statistical perspective.

Statistical analyses of the behavioral data were conducted in Stata 15 (StataCorp, 2017). To explore differences in average performance on these measures by group, the standard scores were entered into one-way ANOVAs and Bonferroni-corrected post-hoc tests were performed. Welch’s ANOVA and Games-Howell-corrected post-hoc tests were reported instead, in cases where the homoscedasticity assumption was violated, as indicated by Levene’s Test for Equality of Variances. Individual samples t-tests were used to compare the EL and BL groups on measures conducted in Spanish. Rather than reporting individual CTOPP-2 subtest scores, I reported the Phonological Awareness (PA) and Phonological Memory (PM) composites. Because the behavioral battery was much longer for Spanish-speaking participants, some participants in the BL and EL groups, particularly those with stronger literacy skills, did not have sufficient time to complete the CTOPP-2. Complete data were obtained for the remaining measures, although one outlier, an extremely low English vocabulary score (SS=16) on the WMLS-RNU, was removed for an EL participant, as the ANOVA violated both the normality and homoscedasticity of residuals assumptions upon its inclusion. Regression analyses were also conducted exploring whether the
significant group differences in performance on reading and language measures held when controlling for non-verbal IQ, gender, and maternal education\textsuperscript{14}. The raw group differences were prioritized over those including controls because the KBIT-2 Matrices may not have been a valid indicator of non-verbal IQ for those in the BL and EL groups, Groups were roughly matched on gender, and the differences in maternal education observed in the sample may be reflective of those in the population, consistent with prior findings (e.g., Crosnoe, Ansari, Purtell, & Wu, 2016). Taken together, these clinical and statistical analyses enable comparison of these groups’ average behavioral reading profiles, the relative strengths and weaknesses in performance on cognitive, language, and reading measures.

**MRI Data Acquisition**

MRI data were acquired on a 3T Siemens MAGNETOM Prisma Scanner (Siemens Medical Systems, Erlangen, Germany) using a 32-channel head coil. The scanner was equipped with a built-in MR-compatible EyeLink 1000 eye tracker that was used to monitor participants’ eyes throughout the session to monitor if the participants fell asleep. Participants were outfitted with fMRI-Compatible Insert Earphones from Sensimetrics Corporation to enable high-quality auditory presentation during scanning while protecting participants’ hearing. A high resolution, a whole-brain T1-weighted structural scan with online motion-correction (Tisdall et al., 2012) was acquired for anatomical coregistration.

\textsuperscript{14} Age was not controlled for because it is already taken into consideration in the calculation of the child’s standard scores.
using a MEMPRAGE sequence (repetition time [TR]= 2260 ms; echo times [TE]=1.69 ms, 3.55 ms, 5.41 ms, 7.27 ms; inversion time [TI]=1100 ms; flip angle=7°; voxel size=1-mm³ isotropic; 256x256 matrix, 176 slices). Whole brain images were acquired during the functional task using T2*-weighted, multiband echo planar imaging (TR=2,000 ms, TE=31 ms, multiband acceleration factor=3, flip angle=80°, bandwidth=2076 Hz/Px, echo spacing=0.59 ms, voxel size=2.3 mm isotropic, Field of View [FoV]=198 mm, 66 slices). This sequence was developed at the Center for Magnetic Resonance Research (CMRR) at the University of Minnesota (Xu et al., 2013).

The functional task was adapted from the “fast localizer” paradigm developed by Malins and colleagues (2016) to identify brain regions sensitive to different component processes of reading in English. During the functional runs, participants passively attended to trials consisting of four monosyllabic written or spoken items presented sequentially. Each run consisted of 48 trials (8 conditions, 6 trials each). Each item during the written trials was presented on the screen for 300 ms with a 200 ms interstimulus interval, while spoken trials had a stimulus-onset asynchrony of 800 ms. Trials were presented in an event-related manner with each onset jittering between 4 and 13 seconds (see Figure 2.1 for diagram). The four written word conditions consisted of semantically

15 Participants also completed additional scans during the MRI session as part of the larger study on language and learning in middle schoolers, however those results will be presented in a subsequent publication.

16 The presentation time was increased by 50 ms as compared to Malins and colleagues’ study because children, particularly those with RD, need a longer time to read the words as compared to adults.
related words (e.g., heart, lung, brain, skin), words that shared orthography and phonology (e.g., fake, make, shake, take), words that shared orthography but differed in phonology (e.g., bow, cow, how, tow), and unrelated words (e.g., list, beach, dark, shape). The remaining conditions consisted of visually-presented nonwords (e.g., bliff, caft, lish, tros) and false font (e.g., ●□○) as well as spoken unrelated words (e.g., cat, song, drill, love) and vocoded speech using the same items (referred to throughout the manuscript as noise). All words were considered to be common English terms likely familiar to middle school students, and all non-words followed the phonotactic rules of English. The stimuli developed by Malins and colleagues\(^\text{17}\) were used, and the experiment was programmed in PsychoPy (Peirce, 2007). Participants completed three runs\(^\text{18}\) that were each 5.5-minutes (158 volumes) long. This yielded a total scanning time of 16.5 minutes for the complete task excluding breaks between runs.

**Figure 2.1.** A sample trial sequence illustrating key features of the “fast localizer” design. The length of visual presentation was increased from 250 to 300 ms per word in order to ensure that the adolescents had sufficient time to read the words. Reprinted from “Dough, tough, cough, rough: A “fast” fMRI localizer of component processes in reading” by J. G. Malins and colleagues, 2016, *Neuropsychologia*, 91, pp. 394-406. Copyright 2016 by Elsevier. Reprinted with permission.

---

\(^\text{17}\) Stimuli were downloaded from the supplementary material associated with the online version of Malins and colleagues (2016) at [http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.027](http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.027).

\(^\text{18}\) The original study contained an additional run, however Dr. Luk and I believed that this would be too long to sustain children’s attention. Malins indicated that his team updated the paradigm for use with children and obtained sufficient data with two to four runs of usable data (Arrington et al., 2019).
Participants were first familiarized with the localizer task outside of the scan room using different stimuli from the test items. Once they were settled in the scanner, participants were told that their eyes would be monitored throughout the session, and they were asked to stay as still as possible during all scans. Instructions were given in the child’s preferred language (i.e., either English or Spanish). The scan session began with two brief localizer scans prior to the T1 anatomical scan, which lasted approximately 4 minutes. Participants were encouraged to take a “wiggle break” between after this scan and between all subsequent scans. Next, participants were asked to keep their eyes open during a 6-minute-long resting state scan. This was followed by a brief sound test to ensure that participants would be able to hear the stimuli. Directly prior to the fMRI task, participants were instructed to read the words and pay attention to the best of their abilities. They were told that a researcher would briefly test their memory of some of the words after they got out of the scanner. Throughout the fMRI task, participants’ eyes were monitored to ensure that they were attending to the stimuli. Between runs, participants were encouraged over the intercom system to continue paying attention and staying still during the task. Once all three runs of the task were completed, participants completed two more scans as part of the larger study for an additional 12 minutes of imaging time. After exiting the scanner, participants completed a brief 8-item recognition memory task, however results were poor on this task (less than 50% correct across all groups), likely due to the high number of items during the task (576 items) and length of
time between the task and assessment. For this reason, the results of the out-of-scanner recognition memory task are not reported in this paper. I believe, however, that this task served its purpose in motivating the participants whose data were successfully obtained to pay attention and read the words.

**MRI Data Analysis**

MRI data were analyzed using AFNI (Cox, 1996). Data were preprocessed and entered into single subject analyses using the `afni_proc.py` program. The functional scans were reconstructed (`to3d`); the first 5 volumes were removed from each run as these were included to ensure stabilization of the magnetic field (`3dTcat`); slice time correction was implemented (`3dTshift`); functional runs were aligned to the T1 structural image (`3dAllineate`), motion-corrected using a six-parameter rigid-body transform (`3dvolreg`), and normalized to the Haskins Pediatric Atlas\(^{19}\) (Molfese et al., 2015) in Talairach space (@auto_tlrc) in a single transformation that resulted in a 2 mm isotropic voxel size; and smoothing was applied using a 6-mm FWHM Gaussian kernel (`3dmerge`). Single subject analyses were conducted using GLM analysis with a gamma basis function (`3dDeconvolve`). Regressors were included for each of the eight conditions and six general linear tests (GLTs) were performed: unrelated vs. semantically-related words (UNREL-SEM), unrelated words vs. pseudowords (UNREL-PSW), unrelated words vs. false font (UNREL-FF), speech vs. noise (SPEECH-NOISE), phonologically-unrelated versus phonologically-related words

\(^{19}\) The Haskins Pediatric Atlas was developed using T1 images from 72 children ages 7-14 years.
with shared orthography (O+P- minus O+P+), and an average of all written words and non-words vs. speech (WORDS-SPEECH). The first five planned contrasts were generated to illustrate regions that respond to semantic similarity, lexical processing, word reading, spelling-sound consistency, and spoken language processing, respectively. The sixth contrast was generated to enable post-hoc analysis beyond the planned contrasts from the original study to explain the results of the Group x Condition Interaction. Nuisance regressors were included for the six motion parameters and third order polynomial drift terms. Volumes that exceeded the thresholds of 0.5 mm timepoint-to-timepoint movement or 20% outliers within the automasked brain were censored.

To explore similarities and differences in neural responses to the various contrasts of interest by group, single subject analyses were entered into a whole-brain multivariate analysis (3dMVM) that tested for main effects of group (BL, EL, ML, RD), condition (6 visual and 2 auditory conditions), and their interaction. A voxelwise threshold of $p=0.001$, cluster-corrected at $p=0.05$ was used to threshold the resulting statistical maps (3dClustSim). This process used Monte Carlo simulations performed on all voxels within the Haskins Pediatric template brain using the average of the spherical autocorrelation values from each participant’s error time series (as output by 3dFWHMx option -acf) as recommended by Eklund, Nichols, and Knutsson (2016) to reduce the likelihood of Type I error when correcting for multiple comparisons in AFNI. The resulting cluster threshold for a corrected alpha level of 0.05 with a voxelwise $p<0.001$ was
90 voxels with second nearest neighbor (faces or edges of voxels must touch to be clustered together) and bi-sided (positive and negative values clustered separately above the two-sided threshold) clustering.

The Main Effect of Condition and Group x Condition Interaction yielded statistically significant results, however the Main Effect of Group was not significant. I further explored the Main Effect of Condition by performing post-hoc t-tests of the five planned contrasts described above (i.e., SPEECH-NOISE, UNREL-FF, UNREL-PSW, UNREL-SEM, O+P- minus O+P+) and reporting the regions that responded selectively to each of these basic word-reading processes. These results illustrate the shared regions implicated in each of these component processes of reading across all four groups of adolescents. Next, I performed post-hoc GLTs exploring differences in neural responses between groups (6 comparisons) for each of the five planned contrasts of interest as well as a written vs spoken language comparison (WORDS-SPEECH). After determining the comparisons that yielded these significant group differences, I extracted the average difference in beta values per participant for each contrast of interest from each of the significant clusters (3dmaskave) and plotted them by group to further illustrate the results. I also performed sensitivity analysis exploring whether these group differences held when controlling for non-verbal IQ (KBIT-2 Matrices), age, and gender\textsuperscript{20}. The same regions were identified, and therefore I present the results of the simpler model, without controls for the sake

\textsuperscript{20} These results held when centering the continuous covariates using both across-group and within-group centering.
of parsimony. Furthermore, I did not perform additional correction for multiple testing (i.e., to account for performing post-hoc tests with 6 contrasts/group comparisons) because the family-wise error rate correction (FWER) used to threshold the resulting statistical maps was already extremely conservative, especially for whole-brain analysis with a relatively small number of participants in each group (Logan & Rowe, 2004).

I aimed to investigate the presence of such correlations across the sample as a whole as well as within-group, as the relationships between neural recruitment and behavioral performance may differ by group. Due to the relatively small number of participants in each group, clusters of voxels would have had to exceed a threshold of $r=0.90$ when conducting within-group brain-behavior correlations to reach significance at a voxelwise $p$ of 0.001 and FWER-corrected alpha of 0.05. Since this threshold was unrealistically high, I decided to relax the voxelwise $p$ value to 0.01 while maintaining a FWER-corrected alpha of 0.05. Of course, relaxation of this threshold increased the likelihood of committing Type I error, so I decided to restrict the analysis to regions determined a priori to be implicated in reading processes among children and adults in a meta-analysis conducted by Martin, Schurtz, Kronbichler, and Richlan, (2015) as well as additional regions implicated in studies comparing TD and RD children (Richlan et al., 2011) and ML vs BL children (Jasinska et al., 2017), respectively. I also added all right hemisphere homologs corresponding with left hemisphere regions. The final set of regions included the bilateral inferior and middle frontal
gyrus; superior, middle, and inferior temporal gyrus; precentral gyrus; superior temporal sulcus; supramarginal gyrus; inferior parietal lobule, and fusiform gyrus. I used the Haskins Pediatric Nonlinear Atlas in AFNI corresponding with the template used for normalization to select these regions (using wherami), combine them into a mask (using 3dcalc), resample them to match the data (using 3dresample), and generate the final small-volume corrected mask used to restrict the brain-behavior analyses (see Figure 2.2).

**Figure 2.2.** Small-volume corrected mask used to constrain brain-behavior correlation analyses to regions related to reading and language processing. Regions were specified using the Haskins Pediatric Brain Nonlinear atlas and comprised the bilateral inferior and middle frontal gyrus; superior, middle, and inferior temporal gyrus; precentral gyrus; superior temporal sulcus; supramarginal gyrus; inferior parietal lobule, and fusiform gyrus. The numbers beneath each slice indicate Talairach coordinates in the axial plane. Images are presented in radiological orientation (the left hemisphere is presented on the right side).
Across all groups and within each group, Pearson correlations were performed within the small volume corrected mask (using 3dTcorr1D) between individual’s differences in beta scores for the five planned contrasts of interest (listed above) and their performance on timed single word reading (TOWRE-2 SWE) and timed pseudoword reading (TOWRE-2 PDE). I selected these behavioral measures because they were most closely related to the neuroimaging task and because these were two of the same measures that Malins and colleagues (2016) used when performing brain-behavioral correlations when using this task with adult participants. Thresholding the resulting statistical maps at a voxelwise p<0.01 meant that clusters needed to exceed 248 voxels in size to achieve statistical significance at the 0.05 alpha level. I then extracted the average difference in betas from each significant cluster per participant for the conditions of interest and used those values to graphs the results. I inspected these graphs visually for linearity and the presence of atypical points. I also used these graphs to visually explore the nature of the brain-behavior relationships among the groups for which the relationship was not significant (in the case of within group analyses). As these analyses were exploratory in nature, I did not correct for multiple testing beyond the FWER-correction used to threshold the fMRI results.
Results

Behavioral Results

In Table 2.3, I present the average performance of groups comprising the full behavioral sample on a variety of psychometric assessments as well as the results of one-way ANOVAs and independent samples t-tests comparing unadjusted groups means.

Nonverbal IQ. Participants in the ML and BL groups performed similarly on the KBIT-2 measure of non-verbal IQ, with a standard score of 115 and 113, on average respectively. This was significantly higher than those in the RD and EL groups (ps<0.001), who comparably scored 97 and 98, on average, respectively. These results held when accounting for gender and maternal education. All participants performed in the average range on this measure except for one participant in the EL group (SS=83) and one in the RD group (SS=80).

Word Reading Measures in English. The ML and BL groups each attained a mean standard score of 112 on a measure of word reading efficiency in English (TOWRE-2 SWE), which was significantly higher than those in the RD and EL groups (ps<0.01), who scored 89 and 94 on that measure, respectively. Although the RD group scored the lowest on this measure, this group did not perform significantly lower than those in the EL group; and interestingly neither group’s mean score fell into the below average range (SS<85), although 35 percent of the RD group (n=9, SS=73-83) and 20 percent of the EL group (n=2,
SS=75-84) did so. One participant in the TD group had a standard score in the below average range on this measure (SS=80). These results held when controlling for gender and non-verbal IQ, however when maternal education was also accounted for, the EL group was not significantly lower than the BL and ML groups.

For non-word reading efficiency (TOWRE-2 PDE), both the ML and BL groups had an average standard score of approximately 105 as compared to 96 and 84 for the EL and RD groups, respectively. The EL group performed marginally lower than the ML and BL groups (ps≤0.06) on this measure, and only two participants in this group (20%, SS=79) scored in the below average range. Meanwhile, the RD group performed significantly lower than all other groups, on average (ps≤0.004), and over half of these participants (62%, n=16, SS=65-84) performed in the below average range on this measure. One participant in the BL group also performed in the below average range with a standard score of 83. When controlling for non-verbal IQ, gender, and maternal education, the difference in performance among the ML, BL, and EL groups was not statistically significant, however the RD group still performed significantly lower than all other groups on this measure when taking these factors into account.

Both the ML and BL groups outperformed the RD and EL groups on untimed letter and word identification in English (WMLS-RNU Letter-Word ID; ps≤0.001), and these results held when controlling for the aforementioned variables. Nevertheless, all participants scored in the average range or higher on
this measure except for one participant in the RD group (SS=82). When looking across all three measures of word reading, only one participant in the ML group and one participant in the BL group scored in the below average range (SS≤85) on at least one of these measures, while 38% percent of the EL group (n=3) and 65% of the RD group (n=17), did so, respectively.

**Reading-Related Processes in English.** The ML, BL, and EL groups all performed similarly on the rapid naming measure (RAN/RAS 3-set), on average, which was conducted in their preferred language, yet the RD group scored significantly lower than the other groups (ps<0.05), even when controlling for non-verbal IQ, gender, and maternal education. All students in the ML, BL, and EL groups scored in the average or above average range on this measure, while 15 percent of the RD group scored in the below average range (n=4, SS=76-80).

Conversely, the EL group scored significantly lower than all other groups on the phonological awareness composite (CTOPP-2 PA), on average (ps<0.01), and this result held when controlling for the aforementioned nuisance variables. The RD group scored lower than both the ML and BL groups on this measure (ps<0.01), who performed comparably to each other. Interestingly, the RD group outperformed the EL group, on average (p=0.002), and only 19% of participants in the RD group (n=5, SS=73-80) scored in the below average range on the PA measure as opposed to 71% of the EL group (n=5, SS=58-75). One member of the BL group (SS=77) and one member of the TD group (SS=84) also scored in the below average range on this measure. Furthermore, the difference in scores
between the RD group and the ML and BL groups was not statistically significant when accounting for non-verbal IQ.

The ML group scored significantly higher on the phonological memory composite (CTOPP-2 PM) than both the RD and EL groups ($p$s<0.01), who performed similarly to each other, on average ($p$=0.39). Even so, as was the case with the PA composite, when accounting for non-verbal IQ, the difference in performance between the RD and ML groups on the PM composite was not statistically significant, however results did not change for the EL vs. ML comparison. The BL group did not score significantly different from any other groups on this measure. One participant in the ML group (SS=73) and one participant in the BL group (SS=73) scored in the below average range on this measure, while 50 percent of the EL group ($n=3$, SS=67-79) and 15 percent of the RD group ($n=4$, SS=70-82) did so. These results indicate that the RD and EL groups both received lower scores than the other groups on measures of phonological processing in English, but only members of the RD group demonstrated rapid naming difficulties, possibly because participants in the EL and BL groups were allowed to complete this measure in their language of preference.

**Vocabulary and Reading Comprehension in English.** The ML group performed a standard deviation above the mean on the WMLS-RNU English vocabulary measure (SS=115) and outperformed those in the BL and RD groups ($p$s<0.001) who scored 97 and 104, respectively, on average. When accounting
for non-verbal IQ, the difference between the BL and ML groups did hold, and the BL group had similar mean performance as the RD group ($p=0.12$). The EL group, however, scored significantly and substantially lower than all other groups on English vocabulary ($ps<0.001$), with a mean standard score of 59, over 2.5 standard deviations below the mean even after removing the lowest score in the EL group (SS=16) from this analysis due to its influence as an outlier. No participants in the TD or RD groups scored in the below average range on this measure, yet one participant in the BL group (SS=81) and 89 percent of the EL group ($n=8$, SS=16-64) did.

On the WMLS-RNU English reading comprehension measure, group mean scores were all significantly different from each other ($ps<0.05$) with the ML group (SS=108) outperforming the BL group (SS=98), who in turn scored higher than the RD group (SS=90); and all groups scored substantially higher than the EL group (SS=57). When controlling for non-verbal IQ, gender, and maternal education, however, the differences in average reading comprehension scores among the ML, BL, and RD groups were not statistically significant. Nevertheless, the EL group’s performance was still significantly lower on this measure than all other groups ($ps<0.001$), with a mean score over 2.5 standard deviations below that of the BL group. One participant in the BL group performed in the below average range on this measure (SS=83), while 23 percent of the RD group ($n=6$, SS=58-84), and 90 percent of the EL group ($n=9$, SS=27-74) did so.
**Spanish Measures.** The EL group outperformed the BL group on all of the Spanish measures ($p<0.05$) when controlling for maternal education, gender, and non-verbal IQ. This makes sense given that many participants in the BL group received little to no school-based instruction in Spanish, while all members of the EL group received multiple years of instruction in Spanish. For instance, the BL group scored over a standard deviation below the mean on the Spanish vocabulary measure ($SS=82$), on average, while ELs’ mean score was in the average range ($SS=94$) possibly because many of the assessment items involved terms often used in the academic as opposed to the home context (e.g., ruler). One participant in the EL group scored in the below average range on this measure ($SS=71$), while 67 percent of the BL group ($n=12, SS=61-83$) did so.

Both the BL and EL groups scored over a standard deviation above the mean on the Spanish letter-word identification measure (121 and 137, respectively), with no participants scoring in the below average range. The average reading comprehension score in Spanish, however, was in the below average range for the BL group ($SS=83$) and in the average range for the EL group ($SS=93$). Twenty percent of the EL group performed in the below average range on this measure ($n=2, SS=72-74$), while 56 percent of the BL group did so ($n=10, SS=58-83$). Although participants in the BL group have seemingly sufficient word-level reading skills in Spanish, many seemingly lack the academic vocabulary necessary to comprehend Spanish passages.
Table 2.3.

Performance on Psychometric Tests of IQ, Reading, and Language Presented by Group for the Full Behavioral Sample (N=76)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>(Min-Max)</th>
<th>n</th>
<th>Test of Difference Among Groups</th>
<th>Post hoc tests/Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>KBIT-2 Matrices</td>
<td>Nonverbal IQ</td>
<td>ML</td>
<td>114.68</td>
<td>11.00</td>
<td>(90-130)</td>
<td>22</td>
<td>ML, BL &gt; RD, EL</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>96.61</td>
<td>8.56</td>
<td>(80-121)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>112.78</td>
<td>8.63</td>
<td>(99-125)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>98.20</td>
<td>10.05</td>
<td>(83-117)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWRE-2 Sight Word Efficiency</td>
<td>Word reading efficiency in English</td>
<td>ML</td>
<td>111.68</td>
<td>14.46</td>
<td>(80-136)</td>
<td>22</td>
<td>F (3, 72)=19.73, p&lt;0.001</td>
<td>ML, BL &gt; RD, EL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>88.50</td>
<td>9.99</td>
<td>(73-108)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>112.06</td>
<td>10.94</td>
<td>(102-142)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>94.40</td>
<td>10.66</td>
<td>(75-108)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWRE-2 Phonemic Decoding Efficiency</td>
<td>Non-word reading efficiency (English)</td>
<td>ML</td>
<td>105.50</td>
<td>8.09</td>
<td>(89-123)</td>
<td>22</td>
<td>F (3, 72)=30.00, p&lt;0.001</td>
<td>ML &gt; RD, EL, BL, EL &gt; RD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>84.00</td>
<td>7.95</td>
<td>(65-101)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>105.39</td>
<td>10.54</td>
<td>(83-131)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>95.90</td>
<td>10.34</td>
<td>(79-109)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAN/RAS 3-set</td>
<td>Rapid naming of letters, numbers, and colors (in preferred language)</td>
<td>ML</td>
<td>111.14</td>
<td>12.60</td>
<td>(87-133)</td>
<td>22</td>
<td>F (3, 72)=10.93, p&lt;0.001</td>
<td>ML, BL, EL &gt; RD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>95.54</td>
<td>11.16</td>
<td>(76-125)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>113.11</td>
<td>11.70</td>
<td>(94-135)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>108.20</td>
<td>9.73</td>
<td>(95-121)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTOPP-2 Phonological Awareness</td>
<td>Composite of elision, blending words, and phoneme isolation subtests (in English)</td>
<td>ML</td>
<td>104.41</td>
<td>11.38</td>
<td>(84-125)</td>
<td>22</td>
<td>F (3, 64)=12.77, p&lt;0.001</td>
<td>ML, BL &gt; RD, EL, RD &gt; EL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>93.19</td>
<td>11.84</td>
<td>(73-114)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>97.46</td>
<td>9.42</td>
<td>(77-114)</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>74.14</td>
<td>14.58</td>
<td>(58-98)</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTOPP-2 Phonological Memory</td>
<td>Composite of memory for digits and non-word repetition subtests (in English)</td>
<td>ML</td>
<td>107.73</td>
<td>14.27</td>
<td>(73-134)</td>
<td>22</td>
<td>F (3, 61)=8.33, p&lt;0.001</td>
<td>ML &gt; RD, EL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>93.64</td>
<td>11.52</td>
<td>(70-116)</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>95.75</td>
<td>11.85</td>
<td>(73-116)</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>82.83</td>
<td>12.21</td>
<td>(67-98)</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>English vocabulary</td>
<td>ML</td>
<td>114.86</td>
<td>10.12</td>
<td>(100-133)</td>
<td>22</td>
<td>F (3, 71)=76.33,</td>
<td>ML &gt; BL, RD &gt; EL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU</td>
<td>Picture Vocabulary</td>
<td>RD</td>
<td>104.12</td>
<td>7.54</td>
<td>(90-118)</td>
<td>26</td>
<td>p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU Letter-Word ID</td>
<td>English letter and word identification</td>
<td>BL</td>
<td>97.17</td>
<td>8.45</td>
<td>(81-110)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU Passage Comprehension</td>
<td>English reading comprehension</td>
<td>EL</td>
<td>58.78</td>
<td>14.33</td>
<td>(35-88)</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU Vocabulario Sobre Dibujos</td>
<td>Spanish vocabulary</td>
<td>ML</td>
<td>117.45</td>
<td>11.55</td>
<td>(99-140)</td>
<td>22</td>
<td>Welch’s F (3, 29.01)=39.15, 15, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>93.65</td>
<td>5.59</td>
<td>(82-109)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>113.50</td>
<td>9.17</td>
<td>(99-135)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>101.60</td>
<td>8.10</td>
<td>(91-114)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU Identificación de letras y palabras</td>
<td>Spanish letter and word identification</td>
<td>ML</td>
<td>108.32</td>
<td>15.76</td>
<td>(85-138)</td>
<td>22</td>
<td>Welch’s F (3, 29.20)=21.90, 90, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RD</td>
<td>90.19</td>
<td>11.14</td>
<td>(58-117)</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>97.56</td>
<td>6.47</td>
<td>(83-109)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>57.20</td>
<td>17.75</td>
<td>(27-86)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMLS-RNU Comprensión de textos</td>
<td>Spanish reading comprehension</td>
<td>ML</td>
<td>121.06</td>
<td>15.01</td>
<td>(96-150)</td>
<td>18</td>
<td>t(26)=-2.10, p=0.045</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>82.28</td>
<td>11.13</td>
<td>(61-104)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>94.20</td>
<td>10.17</td>
<td>(71-106)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>137.00</td>
<td>14.46</td>
<td>(118-164)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>83.00</td>
<td>11.63</td>
<td>(58-102)</td>
<td>18</td>
<td>t(26)=-2.73, p=0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL</td>
<td>93.30</td>
<td>13.81</td>
<td>(72-115)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. One-way ANOVAs and Bonferroni-corrected post-hoc tests were used for continuous variables when comparing more than two groups. Welch’s ANOVA and Games-Howell-corrected post-hoc tests are presented instead, in cases where the homoscedasticity assumption was violated, as indicated by Levene’s Test for Equality of Variances. Independent samples t-tests were used when comparing solely the BL vs. EL groups. Standard scores are reported for all measures.*

**Neural Correlates of Component Reading Processes**

In Table 2.4, I present the results of the multivariate model exploring similarities and differences in the neural correlates of component reading processes among adolescents in the MRI sample. These results are also illustrated in Figures 3 and 4.
**Table 2.4.**

*Clusters Showing a Main Effect of Condition with Significant Post-hoc t-tests for Planned Comparisons (t=3.504) and Clusters Showing a Group x Condition Interaction (F=2.327) Voxelwise p=0.001, cluster corrected at p=0.05.*

<table>
<thead>
<tr>
<th>Region</th>
<th>Talairach Coordinates of Peak Voxels</th>
<th>Maximum t/F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effect of Condition: Planned Post-hoc t-tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SPEECH &gt; NOISE (Spoken Language Processing)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Superior Temporal Gyrus, Superior Temporal Sulcus, and Middle Temporal Gyrus</td>
<td>-64</td>
<td>-14</td>
<td>7</td>
</tr>
<tr>
<td>R Superior Temporal Gyrus, Superior Temporal Sulcus, and Middle Temporal Gyrus</td>
<td>66</td>
<td>-14</td>
<td>7</td>
</tr>
<tr>
<td><strong>FF &gt; UNREL (Visual Symbol Processing)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Fusiform Gyrus and Inferior Parietal Lobule</td>
<td>24</td>
<td>-64</td>
<td>51</td>
</tr>
<tr>
<td>L Fusiform Gyrus and Inferior Parietal Lobule</td>
<td>-35</td>
<td>-88</td>
<td>12</td>
</tr>
<tr>
<td><strong>UNREL &gt; FF (Word Reading)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Superior Temporal Gyrus, Superior Temporal Sulcus, and Middle Temporal Gyrus</td>
<td>-66</td>
<td>-32</td>
<td>6</td>
</tr>
<tr>
<td>L Inferior Frontal Gyrus (Pars Opercularis)</td>
<td>-52</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td>R Cerebellum</td>
<td>22</td>
<td>-73</td>
<td>-40</td>
</tr>
<tr>
<td>R Cuneus</td>
<td>12</td>
<td>-96</td>
<td>23</td>
</tr>
<tr>
<td>L Precentral Gyrus</td>
<td>-54</td>
<td>-4</td>
<td>44</td>
</tr>
<tr>
<td>R Superior Temporal Gyrus</td>
<td>58</td>
<td>-31</td>
<td>4</td>
</tr>
<tr>
<td>L Fusiform Gyrus and Inferior Temporal Gyrus (VWFA)</td>
<td>-47</td>
<td>-56</td>
<td>-15</td>
</tr>
<tr>
<td><strong>UNREL-PSW (Lexical Processing)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Superior Temporal Sulcus</td>
<td>-62</td>
<td>-44</td>
<td>12</td>
</tr>
<tr>
<td>L Middle Temporal Gyrus</td>
<td>-68</td>
<td>-28</td>
<td>1</td>
</tr>
<tr>
<td>L Parahippocampal Gyrus</td>
<td>-28</td>
<td>-35</td>
<td>-10</td>
</tr>
<tr>
<td>L Inferior Parietal Lobule</td>
<td>-35</td>
<td>-77</td>
<td>39</td>
</tr>
<tr>
<td>L/R Supplementary Motor Area</td>
<td>1</td>
<td>-7</td>
<td>52</td>
</tr>
<tr>
<td>R Supramarginal Gyrus</td>
<td>52</td>
<td>-27</td>
<td>28</td>
</tr>
<tr>
<td><strong>UNREL &gt; SEM (Semantic Similarity)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Rostral Middle Frontal Gyrus</td>
<td>-35</td>
<td>59</td>
<td>8</td>
</tr>
<tr>
<td>R Rostral Middle Frontal Gyrus</td>
<td>33</td>
<td>64</td>
<td>1</td>
</tr>
<tr>
<td><strong>O+P- &gt; O+P+ (Spelling-Sound Consistency)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Precentral and Caudal Middle Frontal Gyrus</td>
<td>-52</td>
<td>12</td>
<td>37</td>
</tr>
<tr>
<td><strong>Group x Condition Interaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Fusiform Gyrus and Lateral Occipital Cortex</td>
<td>-39</td>
<td>-81</td>
<td>-4</td>
</tr>
<tr>
<td>L Supramarginal Gyrus</td>
<td>-43</td>
<td>-40</td>
<td>37</td>
</tr>
<tr>
<td>R Superior Temporal Gyrus</td>
<td>66</td>
<td>-28</td>
<td>12</td>
</tr>
</tbody>
</table>

Voxels are 2 mm isotropic (8 cubic mm) in size, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001
Main Effect of Condition. The main effect of condition explored if regions differed in neural activity among any of the pairwise comparisons, accounting for group. As might be expected, this yielded a multitude of statistically significant regions across the entire brain and was not readily interpretable. To account for this, I followed up with pairwise t-tests of the five condition comparisons that this task was developed to explore (Malins et al., 2016). Significant regions were identified for all five of the comparisons of interest (100%), which is higher than would be expected by chance (5%). These results illustrate the neural regions that support speech and various component reading processes across adolescents in the MRI sample, controlling for group membership.

As illustrated in Figure 2.3, bilateral temporal clusters responded significantly more to speech than vocoded speech, accounting for group, confirming that these regions are involved in processing spoken language. Similarly, bilateral clusters composed of the fusiform gyrus and parts of the inferior parietal lobule were more responsive to false font than unrelated words, suggesting that these regions support visual processing of non-alphabetic symbols.
Figure 2.3. Clusters that showed a main effect of stimulus type across the eight conditions and significant post-hoc pairwise t-tests for spoken language processing or visual processing (see key above). Please refer to Table 2.4 for details. The numbers beneath each slice indicate Talairach coordinates in the axial plane. Images are presented in radiological orientation (the left hemisphere is presented on the right side).

Of particular interest are the regions that are associated with word reading and component reading processes. In Figure 2.4, the green clusters demonstrate the regions that were more responsive to unrelated words than false font, across participants, which included a large left temporal cluster incorporating the superior and middle temporal gyrus and superior temporal sulcus, and smaller clusters within the right superior temporal gyrus, left inferior frontal gyrus (pars opercularis), left precentral gyrus, right cuneus, right cerebellum, and left lateral
fusiform and inferior temporal gyrus (VWFA). Several regions were also elicited when comparing unrelated words to pseudowords, illustrated in purple. Clusters in the left superior temporal sulcus, middle temporal gyrus, parahippocampal gyrus, and inferior parietal lobule as well as the supplementary motor area and right supramarginal gyrus responded significantly more to unrelated words than to pseudowords, accounting for group. A portion of the left middle temporal gyrus was the only region that responded selectively to unrelated words as compared to both false font and pseudowords. No regions responded significantly more to pseudowords or semantically-related words as compared to unrelated words. Bilateral clusters in the rostral middle frontal gyrus did respond significantly more to unrelated words versus semantically related words, however. Lastly, a cluster straddling the left precentral and caudal middle frontal gyrus responded significantly more to phonologically inconsistent versus phonologically consistent with similar orthography.
Figure 2.4. Clusters that showed a main effect of stimulus type across the eight conditions and significant post-hoc pairwise t-tests for word reading, lexical processing, semantic similarity, or spelling-sound consistency (see key above). Please refer to Table 2.4 for details. The numbers beneath each slice indicate Talairach coordinates in the axial plane. Images are presented in radiological orientation (the left hemisphere is presented on the right side).

**Group x Condition Interaction.** In addition to exploring the regions that support speech and component reading processes across participants, I also performed a Group x Condition Interaction to explore whether the neural regions that support these processes differ by group. Only three significant clusters emerged, as illustrated in Figure 2.5. Results of post-hoc t-tests of pairwise group comparisons across the five planned contrasts and additional contrast of all
written words versus speech, indicated that the left supramarginal gyrus cluster emerged from a significant difference among groups in processing spoken language as compared to vocoded speech, while the clusters in the left fusiform gyrus and right superior temporal gyrus resulted from group differences in processing spoken versus written language. Of the 36 post-hoc pairwise group comparisons conducted in total (6 comparisons for each of the 6 contrasts), 7 were statistically significant (19%), which is greater than the number that would be expected to occur by chance (5%). Specifically, monolingual participants recruited the left supramarginal gyrus to a greater extent in processing spoken language versus vocoded speech, while the other groups did not, on average, and these differences between the ML group and all other groups were statistically significant ($p<0.01$), as illustrated in Figure 2.5. In the right superior temporal gyrus, all groups recruited the region in the posterior portion for spoken as compared with written language, however the RD and BL groups activated this area significantly more for spoken versus written language than the ML group, on average ($p<0.05$). This was driven by differences in neural activation to speech, as mean beta values in this region for written language conditions were not significantly different from zero across groups. By contrast, all groups activated a region straddling the left posterior fusiform gyrus and lateral occipital cortex more for written versus spoken language, although the ML group recruited this region significantly more than the RD and BL groups, on average ($p<0.01$). This finding was driven by differences in neural activation to print across
conditions as the mean beta values in this region for spoken language conditions were not significantly different from zero across groups.

Figure 2.5. Clusters that showed a significant Group x Condition Interaction across the four groups and eight conditions, with bar graphs illustrating post-hoc t-tests for comparisons that yielded significant differences across groups. "WORDS" refers to the average beta value of all trials involving reading (UNREL, SEM, O+P+, O+P-, and PSW). Error bars represent the 95% confidence interval of the predicted mean difference in betas for each group. Please refer to Table 2.4 for details. The numbers beneath each slice indicate Talairach coordinates in the axial plane. Images are presented in radiological orientation (the left hemisphere is presented on the right side). * p<0.05, ** p<0.01, *** p<0.001
**Brain-Behavior Correlations**

Brain-behavior correlations were performed to explore the relationship between the five planned contrasts and the TOWRE-2 SWE and PDE subtests both within and across groups limited to regions within the small-volume corrected mask. Of the 50 comparisons performed (5 groups, 5 contrasts, 2 behavioral measures), 10 were statistically significant (20%), which is above the number that would be expected by chance (5%). The full results of this analysis are summarized in Appendix B. Figure 2.6 presents the only significant correlation that emerged across the entire sample. Increased activation of the pars opercularis region of the left inferior frontal gyrus for phonologically inconsistent as compared to phonologically consistent words was positively associated with phonemic decoding efficiency ($r=0.43$, $p<0.01$). This means that participants who recruited the left pars opercularis to a greater extent for phonologically inconsistent versus consistent words scored higher on a measure of timed non-word reading. This correlation was not statistically significant at the more stringent voxelwise level of $p<0.001$, however. Two participants in the BL group had extremely high differences in beta values for this comparison, however the correlation held when excluding those participants, suggesting that this finding is fairly robust across participants. The same pattern did not hold within group, however, suggesting that this finding was driven by group differences in both mean neural response and behavior within this region. Specifically, participants in the BL and ML groups recruited this region to a greater extent for
phonologically inconsistent versus consistent words than those in the EL and RD groups and also had higher timed non-word reading scores, on average. The voxelwise significance threshold was more lenient in these correlational analyses ($p<0.01$ vs. $p<0.001$), no Bonferroni correction was performed, and the within-group analyses had particularly small sample sizes (n=8-14), so additional research is necessary to replicate these results.

![Image](image.png)

**Figure 2.6.** Significant correlation between differential in mean response to phonologically inconsistent vs. phonologically consistent words and phonemic decoding efficiency in the left pars opercularis across participants (n=48)

**Discussion**

The current study investigated behavioral reading profiles, neural correlates of component reading processes, and their correlations among adolescents with diverse language and literacy backgrounds. Performance on
non-verbal IQ, reading-related cognitive processes, word reading, vocabulary, and reading comprehension were compared among adolescents who were typically-developing English-speaking monolinguals, Spanish-English bilinguals who were ELs in early elementary school, recent immigrant Spanish-English ELs, and English-speaking monolinguals with diagnosed RDs consistent with a profile of dyslexia. Behavioral results were largely consistent with prior research, such that both children with RDs and recent immigrant ELs demonstrated evidence of reading difficulties in English. Different patterns of relative strengths and weaknesses on these measures were observed between these two groups, however. All groups generally recruited similar regions during an fMRI passive word reading task, suggesting that these adolescents shared similar neural responses to the component reading processes explored: spoken language processing, visual symbol processing, word reading, semantic-similarity, spelling-sound consistency, and lexical processing. Lastly, a number of brain-behavior correlations were identified both within and across groups. These results, particularly within groups, should be interpreted with caution due to low sample sizes (Yarkoni, 2009). Below, I discuss each of these sets of findings and their implications in turn.

**Behavioral Reading Profiles**

While participants in the typically-developing monolingual and bilingual groups generally demonstrated adequate performance on all measures, a majority of ELs and monolinguals with diagnosed RDs underperformed on a
variety of reading and reading-related measures, by comparison. On average, bilingual participants exposed to English prior to age six demonstrated strikingly similar scores as their typically-developing monolingual peers across all measures except for slightly lower vocabulary and reading comprehension. This is consistent with my hypotheses and with prior research that children who enter elementary school as ELs generally develop word reading skills on par with their English proficient peers, but have lower reading comprehension performance in middle school due, in part, to their protracted development of academic vocabulary (e.g., Nakemoto et al., 2007).

By contrast, recent immigrant ELs performed significantly lower than both their monolingual and bilingual peers without RD, on average, on virtually all measures conducted in English, including sight word efficiency, letter-word identification, phonological awareness, vocabulary, and reading comprehension. These results are consistent with the findings of Pasquarella and colleagues (2012), as the EL group scored approximately a standard deviation lower than the ML and BL groups on English measures of word reading and reading-related skills and over 2.5 standard deviations lower on vocabulary and reading comprehension measures. This suggests that, contrary to ELs who enter U.S. schools during the early elementary years, those who immigrate during upper elementary or middle school may underperform substantially on English word reading and comprehension measures in adolescence, despite having adequate reading skills in Spanish. This is likely because those in the BL group have
received literacy instruction in English for years, while those in the EL group have only attended U.S. schools for one or two years.

Since most behavioral research on reading among ELs has focused on the former group (Lesaux, Rupp, & Siegel, 2007; Nakamoto, Lindsey, & Manis, 2007; Lesaux et al., 2010), researchers have suggested that substantial word-level reading difficulties in ELs may be indicative that their reading struggles are due primarily to an underlying RD (Geva, 2000; Lesaux et al., 2007). These results suggest that this guideline for distinguishing reading difficulties from disabilities among ELs may apply in the case of adolescents who entered early elementary school as ELs, but not those who have received fewer years of schooling in the U.S. Therefore, when evaluating such students for suspected RDs, other research-informed best practices in assessment are particularly important, including providing language and literacy assessments appropriate for use with ELs in both English and the child’s native language. For instance, the CTOPP-2 is widely used to assess students for dyslexia, yet this measure has not been validated for use with ELs. Students in the BL group generally scored in the “Average” range on this assessment based on monolingual norms, while all members in the EL group except for one scored in the “Below Average” range on at least one of the two subscales. Although phonological processing was not measured in Spanish, I assume that the challenges recent immigrants faced on this measure were due mainly to aspects related to their relatively early stage of English language development (e.g., unfamiliar vocabulary, influence of accent)
rather than true deficits in these reading-related cognitive processes. Likewise, ELs demonstrated adequate performance on reading measures in Spanish as well as the rapid naming measure, which was conducted in Spanish for all but one participant in this group. This provides evidence that some ELs may underperform on English measures commonly used in schools to assess word reading due to a lack of English proficiency rather than an underlying disorder; and therefore, educators should consider the age of English exposure and immigration to the U.S. when determining how to support and evaluate adolescent ELs struggling with reading in English.

Similar to recently immigrated adolescent ELs, monolinguals currently receiving support in school for a diagnosed reading disability also performed significantly lower than their typically-developing monolingual and bilingual peers on a variety of measures, on average, including sight word efficiency, phonemic decoding efficiency, rapid naming, phonological awareness, letter-word identification, and reading comprehension, although they generally performed below these groups by about one standard deviation across all measures. Interestingly, these children performed similarly to the EL group, on average, on sight word reading efficiency, phonological memory, and letter-word identification in English, although the EL group scored significantly lower on measures of phonological awareness, vocabulary, and reading comprehension, while the RD group scored significantly lower on phonemic decoding efficiency and rapid naming.
It is intriguing that many of the participants with RD performed in the average range on reading comprehension, despite their word level difficulties; however, this measure involved adding missing words to short sentences and paragraphs in an untimed manner and may not be reflective of their reading skills under more naturalistic circumstances. To utilize another piece of evidence, I drew upon participants’ performance on the passages from the Basic Reading Inventory (Johns, 2005) that they read during eye tracking in English and answered comprehension questions about in the language of their preference. Children with RD and recent-immigrant ELs performed similarly on this measure ($p=0.95$) and answered fewer comprehension questions correctly (56% and 47%, respectively) than their typically-developing monolingual (72%) and bilingual (70%) peers ($p<0.10$). This suggests that the use of a more ecologically-valid measure of reading comprehension, such as a reading inventory, may be necessary to identify difficulties in this process among adolescents with RD.

In sum, these results suggest that the behavioral reading profile of reading difficulties among monolinguals with word-level RDs may differ slightly from that of recent immigrant ELs, on average, however there are many similarities. Given the additional variability at the individual level, it is not surprising that educators and practitioners have struggled to distinguish reading difficulties due to normal English language development, from those attributable to an underlying reading disability among ELs, based on behavioral measures alone (McCardle, Mele-McCarthy, & Leos, 2005; Pugh et al., 2005; Simos, Billingsley-Marshall, Sarkari,
Patarea, & Papanicolaou, 2005). Next, I discuss the results of an attempt to initiate such neuroimaging research on this topic using a “fast localizer” paradigm of component reading processes (Malins et al., 2016) among a subsample of participants from the behavioral sample.

**Neural Correlates of Component Reading Processes**

Despite the differences in behavioral performance on a variety of reading measures, few significant differences in neural responses to component reading processes among groups emerged. This suggests that the adolescents with diverse language and literacy backgrounds largely recruit the same brain regions when passively reading basic English monosyllabic words and non-words. The similarities in neural responses for the ML and BL groups is consistent with the findings of Hernandez and colleagues (2015) and Malins and colleagues (2018), and stands in contrast to prior fNIRS research (Jasinska & Petitto, 2014; Jasinska et al., 2017). There are a number of possibilities for the differences in results between the fMRI and fNIRS studies with this population. Participants in the fNIRS studies were asked to read the words aloud, and these studies involved slightly younger children. My stimuli and task may have been too easy and low-level for existing neural differences to emerge among adolescents, or the results may be indicative of a developmental difference. Another key distinction is that fNIRS involves surface-based analyses of signals emitting from the neurons in the cerebral cortex, but I performed traditional volume-based fMRI analysis with BOLD signal from the entire brain, including the subcortical regions.
Perhaps analysis of the fMRI data using newly improved surface-based analyses (e.g., Coalson, Van Essen, & Glasser, 2018) would yield similar results.

**Shared Neural Responses Across Groups.** Almost all identified regions associated with speech and component reading processes were not significantly different across groups. On average, participants recruited bilateral temporal regions during spoken language processing, consistent with prior research using a similar paradigm in adults across language systems (Rueckl et al., 2015). In addition, participants recruited the bilateral fusiform gyrus when processing non-alphabetic symbols as well as regions along both the dorsal and ventral streams during word reading including the left superior and middle temporal gyrus, left superior temporal sulcus, right superior temporal gyrus, left inferior frontal gyrus, left precentral gyrus (pars opercularis), VWFA, and right cerebellum. These regions are consistent with those observed in young, healthy English monolingual adults (Malins et al., 2016). Regions associated with lexicality were located in left temporoparietal regions as well as the bilateral supplementary motor area and right supramarginal gyrus. Nevertheless, this increased activation for word versus non-word reading in the left middle temporal and parahippocampal gyrus may be indicative of a task effect, as participants were asked to try to memorize the words for a brief test at the very end of the fMRI session, and activation of these regions has been reported in fMRI studies using similar tasks (Diaz & McCartney, 2009). Participants also exhibited increased activation of the bilateral middle frontal gyrus for unrelated versus
semantically-related words and the left precentral gyrus for reading similar-looking, non-rhyming words versus similar-looking, rhyming words. These regions within the frontostriatal network may have been recruited in response to the increased item difficulty or attention. These results indicate that the task was successful at identifying regions associated with these basic processes underlying reading among adolescents, but it was not particularly effective at identifying differences in the neural underpinnings of reading among groups. This task does seem well-suited to identify regions associated with component word reading processes that are similar across adolescents to be used as seeds for task-based or resting state network analyses.

**Groups Differences in Neural Response.** Prior research suggests children with RDs demonstrate hypoactivation of the left OT and TP cortices during reading and reading-related tasks (Maisog, Einbinder, Flowers, Turkeltaub & Eden, 2008; Martin, Kronbichler, & Richlan, 2016; Paulesu, Danelli, & Berlingeri, 2014; Richlan, Kronbichler, & Wimmer, 2009, 2011), however reduced activation between participants with RDs and their typically-developing monolingual peers only emerged within the left OT, specifically the posterior fusiform gyrus and lateral occipital cortex, when processing written as compared to spoken language. This region was further posterior than the purported VWFA reported in prior research; the same reduced neural response in this region was also observed among bilingual participants who exhibited reading skills on par with their monolingual peers; and similar results emerged when comparing
symbol versus spoken language processing among groups. Therefore, this task does not seem particularly well-suited to identify the neural differences associated with RDs as compared to reading difficulty experienced due primarily to normal second language development.

Furthermore, the ML group demonstrated a significantly smaller differential in neural response to spoken versus written language within the right superior temporal gyrus than those in the RD and BL groups. Interestingly, these results were driven by differences in baseline neural response to speech, and similar findings emerged for vocoded speech, suggesting that those in the RD and BL groups recruit this region to a greater extent during auditory processing as compared to their typically-developing monolingual peers, on average.

Researchers have consistently found hyperactivation of right hemisphere regions among those with RD during auditory processing and/or reading tasks, and have interpreted such findings as evidence of compensation for impaired left hemisphere language and reading networks (e.g., Pugh et al., 2000; Richlan et al., 2009; Sarkari et al., 2002; Shaywitz et al., 2002; Shaywitz & Shaywitz, 2005; Simos et al., 2000). This interpretation is consistent with the finding that those with RD also recruited the left supramarginal gyrus for spoken language significantly less than their typically-developing monolingual peers. Nevertheless, this pattern of increased neural response in the right superior temporal gyrus and decreased response in the left supramarginal gyrus to spoken language was also present among bilinguals with relatively strong language and literacy skills in
English. This suggests that there is an alternative explanation for such differences in neural response to spoken language besides a neurological impairment. The inclusion of bilinguals in future neuroimaging studies on RD may help prevent researchers from misinterpreting results as due to disability when they may be due other factors that warrant further investigation.

**Brain-Behavior Correlations**

Lastly, I performed exploratory correlations between word reading measures and neural responses associated with component reading processes both across and within groups limited to regions previously found to be involved in word reading among children. Malins and colleagues (2016) found no statistically significant correlations among TOWRE subscores and neural responses to the comparisons of interest using this task among their typical adult monolingual participants. In my study, one significant correlation emerged across participants in the left pars opercularis, such that increased activation to phonologically inconsistent versus phonologically consistent words was positively associated with phonemic decoding efficiency. This means that participants who recruited this region traditionally associated with phonological processing to a greater extent for phonologically inconsistent words during passive word reading scored higher on a measure of timed non-word reading, although this pattern did not hold within group, suggesting that the findings were largely driven by group differences that may not have survived the strict FWER-correction due to limited statistical power. Additional research is necessary to replicate the brain-behavior
correlations across- as well as within-groups as this may provide information beyond behavioral findings alone regarding the neural pathways and reading strategies that predict successful reading outcomes.

**Limitations and Future Directions**

While this study yielded many insights regarding behavioral reading profiles, neural correlates of component reading processes, and their association among adolescents from diverse language and literacy backgrounds, there were a number of key limitations. The primary limitation was the fairly low number of participants within each group in the fMRI sample. Once recruited, participants had to successfully complete a screening phone call as well as the behavioral session, meet group eligibility requirements, demonstrate the ability to stay still and calm in the mock fMRI scanner, and successfully complete a 20-minute word reading task as part of a 1.5 hour MRI session without falling asleep or moving too much. The biggest challenge in this process was due to the high prevalence of braces among this age group. To put it into perspective, my colleagues and I received over one thousand initial inquiries regarding the study, yet only approximately two hundred parents completed screening, as the others were often turned down at the outset of the phone call due to their child’s braces, retainer, or other scanner contraindications. As indicated by the initial sample size, approximately half of those screened completed the behavioral session, yet over 20% of these participants were excluded from behavioral analyses due to failure to meet eligibility requirements. These participants were largely excluded
due to an immediate family history of RD or bilingual experience with a language other than Spanish. It was also challenging to find and successfully recruit recent immigrants to participate in an fMRI study, as many members of the local Spanish-speaking community expressed hesitancy at the thought of undergoing an MRI session and others seemed distrustful of research participation in general. In total, it took over a year and a half to collect these data, and most of the recruitment success came from Facebook ads, however future studies may fare better in this process through partnering up with school districts and community centers in the area early during project development to increase buy in and hopefully generate a high response rate.

Participants had a particularly difficult time staying awake and focused during this task as compared to the other fMRI tasks likely due to its passive nature and the rhythmic presentation of the stimuli. It was thought that excluding a functional run would allow adolescents to better tolerate this task, however many still struggled to stay engaged, and the task may have been underpowered to detect group differences among more subtle comparisons (e.g., O+P- minus O+P+, SEM minus UNREL). Future studies should use a more active variant of the task including all functional runs, such as the one used by Arrington and colleagues (2019), who added oddball trials to the “fast localizer” task to both keep the participants engaged and allow for the simultaneous analysis of attentional control networks. Increasing the complexity of the stimuli (e.g., inclusion of multisyllabic words), and/or using more ecologically-valid reading
tasks (e.g., sentence or paragraph reading) may also provide additional insight into the differences in neural processing underlying reading among groups.

Unfortunately, due to recruitment challenges, the groups were not matched on non-verbal IQ or socioeconomic status. The mean performance of the typically-developing monolingual group on non-verbal IQ was a full standard deviation above the estimated population average, which indicates that this group is not representative of the intended population. Even so, the BL group demonstrated similar mean performance on this measure as their ML group, and the RD and EL groups were comparable to each other on this measure as well. Similarly, the ML and RD groups were comparable on measures of socioeconomic status, and the BL and EL groups also demonstrated similar socioeconomic backgrounds to each other. Although I controlled for these factors statistically, in future studies, it would be preferable to have well-matched groups. Using a school- or community-based recruitment approach as described above may be helpful in this endeavor as students from the same school districts will likely all be coming from similar socioeconomic backgrounds. Nevertheless, given the correlational nature of this research, it is not possible to ensure that groups would be perfectly matched across all measured and unmeasured factors.

Assessment validity was also a challenge due to the inclusion of students from diverse language backgrounds and literacy abilities. Of all of the behavioral measures administered, the WMLS-RNU measures were the only ones deemed
appropriate for use with bilingual and EL populations (Schrank, Wendling, & Alvarado, 2010), yet I decided to include the other measures as they are typically used in school and clinical settings when assessing students with reading difficulties (see Study 1). Of particular note is that many participants in the RD group had been receiving academic supports and personalized reading interventions for many years prior to data collection, and some had been administered these specific measures many times during neuropsychological evaluations. Three participants in the behavioral RD group received a standard score of 90 or above on all reading measures, yet all reported prior experience with a subset of these measures and indicated that they still receive services for current reading struggles in school. Only one such participant was included in the MRI sample. I decided not to exclude these participants from analyses because it is likely that at least a portion of adolescents with an RD diagnosis still receive services for subtle deficits despite remediation, and I wanted to represent the full spectrum of RD severity in my sample. Additional studies on this topic may benefit from the use of less common assessments to avoid practice effects, as well as assessments appropriate for use with bilingual participants.

Additional research is also necessary to facilitate interpretation of these results, provided their replication. For instance, the relatively strong performance of Spanish-English bilingual participants who entered as ELs in early elementary school on English reading measures suggests that recent immigrant ELs may eventually catch up to their peers over time provided adequate English language
instruction. Nevertheless, longitudinal research exploring English language and literacy development among ELs who have arrived in the U.S. later in their schooling is necessary to verify this. Other factors may lead to different developmental trajectories among these students, such as the absence of basic literacy instruction in English or exposure to English beyond a sensitive period for phonological processing. Further research is also necessary to explore whether differences in fMRI results from those of Malins and colleagues (2016) resulted from slight task differences, a lack of statistical power, developmental factors, or differences in participant characteristics. Additionally, the exploratory brain-behavior correlations need to be replicated prior to serious consideration, as brain-behavioral correlations do not often replicate in fMRI research (Yarkoni, 2009). Collection of more detailed qualitative information regarding the reading strategies that students use during the fMRI task may be helpful in interpreting such correlations, however. Lastly, this work provides many implications for the process of identifying RDs among ELs, yet future research using this approach should also include ELs with a diagnosed RD to fully explore this issue. Less than five such participants were identified via screening phone calls, so this study was not able to include this important group of participants. Furthermore, comparisons of ELs with RDs to both monolinguals with RDs in word reading and comprehension is necessary to fully explore this topic and provide sufficient basic knowledge to inform educational practice; however, the sample was limited to those with word-level RDs as more research has been conducted with this
population and the fMRI task explored the underpinnings of word reading processes.

**Conclusion**

The present study utilized cognitive neuroscience methods to compare the behavioral reading profiles, neural correlates of component reading processes, and their association among adolescents who were typically-developing monolingual English speakers, Spanish-English bilinguals, Spanish-speaking ELs, and monolingual English speakers with a diagnosed RD. Spanish-English bilinguals who had entered early elementary school as ELs performed similarly to their typically developing monolingual peers on all measures except for vocabulary and comprehension, which they were slightly lower on. Conversely, recent immigrant ELs literate in Spanish and monolinguals with RD performed approximately a standard deviation lower than the other groups, on average, on English measures of word reading and phonological memory, and ELs performed over 2.5 standard deviations lower than their typically-developing monolingual peers on English vocabulary and reading comprehension. This suggests that ELs who struggle with reading due to English language development share similarities in their behavioral reading profiles as monolinguals with RDs.

To explore the underpinnings of reading at a level beyond behavior, neural correlates of basic word reading processes were explored using fMRI and correlated with behavior. Similar regions across participants were associated with
non-alphabetic symbol processing, speech processing, word reading, lexical processing, semantic-similarity, and spelling-sound consistency, despite differences in behavioral performance across groups. Typically-developing monolinguals exhibited different response patterns than the other groups to spoken versus written language in the left fusiform gyrus and right superior temporal gyrus as well as spoken language versus vocoded speech in the left supramarginal gyrus. Finally, correlations between timed word reading measures and neural response to comparisons of interest were found both within and across groups in the bilateral inferior frontal gyrus, right middle frontal gyrus, and left precentral gyrus demonstrating the key role of frontal regions in facilitating efficient word reading. These findings need to be replicated in a larger sample with groups matched on IQ, socioeconomic status, and other factors. Taken together, however, this is a strong initial step in exploring the similarities and differences underlying reading processes using a sample that is more representative of the language and literacy diversity present in U.S. schools. Such an understanding may eventually lead to improved identification and treatment of reading difficulties and/or disorders among both English proficient and EL adolescents.

**GENERAL DISCUSSION**

ELs face unique challenges in U.S. schools as they often must learn academic skills and content in a language that they are still in the process of developing. Due to the dependence of reading on language processes, ELs often
struggle with reading as compared to their English proficient peers, particularly in
the reading comprehension domain (e.g., Lesaux, Crosson, Kieffer, & Pierce, 2010). Such difficulties are fairly widespread among such students, as only 9% of
fourth grade ELs across the U.S. achieved reading proficiency as measured by
the 2017 National Assessment for Education Progress (DoE, 2017).

With proper English language and literacy supports, typically-developing
ELs can achieve proficient conversational and academic English (Cummins,
1979; Thomas & Collier, 2002) as well as word reading skills on par with their
monolingual peers and improved reading comprehension skills, after many years
of schooling in English (Lesaux & Geva, 2006; Lesaux, Rupp, & Siegel, 2007;
Nakamoto, Lindsey, & Manis, 2007). However, ELs with underlying reading
disabilities (RD) are doubly disadvantaged, as their underlying disability may be
mistaken for reading difficulties due to normal second language learning.
Unfortunately, such students often do not receive the supports they need to
succeed in a timely manner, while others ELs are improperly labeled as having a
disability, as evidenced by the widespread disproportionality of ELs identified with
RDs reported at the national-, district-, and grade-levels (Artiles, Rueda, Salazar,
& Higareda, 2005; De Valenzuela, Copeland, Qi, & Park, 2006; Hibel & Jasper,
2012; IDEA Data Center, 2015; Linn & Hemmer, 2011; Shifrer, Muller, &
Callahan, 2011; Sullivan, 2011; Yamasaki & Luk, 2018; Zehler et al., 2003). In
this thesis, I utilized two different approaches aimed at advancing the knowledge
needed to improve RD identification processes for ELs. In this chapter, I will
briefly summarize the key findings, implications, and future directions for each study before presenting concluding remarks that synthesize the findings across both studies.

**Study 1: Investigating Current Practices Used to Identify RDs in ELs**

While prior research has provided guidance regarding best practices in assessment of ELs for suspected RDs (e.g., Abedi, 2006; Geva, 2000), relatively little research has been conducted to identify and address the challenges to this process in practice under ecologically-valid conditions. The most recent studies exploring identification practices for ELs in U.S. schools were conducted 10 to 20 years ago and/or only include information on a small number of schools or districts (Klingner & Harry, 2006; Orosco & Klingner, 2010; Sánchez et al., 2010; Thorius & Sullivan, 2013; Zehler et al., 2003). In my first study, I addressed this gap through the development and dissemination of an online survey on the current practices used to identify RDs among ELs to 598 school- and clinic-based practitioners across the U.S.

**Summary of Key Findings**

Results indicate that practitioners across a wide variety of professions participate in the RD identification process for ELs. These practitioners are largely monolingual or speak a language that is not widely represented by ELs. Practitioners reported using a wide range of cognitive, reading, language, and writing assessments with ELs, however many of these, particularly in the reading and writing domains, are not appropriate for use with this population and are only
available in English. Despite federal policy mandating that ELs are assessed in their native language, when feasible, using valid and appropriate assessments (IDEA, 2004), over 80% of practitioners reported conducting standardized and informal assessments in English only with ELs at least some of the time. Furthermore, roughly a third of the practitioners with experience assessing ELs for RDs did not know that such students could be reliably determined to have a disability prior to achieving English proficiency. Reported challenges to accurate and timely identification for ELs included a shortage of bilingual practitioners, dearth of appropriate assessments for ELs in English and non-English languages, difficulty involving families in the evaluation process, and insufficient training on best identification practices with ELs, consistent with prior research on this topic (Klingner and Harry, 2006; Zehler et al., 2003). Lastly, a majority of practitioners with experience identifying RDs in ELs reported lacking confidence in doing so.

Implications

The results of Study 1 suggest that many of the difficulties faced in identifying RDs in ELs have largely remained the same over the past 20 years. Practitioners identified challenges that would be difficult to address immediately, such as a lack of bilingual practitioners and appropriate assessments for ELs. Nevertheless, one immediately actionable item identified by the findings is a lack of adequate training on best practices for instructing and evaluating ELs. For instance, many of the participants indicated a) receiving little to no training on this
topic in graduate school and/or professional development, b) using identification criteria that have been determined to be biased against ELs, c) failing to use dynamic assessment with ELs despite its usefulness with this population, d) failing to review assessments to their valid and reliable use with ELs, e) believing falsely that ELs could not be accurately identified with an RD prior to developing full English proficiency, and f) a relatively low confidence level in evaluating ELs for RD as compared to the general population. An overwhelming majority of participants indicated interest in learning more about this topic, suggesting that future training efforts would be successful with this population. Many participants also indicated that the ELs that they have evaluated for special education services were determined to be struggling with reading due to inadequate instruction and/or a lack of English proficiency. If instruction at the classroom-level and prereferral strategies were working adequately for ELs in most settings, one would expect a fewer number of participants to report the occurrence of inappropriate referrals due to these reasons. This suggests that general educators and ESL teachers may also benefit from additional instruction on effective language and literacy instruction for ELs, including information about prereferral strategies and alternative supports for students struggling with reading due to factors other than a disability (e.g., Title I services, community services).

**Future Research Directions**

In addition to this survey, I also collected 20 interviews with practitioners in Massachusetts public schools with experience in identifying RDs in ELs.
Originally, I had planned to combine both approaches in a single mixed-methods study, however the interviews provided additional novel insights that warranted their own separate study. For instance, in the interviews, some practitioners revealed prejudices and deficit-oriented views of ELs and their families. In addition, many practitioners discussed the role of the official flowchart provided by the state in influencing their eligibility decisions. For my next step, I plan to finish analyzing these data, examine the state flowchart used to identify RDs, and potentially consult with Massachusetts Department of Elementary and Secondary Education officials regarding the development of this flowchart and other official documentation for identifying RDs in ELs. Hopefully this analysis can provide even further insight into the ways that official guidelines can be improved to ensure that more practitioners can access research-informed best practices in this endeavor. In particular, this approach allows me to explore the ways in which contextual and student-level factors shape the identification process for ELs, which was beyond the scope of Study 1. I would also like to use a qualitative approach to explore the identification practices used for ELs in clinical settings, and how they may differ from those used in school settings. Study 1 was limited in the ability to generalize findings to clinical populations due to difficulty in garnering responses from such practitioners; however, I believe that I would be able to more effectively recruit such subjects by establishing a personal connection through a qualitative approach.
Study 2: Behavioral and Neural Reading Profiles of Adolescents with Diverse Language and Reading Backgrounds

In a special issue of Learning Disabilities Research and Practice, in 2005, a number of leading researchers in the field of reading disabilities discussed the promise of cognitive neuroscience approaches in helping distinguish between reading difficulties due to normal second language learning and an underlying RD in ELs (McCardle, Mele-McCarthy, & Leos, 2005; Pugh et al., 2005; Simos, Billingsley-Marshall, Sarkari, Pataria, & Papanicolaou, 2005). Even with this advocacy, relatively few studies utilizing such approaches have been conducted with bilingual and/or EL children in the 14 years since then. In my second study, I addressed the lack of basic cognitive neuroscience research on reading among these populations by exploring behavioral reading profiles, neural correlates of component reading processes, and their association among adolescents who were typically-developing monolingual English speakers, Spanish-English bilinguals, Spanish-speaking ELs, and monolingual English speakers with a diagnosed RD. Seventy-six participants completed reading and language measures as well as a demographic questionnaire during a behavioral testing session, and 46 of them successfully completed a 17-minute fMRI task in which they passively read and listened to words, non-words, and other word-like stimuli.

Summary of Key Findings

In comparing adolescents’ behavioral reading profiles, I found that Spanish-English bilinguals displayed adequate English reading skills on par with
their typically-developing monolingual peers, yet Spanish-English ELs and monolinguals with word-level RDs demonstrated reading difficulties. ELs performed approximately a standard deviation lower on word reading measures and over 2.5 standard deviations lower on vocabulary and reading comprehension in English than their typically-developing peers, on average, despite having adequate Spanish reading abilities. Monolinguals with RDs also performed approximately one standard deviation below their typically-developing monolingual and bilingual peers on word reading measures, however reading comprehension difficulties on par with ELs were only observed using a grade-level passage reading measure. ELs scored significantly higher than the RD group on a rapid naming measure performed in their preferred language and phonemic decoding efficiency in English, while the RD group outperformed the ELs on English measures of phonological awareness and cloze reading.

Despite behavioral differences, few group differences in neural activation associated with component reading processes were observed. Across all participants, I found that regions in the bilateral superior temporal gyrus responded to spoken language; regions in the bilateral temporal cortices, left fusiform gyrus, left inferior frontal gyrus, left precentral gyrus, among others, were associated with word reading; and a cluster in the left precentral gyrus was associated with spelling-sound consistency. Three group differences in the relative mean neural response to conditions were found, with monolinguals exhibiting higher activation for speech versus vocoded speech than all other
groups in the left supramarginal gyrus; monolinguals responding significantly less
to spoken language than written language in the right superior temporal gyrus
than the bilingual and participants with RD; and monolinguals responding more to
written versus spoken language than these groups in the left fusiform and lateral
occipital gyrus. Lastly, a number of significant brain-behavior correlations within
and across groups were observed, yet additional research is necessary to
replicate these findings.

**Implications**

While basic research, particularly in the cognitive neuroscience domain,
is typically limited in its ability to directly inform educational practice, these results
provide a number of potential implications for the process of identifying RDs in
ELs. First of all, it is no mistake that for Study 2 I selected measures reported in
Study 1 as typically used among practitioners in the assessment of students with
suspected RD, including ELs. The behavioral portion of the study provided me
the opportunity to put on my practitioner hat, so to speak, and explore what such
neuropsychological data might look like in practice when assessing reading and
reading-related measures among a linguistically diverse population of
adolescents with and without reading difficulties. Using this approach, I was able
to witness firsthand the similarities in profiles of reading difficulty between
individual monolinguals with a diagnosis of RD and ELs who recently immigrated
to the U.S. Practitioners rarely have the opportunity to collect all of the student
data at once, compare average performance across groups to look for patterns,
and utilize the other advantages that I have as a researcher, however even
looking at the average performance of the RD and EL groups demonstrates
similar word reading, phonological, and grade-level passage reading
comprehension difficulties in English. This is the case even though the ELs in the
study were adequate readers in Spanish and had no personal or family history of
RD, therefore suggesting that such performance resulted from being in the early
stages of normal second language acquisition.

This stands in contrast to the widely held notion that ELs who
demonstrate word reading difficulties likely do so due to an underlying RD (Geva,
2000), and provides evidence that, particularly in the case of recent immigrants,
practitioners should avoid relying solely on English performance when assessing
Spanish-English adolescent ELs, as such results may be misleading. Of course,
in Study 1, I found that many ELs are reportedly not tested in their native
language, primarily due to a dearth of non-English assessments and bilingual
practitioners. In such cases, it may be helpful to allow students to complete the
RAN/RAS measure in Spanish, as the same materials can be used, and very
little knowledge of Spanish is necessary to score the assessment (e.g., basic
letter, number, and color names). In the sample, all of the ELs performed in the
“Average” range on this measure despite having poor performance on all other
reading measures in English, which demonstrates the utility of such a measure in
ruling out difficulties with automaticity in the child’s native language as a potential
contributing factor in his or her English reading difficulties.
Additionally, the widespread difficulty in word-level reading performance in English among ELs in the sample is consistent with the reports of practitioners in Study 1 that ELs referred to assessment often struggle with word level reading difficulties due to a lack of English proficiency. This story is inconsistent with the majority of research on English reading development among ELs who enter English-speaking schools in the early elementary years, as such research indicates that these students mainly struggle with reading comprehension skills despite adequate word-level skills in English (Lesaux et al., 2007; Nakemoto et al., 2007). Even so, Study 2 demonstrates that both stories can coexist in which adolescents who entered early elementary school as ELs struggle primarily with reading comprehension by middle school, if at all, while adolescent ELs who recently immigrated to the U.S. demonstrate difficulties across a wide range of reading skills, including word level measures, in English. Due to the relatively small sample size and non-random sampling approach, additional research is necessary to replicate such results.

Neuroimaging results generally provide fewer direct implications to educational practice, particularly given that this is among the first studies with this population. Even so, similar neural correlates of basic word reading processes suggest that low level reading processes in English are supported by similar mechanisms in Spanish-English bilinguals and Spanish-speaking ELs as their monolingual peers with and without RD. Reading difficulties among a subset of these students likely results from difficulties in higher level processing during
reading or low level difficulties not detected using this paradigm, however additional research utilizing more active tasks and complex stimuli are necessary to reach such conclusions. Correlations between behavioral and neural measures of reading provide preliminary evidence that adolescents may rely on different approaches (e.g., reliance on articulatory as opposed to sight word reading) to achieve successful reading in English depending on their language background or literacy ability. This study was underpowered to detect reliable correlations; however, these preliminary results suggest that this approach would be promising among a larger sample of participants.

**Future Research Directions**

There are several additional analyses that I would like to perform to continue this line of work. First of all, Malins and colleagues (2018) identified increased neural variability during basic word reading among bilinguals in the right middle frontal gyrus, which is a region associated with attentional control. I’ve already performed the first step necessary to conduct such an analysis, and I’d like to carry this project to fruition to explore whether I can replicate those results using this different paradigm and sample. Such findings would indicate that even though bilinguals and ELs draw upon resources to a similar extent as monolinguals, their response is less consistent within particular regions as their monolingual peers. It would be interesting to explore how this result may differ for participants with RD, as well. A surface-based analysis of the fMRI data using novel methods, such as those described by Coalson and colleagues (2018), may
also be worth pursuing given that fNIRS studies that identified significant differences between monolingual and bilingual children involve surface-based analyses (e.g., Jasínska et al., 2017).

In addition, I would like to set regions identified during the “fast localizer” task, such as the VWFA and left inferior frontal gyrus, as seeds in an analysis exploring group differences in resting state functional connectivity among regions implicated in basic reading processes. Prior research has indicated that children and adolescents with RDs demonstrate reduced functional connectivity among such regions during rest as compared to their typically-developing peers (e.g., Schurz et al., 2014), however no such research has compared resting-state connectivity among these key hubs among bilingual and EL participants, as well. It would be particularly interesting to also perform functional connectivity during naturalistic reading in English among such groups to explore how experience expectant versus experience dependent mechanisms may modulate the connectivity among regions in the reading network both at rest and during reading. In the present study, I did not collect naturalistic reading data in the MRI, so additional research would need to be conducted to do so. Students did engage in a naturalistic learning paradigm where they watched a grade-appropriate video on a science-related topic, so further connectivity analyses can be performed using these data to explore differences among groups in the networks that support learning under more ecologically-valid conditions.
Eventually, it would also be informative to include ELs with diagnosed RDs and monolinguals with RDs in reading comprehension into the fold to explore how their behavior and neural reading profiles compare. Further down the line, with replication of results, such studies may be useful in identifying behavioral markers and/or developing assessments that can be used to distinguish reading difficulty due to normal second language development from that due to an underlying RD. The ability to do so now is currently limited by a lack of basic research on this topic; however, with bilinguals comprising nearly a quarter of all school-age children (U.S. Census Bureau, 2019), and ELs comprising 10% of the public-school population (McFarland et al., 2019), basic research on reading must expand its population of interest to include such students. This is necessary to ensure that all children in U.S. classrooms have access to appropriate, research-informed educational instruction.

**Concluding Remarks**

Taken together, this thesis serves to inform the next steps in improving the identification of RDs in ELs. In Study 1, I investigated the range of current practices and reported challenges involved in this process to identify circumstances that could be directly addressed through changes in practice and/or policy as well as those that call for additional research. Using this approach, I was able to identify weaknesses in practitioners’ knowledgebase suggesting that additional training covering such information would be helpful to directly improve identification processes for ELs. Other findings, such as a lack of
bilingual practitioners and appropriate assessments for ELs, indicate that additional basic research exploring the behavioral reading profiles of ELs conducted using measures and methods similar to those commonly used in schools may be helpful to explore the areas in which such students demonstrate weaknesses that may be misinterpreted as evidence of RD. Given that a majority of practitioners in Study 1 reported relying on standardized measures conducted in English that were often not valid or reliable for use with ELs to make their determinations for RD (e.g., using an IQ/Achievement discrepancy), additional research comparing performance on such measures among various subgroups of ELs and students with RD may help identify similarities and differences in the reading profiles of such students that could be useful to consider in practice.

In Study 2, I address this research gap by performing behavioral comparisons of adolescents’ reading profiles using this approach. Interestingly, I find that Spanish-English bilingual students who entered early elementary school as ELs showed similar reading profiles to their typically-developing monolingual peers, with few students performing in the “Below Average” range on measures of English vocabulary or reading comprehension despite these being relative areas of weakness for such students. Consistent with prior research (e.g., Geva, 2000), this suggests that adolescents with this language history may have an RD if they exhibit reading difficulties on such measures, particularly on word-level skills. Few studies have investigated the reading profiles of adolescent ELs who recently immigrated to the U.S., however. My findings suggest that such students
underperform on many of these measures, similar to monolinguals with RD, despite being typically-developing and having adequate reading skills in Spanish. This might explain the difficulties that practitioners have when evaluating ELs for RDs using similar methods. While I suggest that monolingual English-speaking practitioners use assessments valid and reliable for ELs whenever possible (e.g., nonverbal IQ measures, WMLS-RNU) and allow Spanish-speaking students to complete the RAN/RAS in their preferred language, these findings illustrate the difficulty in using behavioral methods alone to distinguish between reading difficulties and RD among ELs. I therefore initiated fMRI research on the neural correlates of component reading processes among a subset of the behavioral participants to compare neural measures among groups. Few group differences in neural response to processes of interest (e.g., word reading, lexicality, spoken language) were identified, suggesting that the low-level neural processes underlying basic word reading are largely similar across adolescents with diverse language background and reading abilities. This suggests that a basic word reading task of this nature may not be best-suited to identify neural differences underlying reading difficulties among adolescents. Even so, additional research in this domain may be useful in identifying the similarities and differences in the neural underpinnings of reading and their behavioral correlates among monolinguals and bilinguals with and without RDs, which can eventually be used to inform educational practice.
While additional research is necessary to replicate these findings, by exploring the range of identification practices currently in use and incorporating ELs into the burgeoning line of behavioral and fMRI research on reading, this thesis sought to provide additional insight into the ways that the identification process can be improved so that ELs can receive appropriate and efficient reading instruction in U.S. schools.
APPENDIX A: Reading Diagnostics Survey

Diagnostic Approaches for Reading Disabilities: A Survey of Practitioners
Please read the following information carefully before deciding whether to participate in this survey. You are invited to take part in a research survey about assessment practices for reading disabilities. Your participation will require approximately 15 minutes and is completed online at your computer. There are no immediate benefits for participating in this survey. Once analyzed, results will be aggregated and disseminated to the broader public without any reports of individual level data. Information gained through this study may help to identify the prevalence of current reading assessment approaches and inform future studies aimed at guiding practice.

PARTICIPATION: Taking part in this study is completely voluntary. You may refuse to take part in the research and are free to decline to answer any particular question you do not wish to answer for any reason. If you choose to participate in the study, you can discontinue the survey at any time by closing your browser window. The information collected in this survey will be stored in a secure database and will be used for research purposes only. Your responses will be anonymous and any report of this research that is made available to the public will only be shared in aggregated form with no identifying information.

RISKS: There are no known risks or discomforts associated with this survey.

CONFIDENTIALITY: Your survey answers will be sent to a link where data will be stored in a password protected electronic format. Identifying information such as your name, email address, or IP address will not be collected. Therefore, your responses will remain anonymous. No one will be able to identify you or your answers, and no one will know whether or not you participated in the study.

CONTACT: If you have questions, concerns, or comments about this study, please contact Joanna Christodoulou at jchristodoulou@mghihp.edu or by phone at (617) 643-1482. If you have questions about your rights in this research, have a complaint or want to report a problem to someone not directly involved in this research study, please contact the Partners Human Research Committee office by phone at (617) 424-4100.

CONSENT: Please feel free to print a copy of this consent page to keep for your records. Clicking "Agree" below indicates your consent to participate in this survey and that: You have read the above information. You voluntarily agree to participate. You are 18 years of age or older.

☐ Agree

☐ Disagree
Important Note: Please read the following information carefully.
Throughout the survey, the term RD will be used to mean either reading disability or reading disorder unless otherwise specified. This includes specific learning disability in reading, specific learning disorder in reading, dyslexia, and other terms commonly used to describe a diagnosed or identified reading problem. The specific term (e.g., dyslexia) will be used when solely referring to a specific type of diagnosed reading problem. Throughout the survey, the term ELLs will be used to refer to students who are not yet proficient in English. These students are typically referred to as English Language Learners (ELLs), English Learners (ELs), Limited English Proficient (LEP) students, and English as a Second Language (ESL) students. All questions specific to ELLs will be preceded by ELL: Click "I understand" to indicate that you've read the important note.

☐ I understand

Q1. Are you currently responsible for identifying or diagnosing RDs in school-age children?
☐ Yes, I make diagnostic decisions independently.
☐ Yes, I work on a team that is responsible for making diagnostic decisions.
☐ No, I have no role in identifying or diagnosing RDs in school-age children.

Q2. Do you conduct assessments used to determine whether a child has a RD?
☐ Yes, I personally conduct all of the assessments used to make the determination.
☐ Yes, I conduct some of the assessments used to make the determination.
☐ No, I do not conduct any of the assessments used to make the determination.
Q3. Which types of assessments are you responsible for conducting when evaluating a child for a RD? (Select all that apply)

☐ cognitive assessments

☐ language assessments

☐ reading skill assessments

☐ speech assessments

☐ writing assessments

☐ Other(s) (please specify) ____________________________________________
Q4. Which member(s) of your team are responsible for conducting the other assessments used to make the determination for a RD? (Select all that apply)

☐ Clinical Psychologist
☐ Educational Specialist
☐ ESL Teacher
☐ General Education Teacher
☐ Literacy Coach
☐ Neurologist
☐ Neuropsychologist
☐ Pediatrician
☐ Reading/Literacy Specialist
☐ School Psychologist
☐ Special Education Teacher
☐ Speech-Language Pathologist
☐ Other (please specify) ____________________________________________
☐ I don't know
Q5. Which member(s) of your team are responsible for conducting the assessments used to make the determination for a RD? (Select all that apply)

☐ Clinical Psychologist

☐ Educational Specialist

☐ ESL Teacher

☐ General Education Teacher

☐ Literacy Coach

☐ Neurologist

☐ Neuropsychologist

☐ Pediatrician

☐ Reading/Literacy Specialist

☐ School Psychologist

☐ Special Education Teacher

☐ Speech-Language Pathologist

☐ Other (please specify) ______________________________________________________

☐ I don't know
Q6. ELL: Do you have any experience identifying or diagnosing RDs in ELLs?

- Yes, I have made diagnostic decisions regarding ELLs independently.
- Yes, I have worked on a team that has made diagnostic decisions regarding ELLs.
- No

Q7. ELL: Have you conducted assessments used to determine whether an ELL has a RD?

- Yes, I have personally conducted all of the assessments used to make the determination for an ELL with a suspected RD.
- Yes, I have conducted some of the assessments used to make the determination for an ELL with a suspected RD.
- No, I have not conducted any of the assessments used to make the determination for an ELL with suspected RD.
Q8. ELL: Which types of assessments are you responsible for conducting when evaluating an ELL for a RD? (Select all that apply)

☐ cognitive assessments in English

☐ cognitive assessments in the child's native language

☐ language assessments in English

☐ language assessments in the child's native language

☐ reading assessments in English

☐ reading assessments in the child's native language

☐ speech assessments in English

☐ speech assessments in the child's native language

☐ writing assessments in English

☐ writing assessments in the child's native language

☐ Other(s) (please specify) __________________________________________________________
Q9. ELL: Which member(s) of your team are responsible for conducting the other assessments used to make the determination for a RD in ELLs? (Select all that apply)

☐ Clinical Psychologist

☐ Educational Specialist

☐ ESL Teacher

☐ General Education Teacher

☐ Literacy Coach

☐ Neurologist

☐ Neuropsychologist

☐ Pediatrician

☐ Reading/Literacy Specialist

☐ School Psychologist

☐ Special Education Teacher

☐ Speech-Language Pathologist

☐ Other (please specify) ____________________________________________

☐ I don't know
Q10. ELL: Which member(s) of your team are responsible for conducting the assessments used to make the determination for a RD in ELLs? (Select all that apply)

- Clinical Psychologist
- Educational Specialist
- ESL Teacher
- General Education Teacher
- Literacy Coach
- Neurologist
- Neuropsychologist
- Pediatrician
- Reading/Literacy Specialist
- School Psychologist
- Special Education Teacher
- Speech-Language Pathologist
- Other (please specify) ________________________________
- I don't know

End of Block: Branching

Start of Block: Demographics
Q11. Please select the best description of your current occupation: (Select all that apply)

- [ ] Clinical Psychologist
- [ ] Educational Specialist
- [ ] ESL Teacher
- [ ] General Education Teacher
- [ ] Literacy Coach
- [ ] Neurologist
- [ ] Neuropsychologist
- [ ] Pediatrician
- [ ] Reading/Literacy Specialist
- [ ] School Psychologist
- [ ] Special Education Teacher
- [ ] Speech-Language Pathologist
- [ ] Not currently employed (e.g., graduate student training for one of these occupations)
- [ ] Other (please specify) ________________________________
Q12. How many years of experience do you have in this occupation?

- [ ] Less than 1 year
- [ ] 1-3 years
- [ ] 4-6 years
- [ ] 7-9 years
- [ ] 10-12 years
- [ ] 13-15 years
- [ ] 16 years or more
- [ ] N/A
Q13. In which state do you work?

- [ ] AK
- [ ] AL
- [ ] AR
- [ ] AZ
- [ ] CA
- [ ] CO
- [ ] CT
- [ ] DE
- [ ] FL
- [ ] GA
- [ ] HI
- [ ] IA
- [ ] ID
- [ ] IL
- [ ] IN
- [ ] KS
- [ ] KY
- [ ] LA
- [ ] MA
- [ ] MD
- [ ] ME
<table>
<thead>
<tr>
<th>State Abbreviation</th>
<th>State Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI</td>
<td>MN</td>
</tr>
<tr>
<td>MO</td>
<td>MS</td>
</tr>
<tr>
<td>MT</td>
<td>NC</td>
</tr>
<tr>
<td>ND</td>
<td>NE</td>
</tr>
<tr>
<td>NH</td>
<td>NJ</td>
</tr>
<tr>
<td>NM</td>
<td>NV</td>
</tr>
<tr>
<td>NY</td>
<td>OH</td>
</tr>
<tr>
<td>OK</td>
<td>OR</td>
</tr>
<tr>
<td>PA</td>
<td>RI</td>
</tr>
<tr>
<td>SD</td>
<td>SC</td>
</tr>
<tr>
<td>TN</td>
<td></td>
</tr>
</tbody>
</table>
Q14. Please indicate your primary work setting:

- Public School: Non-Charter
- Public School: Charter
- Private/Independent School: Non-Special Education
- Private/Independent School: Special Education
- Hospital
- Private Practice
- Other (please specify) ________________________________
- N/A
Q15. Which of the following best describes your primary work setting?

- Urban
- Suburban
- Rural
- N/A

Q16. Please indicate your secondary work setting:

- Public School: Non-Charter
- Public School: Charter
- Private/Independent School: Non-Special Education
- Private/Independent School: Special Education
- Hospital
- Private Practice
- Other (please specify) ______________________________________________________
- N/A
Q17. What is your highest degree?

- High School
- B.A./ B.S.
- Ed.M./ M.A./ M.S./ M.A.T.
- Ed.D./ Ph.D.
- M.D.
- Psy.D.
- Other (please specify) ________________________________________________

Q18. Please estimate the amount of time in your graduate work dedicated to learning about:

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>none</th>
<th>single lecture</th>
<th>multiple lectures</th>
<th>single course</th>
<th>multiple courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of oral language skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of reading skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of writing skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of reading skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of writing skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification/diagnosis of RDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification/diagnosis of RDs in ELLs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q19. Excluding graduate coursework, please estimate the amount of time over the course of your entire career (i.e., in professional development) dedicated to learning about:

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>none</th>
<th>single session</th>
<th>multiple sessions</th>
<th>single workshop or conference</th>
<th>multiple workshops or conferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of oral language skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of reading skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of writing skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of reading skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of writing skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification/diagnosis of RDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification/diagnosis of RDs in ELLs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q20. Which of the following grades are the children you work with in? (Select all that apply)

☐ Pre-K

☐ K

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ 7

☐ 8

☐ 9

☐ 10

☐ 11

☐ 12

☐ N/A
Q21. Approximately what percentage of your caseload of school-age children includes those identified or diagnosed with a RD?

<table>
<thead>
<tr>
<th>% of caseload with RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>100</td>
</tr>
</tbody>
</table>

Q22. ELL: Approximately what percentage of the children on your caseload with RDs are ELLs?

<table>
<thead>
<tr>
<th>% of RD students who are ELLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>100</td>
</tr>
</tbody>
</table>
Q23. What language(s) do you speak other than English? (Select all that apply)

- None
- Arabic
- Cantonese
- Cape Verdean
- French
- German
- Haitian Creole
- Hmong
- Italian
- Japanese
- Khmer/Cambodian
- Korean
- Mandarin
- Polish
- Portuguese
- Russian
Q24. ELL: Have you ever used your non-English language(s) when assessing an ELL for a suspected RD?

- [ ] Yes
- [ ] No
- [ ] N/A

☐ Spanish
☐ Tagalog
☐ Vietnamese
☐ Other(s) (please specify) ____________________________
Q25. ELL: What language(s) do the ELLs you have evaluated for suspected a RD speak, other than English? (Select all that apply)

- [ ] I don't know
- [ ] Arabic
- [ ] Cantonese
- [ ] Cape Verdean
- [ ] French
- [ ] German
- [ ] Haitian Creole
- [ ] Hmong
- [ ] Italian
- [ ] Japanese
- [ ] Khmer/Cambodian
- [ ] Korean
- [ ] Mandarin
- [ ] Polish
- [ ] Portuguese
- [ ] Russian
☐ Spanish

☐ Tagalog

☐ Vietnamese

☐ Other(s) (please specify) _______________________________________

☐ N/A

End of Block: Demographics

Start of Block: RD assessment
Q26. Including you, which professionals in your setting make up the child study team responsible for identifying or diagnosing RDs? (Select all that apply)

- Clinical Psychologist(s)
- Educational Specialist(s)
- ESL Teacher(s)
- General Education Teacher(s)
- Literacy Coach(es)
- Neurologist(s)
- Neuropsychologist(s)
- Pediatrician(s)
- Reading/Literacy Specialist(s)
- School Psychologist(s)
- Special Education Teacher(s)
- Speech-Language Pathologist(s)
- Other (please specify) ________________________________________________
- I don't know
Q27. ELL: Including you, which professionals in your setting make up the child study team responsible for identifying or diagnosing RDs in ELLs? (Select all that apply)

☐ Clinical Psychologist(s)

☐ Educational Specialist(s)

☐ ESL Teacher(s)

☐ General Education Teacher(s)

☐ Literacy Coach(es)

☐ Neurologist(s)

☐ Neuropsychologist(s)

☐ Pediatrician(s)

☐ Reading/Literacy Specialist(s)

☐ School Psychologist(s)

☐ Special Education Teacher(s)

☐ Speech-Language Pathologist(s)

☐ Other (please specify) ____________________________

☐ I don't know
Q28. When you assess students for a suspected RD, how often do you evaluate:

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral Language Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid Naming (RAN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological Processing Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar/Morpho-Syntactic Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoword Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Fluency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spelling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written Expression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q29. Which of the following assessments do you typically use when assessing children for RDs? 
NOTE: Please select the test regardless of the edition, form, or version. For example, if you use the TOWRE-2, select TOWRE. (Select all that apply)

☐ AIMSweb

☐ Analytical Reading Inventory (ARI)

☐ Clinical Evaluation of Language Fundamentals (CELF)

☐ Comprehensive Assessment of Spoken Language (CASL)

☐ Comprehensive Test of Phonological Processing (CTOPP)

☐ Diagnostic Achievement Battery (DAB)

☐ Diagnostic Assessments of Reading (DAR)

☐ Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

☐ Developmental Reading Assessment (DRA)

☐ Expressive One-Word Picture Vocabulary Test (EOWPVT)

☐ Expressive Vocabulary Test (EVT)

☐ Fountas and Pinnell Benchmark Assessment System (BAS)

☐ Gates-MacGinitie Reading Test (GMRT)

☐ Gray Oral Reading Tests (GORT)

☐ Group Reading Assessment and Diagnostic Evaluation (GRADE)
Grey Silent Reading Tests (GSRT)
Kaufman Assessment Battery for Children (KABC)
Kaufman Brief Intelligence Test (KBIT)
Kaufman Test of Educational Achievement (KTEA)
Nelson-Denny Reading Test
Oral and Written Language Scales (OWLS)
Peabody Picture Vocabulary Test (PPVT)
Process Assessment of the Learner (PAL)
Qualitative Reading Inventory (QRI)
Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN/RAS)
Receptive One-Word Picture Vocabulary Test (ROWPVT)
Scholastic Reading Inventory (SRI)
Stanford-Binet Intelligence Scales (SBIS)
Test of Auditory Comprehension of Language (TACL)
Test of Auditory Processing Skills (TAPS)
Test of Early Written Language (TEWL)
<table>
<thead>
<tr>
<th>Test of Nonverbal Intelligence (TONI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test of Integrated Language and Literacy Skills (TILLS)</td>
</tr>
<tr>
<td>Test of Orthographic Competence (TOC)</td>
</tr>
<tr>
<td>Test of Reading Comprehension (TORC)</td>
</tr>
<tr>
<td>Test of Silent Word Reading Fluency (TOSWRF)</td>
</tr>
<tr>
<td>Test of Word Reading Efficiency (TOWRE)</td>
</tr>
<tr>
<td>Test of Written Language (TOWL)</td>
</tr>
<tr>
<td>Test of Written Spelling (TWS)</td>
</tr>
<tr>
<td>Wechsler Individual Achievement Test (WIAT)</td>
</tr>
<tr>
<td>Wechsler Intelligence Test for Children (WISC)</td>
</tr>
<tr>
<td>Wilson Assessment of Decoding and Encoding (WADE)</td>
</tr>
<tr>
<td>Word Identification and Spelling Test (WIST)</td>
</tr>
<tr>
<td>Woodcock-Johnson Tests of Achievement (WJ ACH)</td>
</tr>
<tr>
<td>Wide Range Achievement Test (WRAT)</td>
</tr>
<tr>
<td>Woodcock Reading Mastery Tests (WRMT)</td>
</tr>
<tr>
<td>Other(s) (please specify) ________________________________________________</td>
</tr>
</tbody>
</table>
Q30. Please select the main reasons for selecting these assessments (Select all that apply)

☐ Accessibility: These are the measures I have access to.

☐ Training: These are the measures that I have been trained to use.

☐ Affordability: These are the measures we have been able to afford.

☐ Ease of use: These measures are reasonable to learn and administer in terms of ease of use.

☐ Time: These measures are practical for use given time constraints.

☐ Obligation: These are the measures we are obliged to use by superiors/the community.

☐ Other (please specify) _______________________________________________________

☐ N/A
Q31. ELL: Which of the following assessments do you typically use when assessing ELLs for RDs? NOTE: Please select the test regardless of the edition, form, or version. For example, if you use the TOWRE-2, select TOWRE. (Select all that apply)

☐ AIMSweb

☐ Analytical Reading Inventory (ARI)

☐ Bilingual Verbal Ability Tests (BVAT)

☐ Clinical Evaluation of Language Fundamentals (CELF)

☐ Comprehensive Assessment of Spoken Language (CASL)

☐ Comprehensive Test of Phonological Processing (CTOPP)

☐ Diagnostic Achievement Battery (DAB)

☐ Diagnostic Assessments of Reading (DAR)

☐ Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

☐ Developmental Reading Assessment (DRA)

☐ Expressive One-Word Picture Vocabulary Test (EOWPVT)

☐ Expressive Vocabulary Test (EVT)

☐ Evaluacion del desarrollo de la lectura (EDL)

☐ Fountas and Pinnell Benchmark Assessment System (BAS)

☐ Gates-MacGinitie Reading Test (GMRT)
Gray Oral Reading Tests (GORT)

Group Reading Assessment and Diagnostic Evaluation (GRADE)

Grey Silent Reading Tests (GSRT)

IDEA Oral Language Proficiency Tests (IPT)

Indicadores Dinamicos del Exito en la Lectura (IDEL)

Kaufman Assessment Battery for Children (KABC)

Kaufman Brief Intelligence Test (KBIT)

Kaufman Test of Educational Achievement (KTEA)

Nelson-Denny Reading Test

Oral and Written Language Scales (OWLS)

Peabody Picture Vocabulary Test (PPVT)

Process Assessment of the Learner (PAL)

Qualitative Reading Inventory (QRI)

Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN/RAS)

Receptive One-Word Picture Vocabulary Test (ROWPVT)

Scholastic Reading Inventory (SRI)
Stanford-Binet Intelligence Scales (SBIS)

Test de Vocabulario en Imágenes Peabody (TVIP)

Test of Auditory Comprehension of Language (TACL)

Test of Auditory Processing Skills (TAPS)

Test of Early Written Language (TEWL)

Test of Nonverbal Intelligence (TONI)

Test of Integrated Language and Literacy Skills (TILLS)

Test of Orthographic Competence (TOC)

Test of Reading Comprehension (TORC)

Test of Silent Word Reading Fluency (TOSWRF)

Test of Word Reading Efficiency (TOWRE)

Test of Written Language (TOWL)

Test of Written Spelling (TWS)

Wechsler Individual Achievement Test (WIAT)

Wechsler Intelligence Test for Children (WISC)

Wilson Assessment of Decoding and Encoding (WADE)
☐ Word Identification and Spelling Test (WIST)

☐ Woodcock-Johnson Tests of Achievement (WJ ACH)

☐ Woodcock-Munoz Language Survey (WMLS)

☐ Wide Range Achievement Test (WRAT)

☐ Woodcock Reading Mastery Tests (WRMT)

☐ Other(s) (please specify) ______________________________________________________

☐ N/A
Q32. ELL: Please select the main reasons for selecting these assessments for use with ELLs. (Select all that apply)

☐ Accessibility: These are the measures I have access to.

☐ Training: These are the measures that I have been trained to use.

☐ Affordability: These are the measures we have been able to afford.

☐ Ease of use: These measures are reasonable to learn and administer in terms of ease of use.

☐ Time: These measures are practical for use given time constraints.

☐ Obligation: These are the measures we are obliged to use by superiors/the community.

☐ Other (please specify) ________________________________

☐ N/A
Q33. Which of the following is used in your setting as qualification criteria for a reading-related specific learning disability? (Select all that apply)

- [ ] External Diagnosis (please specify from whom)
- [ ] Failure to benefit from Tier 1 RTI/MTSS instruction
- [ ] Failure to benefit from Tier 2 RTI/MTSS instruction
- [ ] IQ/Achievement discrepancy
- [ ] Pattern of Strengths and Weaknesses
- [ ] Standard deviation(s) below mean on one test
- [ ] Standard deviation(s) below mean on multiple tests
- [ ] Other (please specify) ________________________________
- [ ] I don't know
- [ ] N/A

Q34. If you use a standard deviation below the mean, please specify:

- [ ] 1
- [ ] 1.5
- [ ] 2
- [ ] 2.5
- [ ] Other (please specify) ________________________________
Q35. Which of the following labels would you consider using when diagnosing a RD? (Check all that apply)

<table>
<thead>
<tr>
<th>Developmental Delay</th>
<th>Dyslexia</th>
<th>Language-Based Learning Disability</th>
<th>Reading Disability</th>
<th>Reading Disorder</th>
<th>Specific Reading Disability</th>
<th>Specific Learning Disability</th>
<th>Specific Learning Disorder</th>
<th>Other</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades K-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 3 and up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Q36. If you selected 'Other' for Grades K-2, please specify:

________________________________________________________________

Q37. If you selected 'Other' for Grades 3 and up, please specify:

________________________________________________________________

Q38. The exclusionary clause under IDEA holds that a learning disability "does not include a learning problem that is primarily the result of environmental, cultural, or economic disadvantage." If a child with reading difficulties comes from a disadvantaged home but otherwise meets criteria for a RD, is s/he still eligible for a RD designation (e.g., SLD) in your setting?

○ Yes

○ No

○ I don't know

○ N/A
Q39. Is a child with reading difficulties from a disadvantaged home eligible for a different designation instead?

- Yes (please specify) ____________________________
- No
- I don't know

Q40. In your setting, are school-based practitioners permitted to use the label of dyslexia (e.g., in an IEP)?

- Yes, they can use the label of dyslexia to describe any child with a RD.
- Yes, they can use the label of dyslexia to describe any child with a RD specific to single word difficulties.
- Yes, but only for children who have received an external diagnosis of dyslexia from a medical professional.
- Yes, other (please specify) ____________________________
- No (please specify the label they use instead) ____________________________
- I don't know
- N/A
Q41. In your setting, do you think school-based practitioners should be permitted to use the label of dyslexia (e.g., in an IEP)?

- Yes, they should be able to use the label of dyslexia to describe any child with a RD.
- Yes, they should be able to use the label of dyslexia to describe any child with a RD specific to single word difficulties.
- Yes, but only for children who have received an external diagnosis of dyslexia from a medical professional.
- Yes, other (please specify) _________________________________
- No (please specify) _________________________________
- I don't know
- N/A
Q42. How many times per school year do you collect assessment data that is used to make decisions about adequate progress?

- None
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 or more
- N/A
Q43. Which of the following descriptions reflect your evidence for adequate progress?

- Performing at least at benchmark for curriculum-based measures of reading
- Performance improvement on curriculum-based measures of reading
- Performing in the average range or higher on standardized assessments of reading
- Performance improvement on standardized assessments of reading
- No regression
- Other (please specify) ________________________________________________
- N/A

Q44. In your setting, about how long does it typically take for each of these processes to occur relative to the previous step?

<table>
<thead>
<tr>
<th>Duration a student struggles before being referred to evaluation team at school.</th>
<th>2 wks or less</th>
<th>1 mth</th>
<th>2-3 mnths</th>
<th>4-5 mnths</th>
<th>6-7 mnths</th>
<th>8-9 mnths</th>
<th>10-11 mnths</th>
<th>1 yr</th>
<th>&gt; 1 yr</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation team deems testing appropriate.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child is tested.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test results reviewed by child study team.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualification for special education determined.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes to education are implemented.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### How often do you use the following techniques when assessing ELLs for RDs?

**Q45. ELL: Conduct standardized assessments in:**

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>English only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child's native language only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both English and the child's native language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Q46. ELL: Conduct informal assessments (e.g., criterion referenced measures, language sampling) in:**

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>English only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child's native language only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both English and the child's native language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Q47. ELL: Examine assessments used with ELLs:

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>For cultural bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For inclusion of ELLs/bilinguals in the norming sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To ensure that they are being used for the purposes for which they are valid and reliable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
### Q48. ELL: Use the following assessment practices with ELLs:

<table>
<thead>
<tr>
<th></th>
<th>N/A</th>
<th>Never</th>
<th>Rarely</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus on measuring reading skills rather than English proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use interpreters to translate and/or conduct assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assess the child on multiple dates using the same measures (test-retest reliability)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use dynamic assessment practices (test-teach-retest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use the IQ/Achievement discrepancy criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observe the child in multiple classroom contexts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gather information about the student from their teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct interviews with parents and caregivers about the students' language and reading background</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q49. ELL: Which of the following may explain why ELLs in your setting may not be assessed in their native language(s)? (Select all that apply)

- All students are always assessed in their native language in my setting.
- Students who speak languages for which we have non-English assessments and practitioners who speak that language are always assessed in their native language, but other students are assessed solely in English.
- We do not have sufficient assessments in non-English languages to meet the language profiles of our students.
- We do not have sufficient bilingual practitioners to assess students in non-English languages.
- We do not have sufficient time to assess children in both languages.
- We do not believe that it is important to assess ELLs in their non-English language(s).
- I don't know
- Other (please specify) ________________________________________________
Q50. ELL: Based on your experiences, which of the following factors have served as the primary cause(s) of word-level reading difficulties (e.g., phonological awareness, decoding, fluency) in the ELLs that you have evaluated for a suspected RD? (Select all that apply)

☐ A lack of oral language proficiency in English (e.g., low vocabulary)

☐ Inadequate reading instruction (e.g., due to cultural, environmental, and/or economic disadvantages)

☐ An underlying RD

☐ Executive function (e.g., working memory) deficits

☐ Cognitive deficits (e.g., low IQ)

☐ Social-emotional and/or behavioral issues

☐ Difficulties distinguishing between speech sounds not present in their native language

☐ Other (please specify) ________________________________

☐ N/A
Q51. ELL: Based on your experiences, which of the following factors have served as the primary cause(s) of reading comprehension difficulties in the ELLs that you have evaluated for a suspected RD? (Select all that apply)

☐ A lack of oral language proficiency in English (e.g., low vocabulary)

☐ Inadequate reading instruction (e.g., due to cultural, environmental, and/or economic disadvantages)

☐ An underlying RD

☐ Executive function (e.g., working memory) deficits

☐ Cognitive deficits (e.g., low IQ)

☐ Social-emotional and/or behavioral issues

☐ Difficulties with word-level reading skills

☐ Other (please specify) ____________________________________________

☐ N/A
Q52. ELL: Agree or Disagree–It is important to wait until oral language proficiency in English is fully developed before assessing ELLs for a RD.

☐ Agree

☐ Disagree

☐ I don't know

Q53. ELL: What do you believe are the greatest challenges to the accurate and timely identification of RDs in ELLs? (Select all that apply)
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>None: I do not believe that there are any challenges specific to ELLs.</td>
<td></td>
</tr>
<tr>
<td>Policies: Ineffective, inappropriate, and/or absent district and/or state policies regarding eligibility/assessment procedures.</td>
<td></td>
</tr>
<tr>
<td>English assessments: The lack of appropriate assessments in English (racially and culturally and appropriate, validated for use with ELLs).</td>
<td></td>
</tr>
<tr>
<td>Non-English assessments: The lack of appropriate assessments in non-English languages.</td>
<td></td>
</tr>
<tr>
<td>Practitioners: The lack of practitioners able to conduct assessments in non-English languages.</td>
<td></td>
</tr>
<tr>
<td>RTI/MTSS: The improper/ineffective use of tiered intervention models.</td>
<td></td>
</tr>
<tr>
<td>Training: A lack of training on how to conduct assessments with ELLs for RDs.</td>
<td></td>
</tr>
<tr>
<td>Research: A lack of research on best practices in assessing ELLs for RDs.</td>
<td></td>
</tr>
<tr>
<td>Families: Difficulties working/communicating with the families of ELLs.</td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
</tr>
<tr>
<td>I don't know.</td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
Q54. Please rate your general level of confidence in identifying or diagnosing RDs.

- 10 (Most Confident)
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1 (Least Confident)
- N/A
Q55. ELL: Please rate your level of confidence in identifying or diagnosing RDs in ELLs.

- [ ] 10 (Most Confident)
- [ ] 9
- [ ] 8
- [ ] 7
- [ ] 6
- [ ] 5
- [ ] 4
- [ ] 3
- [ ] 2
- [ ] 1 (Least Confident)
- [ ] N/A
Q56. How important is it for you to learn more about diagnostic processes related to reading?

- Extremely important
- Very important
- Moderately important
- Slightly important
- Not at all important

Q57. ELL: How important is it for you to learn more about diagnostic processes related to reading for ELLs specifically?

- Extremely important
- Very important
- Moderately important
- Slightly important
- Not at all important

End of Block: End Block
Appendix B: Brain-Behavior Correlations

All significant brain-behavior correlations performed are presented in Table 2.5, and a subset of these are also graphically displayed in Figure 2.7. Among participants in the ML group, I only found one statistically significant correlation between neural response and TOWRE PDE subtest performance within the small volume corrected mask, such that increased response to pseudowords relative to unrelated words within a right rostral middle frontal gyrus region was negatively associated with phonemic decoding efficiency \((r= \text{-0.72, } p<0.01)\). A negative association between these variables was also observed among the other groups as well, as illustrated in Figure 2.7, although none of those correlations were statistically significant. The RD group also only had one significant correlation between these factors in the pars opercularis region of the left inferior frontal gyrus, such that increased activation for semantically related words versus unrelated words in this region was positively associated with phonemic decoding efficiency in this cluster \((r=0.77, \ p<0.05)\).

There were numerous significant correlations between sight word efficiency performance and neural response to a variety of contrasts among the BL group, however many of these significant correlations were largely driven by an atypical observation with a superior score on the behavioral measure (SS=142) and extreme neural responses to certain contrasts as compared to other group members. Even when accounting for these atypical observations, the negative correlation between this measure and the semantically-related vs. unrelated contrast, as shown in the right inferior frontal gyrus cluster in Figure 2.7, remained statistically significant \((r= \text{-0.77, } p<0.001)\), as were all of the other reported results. This same relationship was also observed within bilateral rostral middle frontal gyrus clusters \((rs= \text{-0.76, } ps<0.01)\), and increased activation to unrelated words versus false font was associated with sight word efficiency in right superior temporal gyrus \((r=0.78, \ p<0.01)\), right inferior frontal gyrus \((r=0.78, \ p<0.05)\), and left precentral gyrus \((r=0.79, \ p<0.05)\) clusters.

Among ELs, clusters needed to exceed a correlation coefficient of \(r=0.83\) due to the small sample size. Only one cluster emerged that exhibited a significant relationship that was robust to the exclusion of outlying observations. As illustrated in Figure 2.7, increased neural response to pseudowords as compared with unrelated words in a left precentral gyrus region was positively associated with sight word efficiency (TOWRE SWE) among ELs \((r=0.89, \ p<0.05)\). This relation remained statistically significant upon removal of the participant with a much lower score than the others (SS=75).

Few of these regions remained statistically significant at a more stringent voxelwise threshold of \(p<0.001\), and none would have survived correction for multiple comparisons, given the large number of comparisons performed (50) and small number of participants in each group.
Table 2.5.

Clusters Showing Significant Brain-Behavior Correlations Overall (All Participants) and Within Each Group (voxelwise $p=0.01$, cluster corrected at $\alpha=0.05$)

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Behavioral Measure</th>
<th>Region</th>
<th>Talairach Coordinates of Peak</th>
<th>Voxels</th>
<th>Mean $r$ value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L/R Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Participants (threshold: $r=0.38$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O+P- minus O+P+</td>
<td>TOWRE-2 PDE</td>
<td>L Inferior Frontal Gyrus (pars opercularis)</td>
<td>-49 9 14 514</td>
<td>r=0.43</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>ML (threshold: $r=0.66$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSW - UNREL</td>
<td>TOWRE-2 PDE</td>
<td>R Rostral Middle Frontal Gyrus</td>
<td>37 58 0 393</td>
<td>r=-0.72</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>RD (threshold: $r=0.71$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEM - UNREL</td>
<td>TOWRE-2 PDE</td>
<td>L Inferior Frontal Gyrus (pars triangularis)</td>
<td>-54 36 -3 342</td>
<td>r=0.77</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>BL (threshold: $r=0.71$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEM - UNREL</td>
<td>TOWRE-2 SWE</td>
<td>R Inferior Frontal Gyrus (pars triangularis and pars opercularis)</td>
<td>33 34 11 704</td>
<td>r=-0.77</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>SEM - UNREL</td>
<td>TOWRE-2 SWE</td>
<td>R Rostral Middle Frontal Gyrus</td>
<td>20 33 22 614</td>
<td>r=-0.76</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>SEM - UNREL</td>
<td>TOWRE-2 SWE</td>
<td>L Rostral Middle Frontal Gyrus</td>
<td>-26 12 35 401</td>
<td>r=-0.76</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>UNREL - FF</td>
<td>TOWRE-2 SWE</td>
<td>R Inferior Frontal Gyrus (pars opercularis)</td>
<td>47 8 10 415</td>
<td>r=0.78</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>UNREL - FF</td>
<td>TOWRE-2 SWE</td>
<td>R Transverse Temporal Gyrus</td>
<td>35 -30 16 304</td>
<td>r=0.78</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>UNREL - FF</td>
<td>TOWRE-2 SWE</td>
<td>L Precentral Gyrus</td>
<td>-18 -7 57 300</td>
<td>r=0.79</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>EL (threshold: $r=0.83$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSW- UNREL</td>
<td>TOWRE-2 SWE</td>
<td>L Precentral Gyrus</td>
<td>-39 9 25 330</td>
<td>r=0.89</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Voxels are 2 mm isotropic (8 cubic mm) in size, * $p<0.05$, ** $p<0.01$, ***$p<0.001$
Figure 2.7. Select correlations between behavioral measures and contrasts of interest within small-volume-corrected mask (see Figure 2.5 for extent of analysis). The key in the bottom right corner of the figure provides color/group correspondence for the scatter plots (left) as well as the color/r-statistic correspondence for the cluster presented on the rendered template brain in each figure (right). The thicker, solid lines and larger dots represent the groups with the statistically significant brain-behavior relationships, while other groups are displayed in the background with smaller dots and dashed lines for reference. See Table 2.5 for details and additional brain-behavior correlations.
REFERENCES


Berken, J. A., Gracco, V. L., Chen, J. K., Watkins, K. E., Baum, S., Callahan, M.,


Bracken, B. A., & McCallum, R. S. (2016). *Universal Nonverbal Intelligence Test*. Austin, TX: PRO-ED.


developmental dyslexia. *Neuron, 44*(3), 411-422.


Geva, E., & Herbert, K. (2013). Assessment and interventions for English


http://dx.doi.org/10.1037/a0027238


Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz,


Langer, N., Benjamin, C., Becker, B. L., & Gaab, N. (2019). Comorbidity of


https://doi.org/10.1016/j.neuroimage.2009.12.009


Logan, B. R., & Rowe, D. B. (2004). An evaluation of thresholding techniques in...


[https://doi.org/10.1101/502625](https://doi.org/10.1101/502625)


Proficiency?. Society for Research on Educational Effectiveness.

Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L.


Qualtrics. Copyright © [2016] Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. https://www.qualtrics.com

Qualtrics. Copyright © [2018] Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. https://www.qualtrics.com


Rivera, M. O., Moughamian, A. C., Lesaux, N. K., & Francis, D. J. (2009). Language and Reading Interventions for English Language Learners and English Language Learners with Disabilities. *Center on Instruction*.


Roid, G. H., & Koch, C. (2017). Leiter-3: Nonverbal cognitive and


