

2022

Third Grade Arabic and English Dual Language Learners' Literacy Achievement Associated With the Scott Foresman Reading Street Program

Katina Latrice Walton
Walden University

Follow this and additional works at: <https://scholarworks.waldenu.edu/dissertations>



Part of the [Education Commons](#)

This Dissertation is brought to you for free and open access by the Walden Dissertations and Doctoral Studies Collection at ScholarWorks. It has been accepted for inclusion in Walden Dissertations and Doctoral Studies by an authorized administrator of ScholarWorks. For more information, please contact ScholarWorks@waldenu.edu.

Walden University

College of Education and Human Sciences

This is to certify that the doctoral dissertation by

Katina L. Walton

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee

Dr. Nicolae Nistor, Committee Chairperson, Education Faculty

Dr. Billie Andersson, Committee Member, Education Faculty

Dr. Beate Baltes, University Reviewer, Education Faculty

Chief Academic Officer and Provost

Sue Subocz, Ph.D.

Walden University

2022

Abstract

Third Grade Arabic and English Dual Language Learners' Literacy Achievement

Associated With the Scott Foresman Reading Street Program

by

Katina L. Walton

Ed.S., University of South Alabama, 2007

MS, University of South Alabama, 2004

BS, University of South Alabama, 2001

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education

Walden University

August 2022

Abstract

The academic success of dual language learners (DLLs) has been studied extensively, but few studies have investigated the effectiveness of the Scott Foresman Reading Street (SFRS) program on DLL struggling readers. The purpose of this quantitative study, grounded in the linguistic interdependence theory, was to compare the Measure of Academic Progress Reading Fluency (MAP RF) Rasch Unit (RIT) scores of third grade DLL struggling readers before and after they participated in 1 year of the SFRS program at Commonwealth School in Riyadh, Saudi Arabia. The RIT scores measure students' progress between grades. A repeated-measures *t* test was used to examine changes in MAP-RF test scores after 1 year of SFRS program participation. Archived data from 142 third grade DLL struggling readers after 1 year of SFRS in 2015, 2016, and 2017 were used. The 2016–2017 academic year focused on the intervention group of DLL students with reading difficulties. This group consisted of students who had previously scored between 69% and 60% on the second grade reading test. The findings revealed the final MAP RF scores for male and female DLL struggling readers after the implementation of the SFRS program were significantly higher than those at the start of the year. The discovery of a positive relationship between second language reading fluency and MAP RF scores shows that the SFRS program may influence explicit fluency instruction. Furthermore, this finding could be used to improve educational quality, including academic success, decision making, problem-solving abilities, empathy, and employment prospects, which could lead to positive social change.

Third Grade Arabic and English Dual Language Learners' Literacy Achievement

Associated With the Scott Foresman Reading Street Program

by

Katina L. Walton

Ed.S., University of South Alabama, 2007

MS, University of South Alabama, 2004

BS, University of South Alabama, 2001

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education

Walden University

August 2022

Acknowledgments

This dissertation is dedicated to my son, Justice M. Walton. I am thankful for all the encouragement and love you always provide in my life. I hope I have inspired you always to aim to make the world a better place and to realize that despite many obstacles, you should never give up on your aspirations. I am also very grateful to my parents, Jimmie and Carolyn Walton, for all the love, support, and encouragement they have provided me throughout my life and while writing this dissertation. I appreciate the encouragement and unconditional love offered to me by my sister and brothers, Varion Walton, Ganester Walton, and Ronald Wesley. An enormous thank you goes to my Dissertation Chair Dr. Nicolae Nistor, whose feedback, knowledge, and support was critical. Committee Member Dr. Billie Andersson, your kind words and excitement for reading were invaluable. Most importantly, this dissertation is dedicated to all readers.

Table of Contents

List of Tables	iv
Chapter 1: Introduction to the Study.....	1
Background.....	3
Problem Statement.....	8
Purpose of the Study.....	10
Research Questions.....	11
Theoretical Framework.....	11
Nature of the Study.....	12
Definitions.....	12
Assumptions.....	13
Scope and Delimitations.....	14
Limitations.....	15
Significance.....	15
Summary.....	16
Chapter 2: Literature Review.....	17
Literature Search Strategy.....	18
Theoretical Foundation.....	18
Literature Review.....	25
Dual Language Learners.....	25
Reading Literacy Achievement.....	28
Summary and Conclusions.....	37

Chapter 3: Research Method.....	39
Research Design and Rationale	39
Methodology.....	40
Population	40
Sampling and Sampling Procedures	41
Procedures for Recruitment, Participation, and Data Collection.....	41
Instrumentation and Operationalization of Constructs	42
Validity	45
Reliability.....	46
Data Analysis Plan.....	46
Threats to Validity	48
Ethical Procedures	50
Summary.....	51
Chapter 4: Results.....	52
Data Collection	52
Results.....	53
Analysis Results for Research Question 1	53
Analysis Results for Research Question 2	57
Chapter 5: Discussion, Conclusions, and Recommendations.....	62
Discussion.....	62
Research Question 1	62
Research Question 2	64

General Discussion	65
Conclusion	69
Limitations and Recommendations for Future Research.....	69
References.....	73

List of Tables

Table 1. 2015 NWEA Measure of Academic Progress Normative Data..... 6

Table 1. 2020 NWEA Measure of Academic Progress Normative Data..... 7

Table 3. Descriptive Characteristics of Pretest Reading Scores for Third Grade DLL
Struggling Reader Female Students..... 53

Table 4. Descriptive Characteristics of the Reading Scores and MAP RF Scores of Third
Grade DLL Struggling Reader Female Students Before and After Participation in the
SFRS Program 54

Table 5. Statistical Test Results 55

Table 6. Repeated Measures t test for MAP RF Score Changes for the Third Grade DLL
Struggling Reader Female Students After 1 Year Participation in the SFRS Program
..... 57

Table 7. Descriptive Analysis of Pretest Reading Scores for Third Grade DLL Struggling
Reader Male Students 58

Table 8. Descriptive Characteristics of the Reading Scores and MAP RF Scores of Third
Grade DLL Struggling Reader Male Students After Participation in the SFRS
Program..... 59

Table 9. Statistical Test Results 60

Table 10. Repeated Measures t test for MAP RF Score Changes the Third Grade DLL
Struggling Reader Male Students After Participation in the SFRS Program 61

Chapter 1: Introduction to the Study

Education is a critical component of a country's economic growth and development. Education encourages a flourishing society and influences a country's ideals, goals, and economic potential. Education policy, on the other hand, is one of the most challenging areas to modify globally. Countries across the world have increased their use of several languages in the early years of primary school via the introduction or expansion of education programs. These regulations have encouraged pupils to use a familiar language for the first few years of school before moving to an official language used in other public institutions. Many international students grow up as dual language learners (DLLs) who simultaneously acquire their heritage language and the language of instruction; their individual language learning histories and home language exposure vary (Heppt & Stanat, 2020; Simon-Cereijido & Mendez, 2020). In high-income Western countries, children from immigrant backgrounds often score lower on reading on standardized tests (Cuticelli et al., 2016; Foorman et al., 2018; Lathouras et al., 2019). Societal language skills are related to literacy and academic performance (Bleses et al., 2016; Prevoo et al., 2016). Willard et al. (2020) emphasized that addressing the needs of DLLs is vital because the number of children with immigrant ancestry in affluent nations is expanding.

For all students, dual academic language proficiency can strengthen their chances of reaping the academic, cognitive, and social benefits of becoming biliterate (Bialystok, 2018). The Ministry of Education (2020) of Saudi Arabia revealed Saudi Arabia's Vision 2030 emphasizes the crucial relevance of an education system linked with market needs

that produces economic opportunities. Furthermore, Mitchell and Alfuraih (2017) stated the Kingdom's goal is to develop creative and critical thinkers with entrepreneurial skills and knowledge who can thrive in an increasingly globalized environment. Improvements and breakthroughs in the educational environment of Saudi Arabia's government schools are closely linked to the growth of English language lessons (Al-Zahrani & Rajab, 2017).

Every year, many children are born into families in which they are expected to learn to read and write in more than one language. According to Galloway et al. (2020) and Heppt and Stanat (2020), many international students grow up as DLLs studying their heritage language and the language of instruction simultaneously. DLL programs focus on developing ideas in the child's native and second language. Understanding classroom directions and engaging in classroom discussions require learning the mainstream language suitable for daily interactions. In the current study, I investigated the instructional strategies used to instruct DLL struggling readers. The study was intended to explore the implementation of the Scott Foresman Reading Street (SFRS) program on Measures of Academic Progress Reading Fluency (MAP-RF) scores of third grade DLL struggling readers at Commonwealth School (a pseudonym) in Riyadh, Saudi Arabia. Commonwealth School offers an international education for learners in kindergarten through Grade 12, culminating in the American Diploma in Years 9 through 12.

This research may promote positive social change in addressing student progress associated with implementing the SFRS program with third grade Arabic and English DLL struggling readers. A thorough review of the literature indicated a lack of

independent (defined as noncommissioned, noncontracted) research on the Reading Street program. Since 2005, Pearson, the publisher of Reading Street, has commissioned an official project report of the Reading Street curriculum from independent consulting organizations such as Gatti Evaluation, Magnolia Consulting, and Claremont Graduate University (Berry et al., 2009; Gatti, 2005; Wilkerson et al., 2006, 2007), none of which focused on DLL struggling readers. Independent studies focusing on DLLs conducted by organizations not associated with or contracted with Pearson appeared to be lacking in the literature at the time of the current study, and independent research was needed to assess the effectiveness of Reading Street. The purpose of this quantitative study was to investigate the differences in MAP RF score changes before and after third grade DLL struggling reader students participated in the SFRS program for one school year at Commonwealth School in Riyadh, Saudi Arabia. I examined common barriers in education for DLL struggling readers' reading and language capacity. To explore Cummins's interdependent hypothesis, I examined the connections between literacy skills and the influence of the implementation of the SFRS program on DLL struggling readers' literacy achievement.

Background

Multiple barriers to high-quality education for struggling DLLs continue to exist as the field of education works to achieve this goal. van Staden (2016) noted that the National Reading Panel had recommended effective instruction in phonics, phonemic awareness, fluency building, vocabulary, and text comprehension for struggling readers. Language instruction and explicit and appropriate assessment are characteristics of

programs that produce the best outcomes for DLLs (Collier & Thomas, 2017; Cuticelli et al., 2016; Dillon et al., 2020; Foorman et al., 2018; Gao et al., 2020; Lathouras et al., 2019). Providing successful interventions with partner-based and computer-assisted training, boosting sociocultural interaction, or a combination of these strategies may be beneficial in meeting the requirements of DLLs (Alanis, 2018; Moon et al., 2017; O’Neal et al., 2019; Zimmerman et al., 2019). Although DLLs vary in reading ability and language capacity, Acosta et al. (2019) emphasized educators must identify and focus on reading programs to address individual deficiencies. Since the early 1970s, researchers have attempted to link the relative growth of L1 (heritage language) and L2 (the language of instruction). Adopting bilingualism as a strength necessitates rethinking preconceived notions about the influence of language use in an educational setting (Arif & Abdullah, 2017; Kim & Piper, 2019). According to Cummins (1979), the interdependent hypothesis postulates that a certain level of first language competence can be favorably transmitted to second language processing. According to Kim and Piper (2019), regardless of language status (L1 or L2), the language in which literacy abilities are acquired aids literacy learning in other languages.

The SFRS program is an interactive English language arts program for kindergarten through sixth grades, emphasizing a student-centered approach to reading, writing, speaking, listening, and thinking. Each of the evidence-based competencies that students need to read effectively (phonological awareness, phonics, vocabulary, comprehension, and fluency, as identified by the National Reading Panel) are taught explicitly and systematically (SAAVAS, 2021). In addition, SAAVAS (2021) stated that

authentic literature, highly engaging trade books, collaborative learning, and project-based inquiry are used to teach and practice 21st-century thinking and social and emotional development competencies. The collaborative learning, social, and development competencies connect to the interdependent hypothesis by Cummins (1979), which proposes that the development of competence in L2 is partially a function of the type of competence already developed in L1 at the time when intensive exposure to L2 begins. The current study may fill the gap in the literature on the influence of reading programs, which could be a successful component of student achievement in literacy (see Cuticelli et al., 2016; Lathouras et al., 2019; Moon et al., 2017).

Table 1*2015 NWEA Measure of Academic Progress Normative Data*

Grade	Begin-year		Mid-year		End-year	
	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>
K	141.0	13.54	151.3	12.73	158.1	12.85
1	160.7	13.08	171.5	13.54	177.5	14.54
2	174.7	15.52	184.2	14.98	188.7	15.21
3	188.3	15.85	195.6	15.14	198.6	15.10
4	198.2	15.53	203.6	14.96	205.9	14.92
5	205.7	15.13	209.8	14.65	211.8	14.72
6	211.0	14.94	214.2	14.53	215.8	14.66
7	214.4	15.31	216.9	14.98	218.2	15.14
8	217.2	15.72	219.1	15.37	220.1	15.73
9	220.2	15.68	221.3	15.54	221.9	16.21
10	220.4	16.85	221.0	16.70	221.2	17.4811
11	222.6	16.75	222.7	16.53	222.3	17.68

Note. MAP RF scores range from 141.0 (kindergarten) to 221.2 (11th grade). Northwest

Evaluation Association (NWEA, 2020) norms enable researchers to compare student achievement in a single term (a “status norm”), as well as across terms (a “growth norm”). The standard deviations of the means are shown. The lower the *SD*, the more compact the scores are around the mean. Adapted from

MAPGrowthNormativeDataOverview.pdf, by NWEA (2020)

(<https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf>). Copyright

[2020] by NWEA.

Table 2*2020 NWEA Measure of Academic Progress Normative Data*

Grade	Fall		Winter		Spring	
	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>
K	136.65	12.22	146.28	11.78	153.09	12.06
1	155.93	12.66	165.85	13.21	171.40	14.19
2	172.35	15.19	181.20	15.05	185.57	15.49
3	186.62	16.65	193.90	16.14	197.12	16.27
4	196.67	16.78	202.50	16.25	204.83	16.31
5	204.48	16.38	209.12	15.88	215.36	16.03
6	210.17	16.46	213.81	15.98	215.36	16.03
7	214.20	16.51	217.09	16.21	218.36	16.38
8	218.01	17.04	220.52	18.73	221.40	19.03
9	218.90	19.02	220.52	18.73	221.40	19.03
10	221.47	17.92	222.91	17.81	223.51	18.20
11	223.53	17.73	224.64	17.80	224.71	18.50
12	223.80	19.32	223.85	21.21	224.33	23.08

Note. MAP RF scores range from 136.65 (kindergarten) to 224.33 (12th grade). NWEA

(2020) norms enable researchers to compare student achievement in a single term (a “status norm”), as well as across terms (a “growth norm”). The standard deviations of the means are shown. The lower the *SD*, the more compact the scores are around the mean.

Adapted from MAPGrowthNormativeDataOverview.pdf by NWEA (2020)

(<https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf>). Copyright

[2020] by NWEA.

Problem Statement

The issue that prompted a literature search was the low 186 or lower MAP RF standardized test score for third grade DLL struggling readers in Riyadh, Saudi Arabia. Table 1 provides an easy-to-understand interpretation of the standards beginning with determining the mean score closest to a student's RIT score in each chart. This indicates the student's performance in comparison to on-grade level students nationally.

According to NWEA (2020), MAP RF scores can give insight into the growth of all children, including low and high performers. The NWEA MAP norms are typical of all schools across the world. Furthermore, stated norms as defined by NWEA (2020) allow schools to compare student success within a single term (a "status norm") as well as between terms (a "growth norm"). Growth norms determined by NWEA (2020) are used to project the amount of growth typical of comparable students, set suitable growth targets, and evaluate observed growth in the next term based on the projection.

In August 2015, Commonwealth School in Saudi Arabia began the implementation of the SFRS program for all elementary students. Before the implementation, the district evaluated and critiqued the reading performance of DLL struggling readers. The district selected the SFRS program because of its inclusive intervention component to increase the reading performance of struggling DLL readers at Commonwealth School in Riyadh, Saudi Arabia.

Kim and Piper (2019) investigated the effects of literacy skills in multilingual contexts and whether the structure of the relationships differed depending on the literacy instruction set. O'Neal et al. (2019) examined individual versus classroom peer effects on

grit as predictors of later literacy achievement among DLLs. Lathouras et al. (2019) sought to evaluate the long-term reading outcomes of the basal book whole-class intervention versus the PrepStart intervention on Australian students. However, researchers had not documented the literacy achievement of Arabic to English DLL struggling readers.

The Kingdom changed its education system to guarantee the realization of Saudi Arabia's Economic Vision 2030. The Kingdom required an educational reform package that focused on various changes, including improving schools' teaching techniques and strategies (Allmnakrah & Evers, 2020). Teaching core subjects in the English language has become a demonstration of a nation committed to the educational process to cultivate future citizens who are academically prepared and possess critical thinking skills. However, as Mitchell and Alfuraih (2017) indicated, much more can be done to promote continuous progress in English learning and instruction as the Kingdom continues to move into the industrialized world, relying less on oil and becoming a knowledge-based society.

Cuticelli et al. (2016) used a concurrent multiple-baseline single-case design to examine what effect performance feedback has on the number of opportunities to respond provided during core reading instruction by primary grade teachers. The Common Core reading program was Scott Foresman Reading Street. The findings in this study indicated that performance feedback could positively influence teachers' use of opportunities to respond; therefore, performance feedback is helpful as a behavioral intervention and

coaching strategy for academic settings. In addition, the findings indicated that providing practical, evidence-based Tier 1 reading instruction could prevent reading difficulties.

Arif and Imran's (2017) primary goal was to examine Iraqi English as a foreign language (EFL) secondary school students' metaphorical comprehension in L1 and how it affects their metaphorical comprehension in L2. Arif and Imran's research was significant in the literature because it was one of the first to address the influence of L1 on L2 using Arabic and English. Furthermore, it had the potential to shed light on Arabic and English DLLs by analyzing whether the nature of literacy instruction (whether DLLs received explicit and systematic training) is associated with their literacy achievement. Researchers had not documented the literacy achievement of Arabic to English DLL struggling readers, so the current study was conducted to address this gap.

Purpose of the Study

The purpose of this quantitative study was to compare the MAP RF scores of third grade DLL struggling readers before and after participating in the SFRS program at Commonwealth School for one school year. As NWEA (2020) noted, an RIT score evaluates a student's instructional level and quantifies the predicted advancement or growth at each grade level. Students at Commonwealth who did not make an RIT score of 188.3 on the fall 2016 assessment were divided into reading intervention groups based on their low RIT score. Students who scored between 185 and 190 were placed in a reading intervention group that participated in a single additional reading intervention lesson embedded within the SFRS program. Students with beginning year scores of 180–175 and below were placed in a reading intervention group that participated in two

additional reading interventions embedded in the SFRS program. This group was labeled the intensive reading interventions group. Boys and girls share classrooms exclusively in kindergarten and can attend the same school up to the third grade. In Grades 1–3, girls and boys are segregated in all single-gender classes but remain in the same building. Once boys enter fourth grade, they are taught exclusively on the only male side of the campus. All teachers were trained in instruction and intervention of the SFRS program.

As NWEA (2020) noted, an RIT score estimates a student’s instructional level and measures student progress or growth expected in each grade level. A repeated measures *t* test was used to investigate score changes associated with the interventions of SFRS by comparing second grade reading scores (pretest) of DLLs to the MAP RF 2017 test score (posttest) after SFRS program implementation of the same 210 DLLs.

Research Questions

RQ1: What is the difference in MAP RF score changes before and after third grade DLL struggling reader female students participated in the SFRS program for one school year?

RQ2: What is the difference in MAP RF score changes before and after third grade DLL struggling reader male students participated in the SFRS program for one school year?

Theoretical Framework

The theoretical foundation in this study was grounded in the linguistic interdependence hypothesis (LIH). Cummins (1979) proposed that the development of competence in L2 is partially a function of the type of competence already developed in

L1 when intensive exposure to L2 begins. Cummins addressed bilingual children's academic and cognitive performance, and the proposed theoretical paradigm highlighted the connection between sociocultural, linguistic, and school program influence. The relation between context, child feedback, and educational intervention factors explained educational outcomes in a DLL education model. Likewise, Kim et al. (2018) noted that if instruction targets metalinguistic awareness, such as phonological awareness and the alphabetic principle, explicitly capitalizing on the link between target languages, this could facilitate transfer to a greater extent than instruction without such intentional attention.

Nature of the Study

The study included a pre-post research design with a repeated measures *t* test to investigate score changes associated with the intervention of third grade DLL struggling readers at a Riyadh, Saudi Arabia Commonwealth School. The repeated measures *t* test was used to investigate the statistical significance of the score changes by comparing DLLs' second grade reading scores (pretest) to the MAP-RF standardized 2017 test scores (posttest) after the SFRS program implementation of approximately 210 struggling reading DLL students.

Definitions

Cummings linguistic interdependence: Children's second language (L2) proficiency is, to some extent, a function of their first language (L1) competence (Kim & Piper, 2019).

Dual language learners: Any young child learning two or more languages to become linguistically and academically proficient in two languages of instruction (Alanis, 2018).

MAP reading assessment: Used to measure a student's growth in reading (NWEA, 2016).

Scott Foresman Reading Street curriculum: A program that has the five Reading First principles embedded (Ladnier-Hicks et al., 2010).

Assumptions

For this study, I made various assumptions about the interventions and implementation fidelity at the study location. I assumed that the teachers responsible for instructing students in the direct instruction corrective reading program strictly adhered to the scripted lessons and suggested reading plan established by the SFRS program. As directed by SAAVAS (2021), intervention requires students to receive at least 45 minutes of instruction per day for 9 to 10 weeks. The students who completed the Reading Street online computer-assisted reading writer's notebook and other online resources completed two to five tasks during each 45-minute session per week, as suggested by the research school site. All students who participated in the direct instruction program read at least two grade levels behind, and all students who received only the computer-assisted reading program read one grade level behind. Teachers reasoned that those pupils reading significantly below grade level required intensive reading intervention that included direct instruction and computer-assisted training.

Additionally, I assumed that each student who participated in the interventions was wholly engaged and motivated to perform optimally during the treatments and evaluation. I anticipated that any archived data scores would reflect accurate score information at their creation and would contain no errors. The final assumption for this research was that all interventions and assessments had been produced, reviewed, and pilot-tested for effectiveness. The school system delivered all interventions and assessments in this study. Therefore, indications of reliability and validity were believed to be acceptable.

Scope and Delimitations

The sample for this study included third grade DLL struggling readers with low reading achievement. The study focused on DLL struggling readers who attended a school in Saudi Arabia in the District of Riyadh. The study was limited to a single school district, one elementary school, and struggling readers' third grade reading achievement scores.

Depending on the different reading abilities of the students, the students who needed the most intense reading interventions received both direct and computer-assisted instruction. This group consisted of 210 students. Students who needed a less intense reading intervention participated in a computer-assisted instruction intervention. Student reading achievement scores included in the study were scores of male and female third grade struggling readers.

Limitations

Changes in the testing location or learning environment between the pretest and posttest constituted a potential limitation for this study. The sample size of 210 students limited the generalizability of student outcomes to other district schools of similar sizes and demographics. Other limitations may have included attrition due to a student's illness or weariness during an evaluation. Another possible limitation was that the pretest and posttest were not administered to the same students during the observation period. A causality limitation existed in the fact that the research design of this study (pre-post design without reference group) would not support causal conclusions, so it would not be clear whether the implementation of the SFRS program was the cause of score changes.

Significance

There was some research on the academic achievement of DLL students. However, there was very little research on score changes associated with the SFRS intervention on the academic achievement of DLLs in Riyadh, Saudi Arabia. This study was significant in that it may fill the gap in the literature on the influence of reading programs, which have been a successful component of student achievement in literacy (see Cuticelli et al., 2016, Kim & Piper, 2019, Lathouras et al., 2019, Moon et al., 2017). The current study could promote positive social change if significant progress is associated with the SFRS program. This method could be applied to improve education quality and promote positive social change.

Summary

The purpose of this study was to determine whether there were score changes in third grade DLL struggling readers' student reading achievement, as measured by the MAP-RF assessment due to the first-year implementation of the SFRS program. Because global systems spend significant money to purchase curricula for school districts encouraged to use research-based materials, independent research documenting a program's effectiveness was necessary. This study could promote positive social change by strengthening student achievement and educational attainment and could be supplemented by further studies in different schools. Other school districts with an Arabic-to-English DLL population and approach to lessening the reading achievement disparity may also benefit from the findings of this study. Chapter 2 addresses this study's theoretical foundation and an extensive review of the literature.

Chapter 2: Literature Review

Researchers had investigated the academic achievement of DLLs. However, there was very little research on score changes associated with the implementation of SFRS on the academic levels of DLL struggling readers. Furthermore, few investigations on the SFRS program and score changes associated with DLL struggling readers' intervention had been conducted. Ladnier-Hicks et al. (2010) used the Stanford Achievement Test-10 to perform research in the state of Alabama in the United States. A modest improvement in student outcome data was discovered. No statistically significant differences in third grade children's test scores were observed before and after implementing the SFRS program. There were no noticeable studies on the SFRS program and Arabic and English DLL struggling readers. The issues that prompted the literature search were the low (186 or lower) MAP RF score for third grade DLL struggling readers in Riyadh, Saudi Arabia, and the school's desire to know whether the SFRS program improved students' outcomes.

Reading comprehension begins in early childhood education and is linked to daily educational tasks in classroom activities from prekindergarten through high school and beyond. The SFRS program is used in classrooms to explore the reading and writing connection. According to the United Nations Educational, Scientific and Cultural Organization (2017), in today's fast-changing world, everyone needs a broad set of competencies, including literacy and numeracy, to participate in social, economic, cultural, and civic life. The SFRS program offers a comprehensive continuum of research-based support to DLLs and their teachers from kindergarten through Grade 6.

These programs, founded on best teaching practices and research from leading language development experts such as Cummins, assist DLLs in preparing for academic success and success in today's global economy (SAVVAS, 2021). In the current quantitative study, I investigate the influence of the SFRS program on the MAP RF scores of third grade DLL struggling readers at a district school in Riyadh, Saudi Arabia. Kim et al. (2018) stated that future quantitative research could contribute to DLL literature by focusing on child outcomes and growth patterns across all development and learning domains. I attempted to fill a gap in the literature on the influence of reading programs that had been a successful component of student achievement in literacy (see Cuticelli et al., 2016; Lathouras et al., 2019; Moon et al., 2017).

Literature Search

Databases searched were Sage Journals, Science Direct, Taylor and Francis, Education Resources Information Center (ERIC), Association for the Advancement of Computing in Education (AACE), Elsevier, and ProQuest. Keywords used for the search of literature were *literacy, fluency, reading comprehension, assessment, Cummings, linguistics interdependence hypothesis, language 1 (L1), language 2 (L2), bidirectional, directionality, sociocultural, Scott Foresman Reading Street, Arabic, English, dual language learners, reading achievement, foundational skills, decoding, and NWEA MAP Reading Assessment.*

Theoretical Foundation

The LIH (Cummins, 1979, 2000) was the theoretical foundation for this study. LIH states that children's second language (L2) proficiency is, to some extent, a function

of their first language (L1). In addressing bilingual children's academic and cognitive performance, Cummins (1979) emphasized the link between sociocultural, linguistic, and school program influence. The findings indicated that context, child feedback, and educational intervention elements explain educational results in a DLL education model. This relationship was relevant to my study because implementing the SFRS program may explain DLL struggling readers' academic and cognitive growth. The LIH is based on the perceived fundamental linguistic properties of human languages, including those that are quite varied. According to the LIH, academic skills may be easily transferred from one language to another. Cummins et al. (2005) pointed out that transfer is predominantly based on conceptual and cognitive factors for different languages and is based on linguistic and conceptual elements for cognate languages. Cummins (1981, 2016) reasoned that learners transfer L1 skills to L2 when they have sufficient exposure to L2 (at school or in the environment) and are sufficiently motivated to study L1. Although languages differ in surface features (pronunciation, lexical differences), the LIH claims bilingual readers' skills in their first language predict language and literacy skills in their second language due to a common underlying proficiency that makes conceptual knowledge, cognitive skills, language, and literacy skills interdependent across languages (Cummins, 1981, 1991, 2016).

Cummins (1991) expanded further details on interdependence in L1 and L2 development. Cummins (1978, 1979) suggested two hypotheses to explain the inconsistent findings of several bilingual investigations. Cummins (1978, 1979) identified the first theory, the threshold hypothesis, which asserts that a child must achieve a

minimum linguistic competence in L1 to be protected from cognitive impairments. Cummins (1978, 1979) noted that if a child's proficiency in L1 is low, the proficiency level in L2 is likely to be low. If, on the other hand, the degree of competence in L1 is high, it may also project that the level of competence in L2 will be high. Likewise, Spies et al. (2018) said the threshold represents the critical degree of second language ability required to realize bilingualism's significant advantages through interaction between the second and native languages. Spies et al. compared the direct and indirect effects of L2 language and reading comprehension on L1 reading comprehension among Spanish-speaking English language learners in regular bilingual classes. In an analysis of L1 and L2 reading abilities in relation to metacognitive knowledge, Guo (2018) found that metacognitive knowledge indirectly influenced L2 reading via L1 reading ability and L2 language competency.

It is probable that the integrated nature of the intervention (i.e., vocabulary, oral retellings, read aloud, comprehension methods) altered how the linguistic factors improved reading comprehension (Spies et al., 2018). Similarly, Daller and Ongun (2018) analyzed data from 100 Turkish-English students and found that children whose parents used more L1 at home outperformed the monolingual control groups by having significantly higher IQ scores ($t = 6.3, df = 98, p < .001$). Both studies contributed to the threshold hypothesis by examining the intervention's integrated nature, parental support, and effect on bilingual children's cognitive development.

The second hypothesis identified by Cummins (1978, 1979) was that if the outside environment is enough to keep L1 alive, more exposure to L2 in school leads to

faster bilingual development with no adverse effects on L1. The debate on the effectiveness of using L1 and L2 in classrooms is ongoing. According to Almoayidi (2018), multilingual teachers must employ the L1 of each student's home country to clarify specific components of their language proficiency. In a study comparing 61 first graders' L1 Spanish literacy, Almoayidi found that those in the experimental group got EFL training throughout K-1, whereas those in the control group did not.

Educational and linguistic studies on literacy development in migration circumstances emphasized the significance of L1 languages. In a literature review on this topic, Riós and Castellón (2018) found that literacy abilities transfer from one language to another and that literacy in both L1 and L2 should be cultivated and supported. Lindahl and Sayer (2018) conducted a one-way independent samples comparison of mean scores on the literacy exam and found that participants in the experimental group who received EFL instruction scored considerably higher on all assessment sections than participants in the control group.

Lindahl and Sayer examined the competency, strategy, motivation, and reading abilities of 269 Iranian EFL students randomly selected to participate in a quantitative study. Cummins (1979) argued that all written languages retain a core cognitive competency. Lindahl and Sayer's (2018) results added to Cummins's understanding by indicating that metacognitive knowledge, an underlying competence, serves as an executive function by monitoring the use of L1 reading ability and L2 language proficiency to promote L2 reading. In another study, Relyea and Amendum (2020) showed that when L1 and L2 reading techniques (or strategies) are transferred to L2,

there must be substantial interdependence between L1 and L2 reading. Maghsoudi's (2021) multiple linear regression analysis revealed that proficiency accounted for a more significant proportion of the variance in reading scores than strategy or motivation in L2 learners accessing their L1 repertoire of reading methods. Later, Cummins expanded on the interdependence hypothesis in subsequent articles, indicating that a reciprocal interaction was involved.

Cummins (1980) proposed differentiation between fundamental interpersonal communicative language abilities and cognitive/academic language skills to understand how dependency between language skills occurs. Although reading comprehension and language skills are considered separate, they are somehow linked. Studies showed that hardware for reading in L2 and neural networks, which study how the brain works, are linked (Citron et al., 2020; Follmer et al., 2018). As Kim and Piper (2019) remarked, cognitive abilities such as reading share underlying resources across languages, and acquisition in one language enhances transfer to other languages; then the bidirectional transfer may occur.

Cummins (1981) elaborated on the hypothesis that training in a particular language increases competency in that language, and transfer to another language will occur providing sufficient exposure to the other language (either in school or in the environment) and sufficient drive to acquire it. Unless the exposure and incentive circumstances are unfavorable, the theory predicts transfer from L1 to L2. For instance, in a study of 124 eighth-grade German students, Maluch and Sache (2020) examined reading speed, text comprehension, and reading efficiency in L2 English and L1 German

students. The results indicated a stronger relationship between L1 and L2 reading speed for learners who read slower. After correcting for L2 reading competence, Maluch and Sache found a nonlinear relationship between L1 and L2 reading speed.

Khaghaninejad (2020) found significant relationships between the proficient readers' use of reading strategies in Persian and English in general ($r = 0.62, p < 0.05$). Moreover, Khaghaninejad revealed that L2 reading comprehension is not language specific but is primarily related to general reading proficiency. According to Kim and Piper (2019), bidirectional interactions and directionality in instructional contexts would be an extension of the main principles of the LIH from a theoretical standpoint. If reading abilities acquired in one language can be transferred to another, the instructional context in which reading is taught is critical. As Wawire and Kim (2018) and Kim and Piper (2019) emphasized, if critical reading component abilities (e.g., basic phonological awareness) are explicitly and routinely taught in a language (L1 or L2), the connection is assumed to be from systematic training to the other language.

I focused on the second hypothesis, the LIH, to investigate the score changes associated with interventions of the SFRS program on the MAP-RF scores of third grade DLL struggling readers at Commonwealth School in Riyadh, Saudi Arabia. To improve educational outcomes in L1 and L2, according to Almoayidi (2018), it is vital to teach letter sounds, decoding abilities, oral reading fluency, and reading comprehension, among other skills. As stated by Cuticelli et al. (2016), Kim et al. (2018), Khaghainejad (2020), Kim and Piper (2019), Lathouras et al. (2019), Moon et al. (2017), Oneal et al. (2019), and Wawire and Kim (2018), if instruction capitalizes on the link between target

languages, such as phonological awareness and the alphabetic principle, this may facilitate transfer to a greater extent than instruction without such intentional attention.

The SFRS program differentiates instruction, emphasizing ongoing progress monitoring and a detailed plan to help manage small groups of students. Students in the SFRS program are taught in small groups with continuous progress monitoring and a structured plan. According to Snow and Matthews (2016), children learn best when they engage with others in meaningful ways. The sociocultural, linguistic, and school program factors explain bilingual children's academic and cognitive development (Cummins, 1979). According to Daller and Ongun (2018), both productive and receptive vocabulary are significantly correlated ($r = .61$ for receptive and $r = .732$ for productive vocabulary), and both vocabularies develop concurrently. However, L2 appears to take over later in life due to school input. These observations support Cummins's (1979) interdependence hypothesis. Daller and Ongun found that the development of the lexicon in L1 positively influences the development of the lexicon in L2. Concerning L2 reading acquisition, Cummins predicted that vocabulary conceptual knowledge, metalinguistic insights regarding print functions, and processing decontextualized language were crucial.

Wawire and Kim (2018) conducted a randomized control experiment to investigate the causal evidence for cross-language transfer of phonological awareness and letter knowledge (names and sounds) among multilingual first-grade kids in Kenya. Wawire and Kim observed that explicit and systematic training improved phonological awareness and letter knowledge in the instruction language (Kiswahili), with impact sizes ranging from .60 to .92. The strategic pairing of students and planned paired learning

activities may enhance collaborative exchanges (Alanis, 2018). The SFRS program may explain DLL children's academic and cognitive growth, as measured by the MAP-RF standardized test. The current study may fill the gap in the literature on the influence of reading programs, which have been a successful component of student achievement in literacy (see Cuticelli et al., 2016, Guo, 2018; Kim & Piper, 2019, Lathouras et al., 2019, Moon et al., 2017). The current study may promote positive social change by improving student achievement and educational attainment, and more studies in various schools might reinforce it. The outcomes of this research may be helpful in other school districts with Arabic-to-English DLL struggling readers and a strategy for reducing reading success disparities.

Literature Review

Dual Language Learners

Diversity is on the rise globally. DLLs are students with at least one parent who speaks a language other than English. Van Norman and Parker (2018) documented that families speak many languages, identify with many races and ethnicities, and have widely varied countries of origin, all essential characteristics influencing DLL children's development. This diversity challenges effectively serving this unique, growing population. According to Goodrich et al. (2021), school employees face the problem of meeting the unique academic needs of multilingual pupils. According to previous research, multilingual students participating in dual language programs obtain the same or higher levels of English academic accomplishment as multilingual students enrolled in English-only programs (Bibler, 2021; Garcia, 2018; Hartanto et al., 2018; Raikes et al.,

2019; Steele et al., 2017). With growing diversity in various countries, instruction should be diverse. Larson et al. (2020) stated that it is critical to better understand development in various languages due to the increasing numbers of DLLs. The objective is for speakers to achieve academic proficiency in two instructional languages.

The central feature of dual language programs is the provision of literacy and content instruction in both English and a partner language, such as Spanish, for promoting bilingualism, biliteracy, grade-level academic achievement, and sociocultural competencies (Howard et al., 2018). DLL applications may be valuable to satisfy academic demands. As defined by Galloway 2020, academic languages include lexical, syntactic, and discursive resources that are often used to explain concepts in the disciplines and explore and negotiate understandings in academic communities. However, as Howard et al. (2018) point out, dual language program design, implementation, and quality vary substantially by context.

When evaluating the effectiveness of a program, it is critical to check DLL growth. Numerous descriptive studies have established beneficial associations between dual language program participation and academic attainment (see Bialystok, 2018; Howard et al., 2018 for reviews). Similarly, Goodrich et al. (2021) discovered that states with a more significant multilingual population had lower achievement gaps, and states with a higher percentage of multilingual pupils who speak Spanish had a stronger achievement gap decrease over time. Moreover, Kieffer and Thompson (2018) suggest that access to high-quality dual language instruction may be partly responsible for bi/multilingual pupils outpacing monolingual counterparts in reading development rates.

The two languages this study will focus on are Arabic and English. Arabic is the official language of 22 countries that stretch across the Arabian Peninsula, Syria, and northern Africa. Saudi Arabia's Vision for 2030 seeks to offer all individuals many chances for exceptional learning; enhance the quality and results of learning; upgrade the skills and talents of education workers; promote innovation and creativity; and create a curriculum (Saudi Arabia's Vision for 2030, 2017). Leaders have understood that long-term economic growth involves focusing on knowledge-based economies, not oil extraction, as noted by Makhoul (2021). (i.e., education). Due to changing demographics, dual language educators must adapt their curriculum to meet the requirements of second language learners enrolled in dual language programs. These curriculum revisions reflect an expansion in the number of DL programs devoted to developing young children's native languages and English (Alanis, 2018). According to Galloway et al., 2020, increased possibilities to absorb academic language materials may contribute to reading development in DLL contexts (when students use two languages for reading, writing, and learning). As documented by previous reviewers Bialystok (2018) and Soto et al. (2019), numerous descriptive studies have established beneficial associations between dual language program participation and academic attainment.

On language and literacy levels, academic success, and appropriateness, the review of Bialystok (2018) concludes that bilingual education has no adverse consequences and numerous positive benefits in many fields. As a result, Bialystok (2018) concluded that some studies included assessments of proficiency in the minority language, which is often the language of instruction (English for Hispanic children in the

United States, French immersion children in Canada, community language for indigenous language programs in the United States and elsewhere) (e.g., Spanish in the US, French in Canada, Maori in New Zealand). No data on Arabic and English students were included in this review.

In addition, Bialystok (2018) cites a prior study by Mondt et al. (2011) showing that teaching a topic in a specific language improves performance in that subject when assessed in that language. Additionally, Soto et al. (2019) assessed the effectiveness of phonological awareness treatments for dual language learners (DLL) from preschool through second grade. This review found that DLL children gain from multilingual phonological awareness education without losing English phonological awareness abilities. Although Soto et al. (2019) focused on Spanish and English, they called for future research to identify efficacious literacy-based interventions and programs mindful of DLLs' unique language, social, and academic needs. Researchers Soto et al. (2019) and Bialystok (2018) noted that DLL education requires a detailed description of the program's structure, the quality of the teaching, and the match between children's needs and abilities and the specific educational program offered.

Reading Literacy Achievement

Literacy Development

Children's development of early literacy skills begins well before formal schooling begins. Literacy development begins with two related skills: foundational and reading comprehension skills. Definitions by Foorman et al. (2016) state that foundation reading skills enable students to read words (alphabetic), relate those words to their oral

language, and read connected text with sufficient accuracy and fluency to understand what they read. Also noted by Piasta (2016), emergent literacy skills are related to understanding how print maps to language (code-focused skills such as phonological awareness and knowledge of the alphabet) and building meaning from text. As defined by Lucas et al. (2021), emergent literacy occurs when youngsters describe what their drawings represent; this may be considered an early stage of “reading.” DLLs begin school with different levels of emergent literacy. Another study by Carroll et al. (2019) noted that these varying levels of emergent literacy must depend on a combination of genetic and environmental factors. Some students are exposed to shared book readings with parents exposed to advertisements or street names. Additionally, as Puglisi et al. (2017) noted, formal literacy interactions, on the other hand, refer to activities in which adults directly teach reading or promote print-related skills at home (e.g., writing the child’s name, teaching letter names and sounds). Children’s culture or exposure to various parenting skills will vary; thus, they will enter their educational settings at various emergent literacy levels.

For example, students begin school with exposure to informal emergent literacy. Informal literacy practices (most shared book reading in the home), as identified by Puglisi et al. (2017), appear to be more closely associated with broad oral language skills, including vocabulary knowledge, and indirectly with reading comprehension later in development. A shared book reading integrates other language modes, such as listening and speaking. Whether formal or informal, exposure to emergent literacy helps children understand the printed text. Despite their restricted language abilities, DLLs must adhere

to the same academic rigor as their native English-speaking counterparts. According to Murphy and Torff (2019), the achievement gap becomes increasingly visible in DLLs, who do not receive the content level of instruction they tested on because of their limited English language proficiency. Overall, as noted by previous researchers (Carroll et al., 2019; Piasta, 2016; Puglisi et al., 2017; Soto et al., 2020), early literacy experiences before formal reading instruction can affect their later reading achievement. Lucas et al. (2021) observed that children describe what they depict in their drawings, which might be regarded as an early stage of “reading.” Emergent literacy integrates the reading and writing experience and can empower students to build their confidence in reading.

SFRS program integrates reading and writing by allowing students to actively participate in reading and writing. As Wang and Lin (2019) stated, integrated research on reading and writing benefited the students’ reading and writing skills and increased their motivation to learn. Students in this study used picture books to provide the opportunity to read for pleasure and write for fun. Similar findings on reading and writing connections in Leal (2016) found that reading children’s literature engaged learners in critical thinking, meaningful discussions, and creative writing. Although successfully conducted in a rural Taiwanese school, adapting, and implementing it in a different sociocultural setting would provide additional insights into integrating reading and writing to increase students’ motivation to read and write in English.

Foundational Skills Development in DLLs

DLLs actively engaged in reading and writing (emergent literacy) build on foundational skills to aid their reading and writing. Foundational skills develop in DLLs

during integrating reading and writing when utilizing a balanced approach to literacy. Williams and Lowrance-Faulhaber (2018) suggest that the balanced approach to literacy lessons, which integrates skills instruction within meaning-based activities and daily opportunities to write and talk about writing with teachers and peers, supports DLLs' literacy achievement. Additionally, Bauer et al. (2017) pointed out that oral language is used to plan their text before writing and scaffold the encoding process by segmenting the sounds in words. Reading and writing are related processes, and teachers can support DLL's growth and proficiency through emergent literacy by integrating reading and writing instruction. Likewise, Piasta (2016) identified that emergent literacy perspectives also emphasize the importance of contexts and experiences in supporting the development and preventing later reading difficulties.

There is a scarcity of research on English to Arabic DLL's emergent literacy. Nonetheless, as previously noted by Bauer et al. (2017), Williams and Lowrance-Faulhaber (2018), and Piasta (2016), DLLs' emergent literacy and understanding of basic literacy concepts serve as a foundation for their literacy development. Nevertheless, as noted by (Duran et al., 2016, Larson et al., 2020, and Soto et al., 2019), the bulk of research on emergent literacy has only been delivered in English to Latinx preschoolers.

Foundational Reading Skills as Predictors of Decoding

Alphabetic knowledge and phonemic awareness are two components of building foundational skills of DLLs for beginning decoding. Gutierrez et al. (2020) noted that alphabetic knowledge refers to the knowledge of alphabet names and sounds and grapheme-phoneme correspondence rules. DLLs practice this during various lessons in

the SFRS program by watching the teacher model writing letters, letter formation, and verbally saying the letter's sound. Within these lessons, students are introduced to writing as a communication tool. Furthermore, Gutierrez et al. (2020) noted that phonemic awareness is the ability to detect, manipulate, and analyze language at the phoneme level, independently of the meaning.

Evidence-based approaches promote the explicit teaching of phonological awareness, sound-spelling correspondences, and orthographic norms through integrated, systematic, organized phonological–orthographic education, as indicated by Williams et al., 2017. When teachers model alphabetic principles and phonemic awareness by teaching reading and writing, they model letter-sound relationships, print conventions, and ways students can read and write words. These skills support students' knowledge of foundational skills. Children learn their language's phonology before reading. According to Ijalba et al. (2020), they learn to distinguish phonemes, establish phonetic categories, identify phonemic segments in words, and combine phonemes to form words.

Teaching explicit instruction on the concepts of print creates space for multiple exposures to literacy. For example, Erickson and Whorton-McDonald (2019) noted that skills-based approaches to literacy instruction stem from behaviorism or the idea that complex processes can be broken into small chunks individually taught and mastered. The SFRS program offers the opportunity for these skills to be modeled in mini lessons in various units, according to SAAVAS (2021), and teaching students skills they recall during reading and writing helps students in decoding. At the same time, Erickson and Whorton-McDonald (2019) adapted activities to support students' need for competence

by including numerous scaffolds (e.g., partner work, visuals, teacher modeling) and options for an additional challenge (e.g., open sort, independent work). These ideas promote a sense of belonging and encourage DLLs to work together toward a similar goal of improving their literacy skills. Additionally, Lonigan and Burgess (2017) noted that decoding would be a stronger predictor of reading comprehension skills than language skills, as their sample included relatively young children (first- and second grade DLLs), and prior research indicates that decoding is the most important predictor of beginning readers' reading comprehension.

Foundational Reading Skills as Predictors of Reading Comprehension

Foundational reading skills are reading subskills taught to DLLs during classroom instruction. According to Foorman et al. (2016), foundational reading skills enable students to read words (alphabetic), relate those words to their oral language, and read connected text with sufficient accuracy and fluency to understand what they read. DLLs require these foundational skills to be taught explicitly during reading and writing when developing literacy. Similarly, Foorman et al. (2016) recommended students are taught academic language skills, including inferential and narrative language and vocabulary knowledge. The SFRS program is a way to teach these recommendations during reading and writing to model foundational literacy skills that could promote increased achievement gains. As noted by Gilanders (2018), DLLs have the advantage of being exposed to two or more languages and understanding the relationship between oral language and the written representation of words. When given multiple exposures to reading and writing, they become aware of specific characteristics in the written language

system. Likewise, Acosta et al. (2019) identified DLLs as linguistically diverse students who view their language abilities as assets instead of limitations. When we allow DLLs to draw from their background knowledge of their culture and connect it to what they learn in English, students use their two languages to write and understand English. For instance, Gilander (2017) states that DLLs follow a similar developmental process as monolinguals and use their knowledge of both language systems in their production.

Durgunolu (2017) highlighted similar underlying competency theories that emphasize the involvement of latent meta-linguistic skills (phonological, morphological, semantic, and discourse awareness), which are heightened during instruction. Others have expanded on the LIH, claiming that bilinguals' language abilities are reciprocally connected (rather than only L1-L2; MacSwan, 2017; Prevoo et al., 2016). When working with students, Sarisahin (2020) notes that it is vital to use research-based reading practices to enhance their reading comprehension abilities. Still, they must also support students' developing English proficiency.

DLLs use foundational reading skills or competencies gained during emergent literacy to build on their foundational reading skills. As noted by previous researchers, Acosta et al. (2019), Edyburn et al. (2017), and Paige et al. (2019), DLLs are more proficient when strong foundational literacy skills have been taught explicitly utilizing the blended approach to literacy.

MAP-RF

This study examines the NWEA's MAP RF score of third grade DLL students at Commonwealth School in Riyadh, Saudi Arabia. MAP-RF is an adaptive computer test

that measures students accurately regardless of whether they are above or below grade level. According to Soland and Sandilos (2021), due to the adaptive nature of the exam, students are presented with items taken from a wide item pool, whose difficulty corresponds to the best estimate of the student's success at that point in the assessment. Educators' understanding of data is inherently interpretative when reviewing log data from the NWEA's MAP tests. As Farley-Ripple et al. (2021) emphasize, assessment systems and their reporting features are instruments that mediate the process of information transformation into action. The tests are aligned to the Common Core Standards for Reading. Each test is administered for roughly 40 to 60 minutes and is typically given three times every academic year in autumn, winter, and spring.

Assessments are critical components of both teaching and learning. NWEA (2016) creates a personalized assessment experience that accurately measures performance whether a student performs on, above, or below grade level, vertical scaling of the assessment enables calculation of gains over time. Johnson (2020) described the RIT scale as used to report test results since it is a linear translation of the Rasch item response theory model's logit scale units. Commonwealth School in Riyadh, Saudi Arabia, began using the MAP-RF test in the 2014-2015 academic year to compare their students to other schools in their district and worldwide. Furthermore, NWEA (2016) uses anonymous assessment data from over 10.2 million students to create national norms. Educators compare their students' performance to national standards to evaluate programs and enhance instruction in individual classrooms and across school districts.

Scott Foresman Reading Street

DLLs can benefit from the SFRS program, produced by Pearson Education, Inc. The SFRS program is primarily intended for young learners, ranging from kindergarten to sixth grade. English language arts in the SFRS program adhere to the Common Core State Standards for English language arts. SAVAS (2021) states that its goal is to help students improve their English language abilities and knowledge. Gatti (2005) concluded that the Reading Street program closely matched national state criteria. In two-year-long independent studies commissioned by Pearson in 2006 and 2007, Magnolia Consulting performed under the leadership of Wilkerson, Shannon, and Herman. Berry et al. (2009) completed an official project report for Pearson to expand on the findings of the Wilkerson et al. tests from 2006 and 2007. The Pearson investigations laid the foundation for more Reading Street research. Additional independent non-commissioned research studies could be undertaken to ensure DLLs received a quality education.

The SFRS program is a coursebook based on multiple stories. The program has several intervention components to support the stories' main book. Grammar books, a spelling and word study book, and an assessment guide are components of this program. The core book separates into units, and each unit consists of five stories that contain both the to-be-learned vocabulary and the relevant word knowledge previously studied. Although a few studies on the Reading Street scripted curriculum have been conducted, research indicates that no studies on the program's effectiveness after implementation in a DLL community have yet been conducted. This study is notable in literature as it is one of the first to use Arabic and English to examine the influence of L1 on L2. It may

provide insight into Arabic and English DLLs by examining whether literacy education (when DLLs receive explicit and systematic training) is linked to literacy attainment. As previously said, there has been little to no research on the literacy accomplishment of Arabic to English DLL struggling readers, and this study might add to that knowledge.

Summary and Conclusions

The United Nations Educational, Scientific and Cultural Organization acknowledged global educational capabilities for social, economic, cultural, and civic participation in 2017. Several studies (Cuticelli et al., 2016; Lathouras et al., 2019; Moon et al., 2017) have documented successful reading programs for students' literacy achievement. Multilingual students enrolled in dual language programs perform as well as or better than multilingual children enrolled in English-only schools in terms of English proficiency (e.g., Bibler, 2021; Garcia, 2018; Hartanto et al., 2018; Raikes et al., 2019; Steele et al., 2017). Much of this research was, however, conducted in Spanish and English. The research discovered using Arabic and English was conducted at the university level only. There is a lack of research on Arabic and English DLL struggling readers at the elementary school level. As a result, there is little evidence linking dual language education to growth trajectories within and across primary grades.

As previously highlighted (Carroll et al., 2019; Piasta, 2016; Puglisi et al., 2017; Soto et al., 2020), literacy development can influence later reading achievement. As identified by Leal, 2016; Wang and Lin, 2019, emergent literacy stated that reading and writing integrated research benefited the students' reading and writing skills and increased their motivation to learn. There is a scarcity of research on English and Arabic

DLL struggling readers. Previous research on DLL's emerging literacy and knowledge of core literacy concepts by Bauer et al. (2017), Williams and Lowrance-Faulhaber (2018), and Piasta (2016) serves as a basis for their literacy development. DLLs employ foundational reading skills or competencies acquired throughout emergent literacy to expand on their basic reading skills. Acosta et al. (2019), Edyburn et al. (2017), and Paige et al. (2019) noted that DLLs are more proficient when foundational solid literacy skills are taught explicitly utilizing the blended approach to literacy.

This research may fill the gap in the literature on the influence of reading programs, which have been a successful component of student achievement in literacy, as identified by previous researchers (Cuticelli et al., 2016, Kim & Piper, 2019, Lathouras et al., 2019, Moon et al., 2017). This research study could promote positive social change if significant progress is associated with the SFRS program. Chapter 3 includes information on the research design, rationale, methodology, threats to validity, and ethical procedures for conducting this research study.

Chapter 3: Research Method

The purpose of this quantitative study was to investigate the differences in MAP-RF score changes before and after third grade DLL struggling readers participated in the SFRS program for one school year. As documented by previous researchers (Barber et al., 2020; Donnelly et al., 2020; Grover et al., 2020), studies of intervention components in literacy programs suggested that supporting DLLs' first language skills can contribute to their successful acquisition of a second language. The current study may supplement the existing research on literacy achievement by DLL struggling readers using literacy programs.

Numerous schools have implemented the SFRS curriculum because they employ research-based ways to increase reading. Howard et al. (2018) indicated that dual-language programs promote bilingualism, biliteracy, grade-level academic success, and sociocultural competence by teaching literacy and content in both English and a partner language. The academic achievement of DLLs has been the subject of previous research. However, there had been relatively little research on the implementation of SFRS to promote the academic achievement of DLL struggling readers. In Chapter 3, I describe the methodology used to achieve the study objectives. I present the research design, targeted population, data instruments, sample size and sampling techniques, data collection methods, operationalization of the study constructs, and data analysis.

Research Design and Rationale

I used a quantitative design to compare pre- and postscores before and after an intervention. This design was chosen over other designs because it helped me determine

the extent and direction of the relationship between dependent and independent variables. To evaluate interventions, it is most appropriate to use causal comparisons (i.e., to compare intervention subjects with subjects from a reference group in which no intervention was administered). However, for causal comparisons are not always possible. Instead, the pre-post approach is often used (Hillier et al., 2019; Merriman et al., 2019).

In the current study, all students were assessed before and after the implementation of the SFRS program. A comparison was made across time to explore the outcomes of the reading program. The scores from pretests were obtained from second grade report cards before implementing the SFRS in 2015, and posttest scores were retrieved from MAP-RF test conducted three times in 2017.

Methodology

Population

The target population of this study was third grade struggling readers at Commonwealth School in Riyadh, Saudi Arabia. Commonwealth School's total population of students was 2,400. There is an all-male and all-female side, but both are instructed in the SFRS program. As noted by Barry (2019) and Geel (2016), gender separation is the norm in all public domains in Saudi society. Grade 3 had 110 female and 100 male students at the time of the study.

The Common Core State Standards are used in the Commonwealth's international program. The Common Core State Standards emphasize students' critical thinking, problem solving, and analytical abilities. The requirements aim to prepare students for

entry-level jobs, first-year college courses, and relevant workforce training programs.

Commonwealth School is committed to retaining student identity during the transition to the international program by providing a rigorous educational curriculum in Islamic, Arabic, and Kingdom of Saudi Arabia history. The student population consists of Saudis preparing for further education at work, abroad, or in Saudi Arabia.

Sampling and Sampling Procedures

I conducted a power analysis to determine the needed sample size for this study. I used G*Power 3.1 software for Mac. For a one-tailed t test assessing the difference between two dependent means of matched pairs (repeated measures), the a priori power analysis expecting a moderate effect size (default .25), α error probability (default .05), and power ($1 - \beta$; default .95) resulted in a minimum sample size of 45. The research was a census study; therefore, no sampling was required. Because I used archival data, data collection time was not a concern.

Procedures for Recruitment, Participation, and Data Collection

I used archival data from the academic years 2015, 2016, and 2017 of third grade DLL struggling readers after completing one full year of the SFRS program. The archival data consisted of information for male and female second grade DLL struggling readers who had proceeded to third grade after completing one full year of the SFRS program. The site administrators preestablished the intervention groups for all students at the beginning of the 2016–2017 academic school year based on students whose second grade reading scores ranged from 69% to 60%. These students received instruction in the SFRS program the same way all students who attend Commonwealth School receive instruction

in the SFRS program. Students with 69%–60% scores were also placed in intervention in the SFRS program. Two hundred and ten students were designated to receive the built-in intervention plan in the SFRS program twice daily for 15 minutes each session. As Snow and Matthews (2016) highlighted, early childhood through third grade instructional and intervention programs are more effective at influencing limited skills or specifically targeted subdomains within unconstrained skills (e.g., the words taught in a vocabulary curriculum). The SFRS program was implemented for all students in Grades K-6. I assumed that teachers who teach all students in the reading program followed the SFRS program's planned lessons and prescribed reading plans for intervention. I used all available data for third grade DLL students who received one full year of instruction in the SFRS program. The MAP-RF assessment was initially administered to students in the fall of 2016 to establish baseline data. The fall 2016 MAP RF score was used to establish other reading intervention groups.

Instrumentation and Operationalization of Constructs

The Commonwealth School leadership team implemented a school improvement plan for struggling readers to improve their reading skills. At the same time, the leadership team started using the MAP reading assessment in all schools to determine individual students' reading ability. The MAP Growth evaluation from NWEA was used in this study. Johnson and Annenberg Institute for School Reform at Brown University (2020) identified that MAP Growth as a computer-adapted exam that gives exact measurements for pupils performing at or above grade level. The assessments are based on the content criteria set by the district.

The MAP (NWEA, 2011) is an adaptive computer exam of academic achievement delivered in groups. The MAP is intended to assess growth, anticipate student success, and influence training. Student progress between grades is measured using an RIT score, which is a stable, equal-interval scale that allows researchers and practitioners to compare student growth across grade levels (NWEA, 2011). A student's RIT score indicates the difficulty of test items that they can adequately answer around half of the time (NWEA, 2016). Additionally, according to the NWEA (2016), customizing test questions to individual levels of success rather than using a static bank of questions that might be above or below a student's ability level results in a more accurate estimate of student performance.

When providing assessment questions, MAP considers the complexity of each test item and each student's proven aptitude, as guided by item response theory (NWEA, 2016). The test questions are tailored to specific levels of success rather than a static bank of questions that may be above or below a student's ability level to achieve a more accurate assessment of student performance (NWEA, 2016). The calibration of the level test to the curriculum makes it easier to compute Rasch unit, $RIT = 0.1$ logit measures that directly associate learners' attainment to the curriculum, providing a linear measure of equal intervals. Therefore, it is possible to accurately sum up the RITs and obtain school or class averages. RITs range from 160 to 240. Students who scored 186 or lower on the MAP-RF assessment participated in additional intervention sessions embedded in the SFRS program.

The questions are chosen from a pool of questions set by a team of proficient instructors to cover a wide range of goals in reading. The questions, which have been field-tested with tens of thousands of learners, ensure quality test informational text understanding, foundational skills, and literature comprehension. The test questions for the comprehension capability of informational tests ask learners to identify the main theme in a sentence, recognize whether a statement is true or false, arrange a chain of events correctly, and pinpoint the purpose of a particular structure in the text. For instance, the student could be asked to rewrite the following sentence without changing the meaning: “John slipped on the stairs, fell, and hurt his ankle.”

- “John hurt his ankle and then slipped on the stairs.”
- “John fell, hurt his ankle, and slipped on the stairs.”
- “John hurt his ankle because he slipped on the stairs.”
- “John fell because he hurt his ankle after slipping on the stairs.”

Questions on fundamental skills focus on phonics, word recognition, and relationship. These questions require learners to classify syllables or identify antonyms and synonyms of words. For example, a student might be asked to identify another word that means the same as “punish” in the following sentence: “Though the pupil entered class late, the teacher did not punish him because he explained that he passed by the hospital.”

- penalize
- abandon
- welcome

- beat

The literature questions assess students' literacy essentials within a passage by asking them to identify the character, purpose of the author, and significant elements in the plot. The following question is an example:

“I don't want to go to bed now!” howled Faith. “I want to stay up and play with my new game!” Faith was the youngest child in the family. Her parents had already allowed her to stay up later than her siblings did when they were her age. “If you don't go to bed, the tooth fairy will not come to visit you,” said Faith's father. When Faith heard that, she ran straight to bed. She checked under her pillow to make sure her tooth was still there.

Which of the following statements can be used to appropriately describe Faith?

- She was afraid of the tooth fairy.
- She was afraid the tooth fairy would not come.
- It was past her bedtime.
- She was tired.

Validity

Validity is defined as the extent to which an instrument of measurement measures what it is projected to measure (Bickman et al., 2009). The MAP scores for the pretest and posttest were calculated separately for each outcome and grade to ensure the instrument's internal validity (Cordray et al., 2013). In addition, the 2015 pretest scores and those of 2017 posttest were compared for students in the same grade to ensure validity. The use of archival data also confirmed that the scores did not pose any threat to

the validity of the research instrument because only available data were used and there were no incomplete data.

Reliability

Reliability is an evaluation of the extent to which a research tool produces constant outcomes or records after multiple tests. Reliability entails how consistently identical measures generate identical outcomes (Cohen et al., 2017). A reliable pointer generates information that does not fluctuate due to the measurement design, indicator, or instrument characteristics. As confirmed by Cordray et al. (2013), numerous features are contained in the MAP evaluations to ensure they give an unbiased measure of learners' ability. Cordray et al. (2013) established that the instrument's reliability indicators include untimed tests, nonaccess of test items by teachers, and non-re-administration of individual questions to learners for 2 years consecutively. These configurations eliminate the possibility of a learner or a teacher gaining the advantage of being familiar with or predicting the tests in the future. Additionally, the procedures incorporated by NWEA ensure the instrument's high reliability for testing achievement by aligning MAP test questions with content standards (Ball & O'Connor, 2016; NWEA, 2011). Further, reliability was ensured because NWEA turned off the scoring function on the MAP test to control learners and teachers to prevent them from viewing their MAP scores and avoid the generation of MAP reports by control teachers (Klingbeil et al., 2017).

Data Analysis Plan

The data were appropriately cleaned by confirming accuracy, clarity, and consistency. The data were then coded to ease entry into the analysis software. When

analyzing the reading scores of the third grade DLL students to achieve the study objective of determining the score changes associated with SFRS program intervention, I conducted a repeated measures *t* test using IBM Statistical Package for Social Sciences Version 28.0. Prior to repeated measures *t* test analysis, I subjected the data to statistical tests to ensure that the analysis assumptions were not violated. Assumptions regarding the measurement scale, random sampling, normality, sample size adequacy, and homogeneity were checked.

The repeated measures *t* test assumes that the scale of measurement used follows an ordinal scale. This assumption was met because the dependent variable was continuous. The assumption that data are randomly selected from a representative part of the population was not applicable because all population data were used. The Shapiro-Wilk test is used to determine normality because of its excellent capacity to identify deviations from a normal distribution (Wijethilake, 2018). Shapiro-Wilk statistics have a range of 0 to 1. If the probability of error *p* value is higher than 0.05 (i.e., the test is not significant), then the data are normally distributed; thus, the assumption is met. The Kaiser-Meyer-Olkin (KMO) metric assesses sample adequacy. The test is a statistical measure of how small partial correlations are compared to zero-order correlations. The KMO scale ranges from 0 to 1. A threshold of 0.5 indicates adequate data and representativeness of the study population (Bickman et al., 2009); the presence of equal variance across observations was tested using Levene's test. A *p* value greater than 0.05 indicates the presence of homogeneity (Zhang & Wang, 2021), implying that the data are suitable for repeated measures *t* test analysis.

Threats to Validity

Proper checks must be performed to ensure that conclusions generated from research findings are accurate or truthful. As Brincks et al. (2018) noted, validity threats are why the conclusions generated from study findings may be inaccurate due to covariance, causal linkages, or causation notions. External, internal, statistical conclusion, and construct validity are the four forms of threats to validity. The usefulness of research designs can help with research, practice, and policy based on how well the design is internally and externally valid.

Peltier (2021) also noted that external validity refers to the generalizability of results through replication within a single case design. In research, as noted by Brincks et al. (2018), the most severe challenge to external validity is the issue of selection bias. This will allow insight into how these findings relate to other implementation trials not included in this synthesis. Multiple treatment interferences could risk the external validity of the studies, but this is not certain. In the study, some students might have had more than one intervention, so it might be hard to conclude the effectiveness of one intervention for improving student reading achievement scores. However, the findings from this study might be generalizable to similar populations using multiple interventions to improve reading comprehension scores.

De Young and Bottera (2018) identified internal validity as the amount to which an intervention, rather than external variables, is responsible for the change. One threat to internal validity is subject characteristics. When the selection of subjects leads to differences across groups connected to the many variables being examined, as noted by

Baldwin (2018), the subject characteristics threat arises. Selection bias occurs when something in the individual's composition favors one group. For example, if one student is more capable, the study could be skewed toward the higher students. Baldwin (2018) also identified mortality as another internal validity threat. This loss of participants could restrict generalizability and inject bias into studies.

To evaluate the efficacy of the SFRS program in this research study, I identified many challenges to internal validity and resolved them to validate the study's findings. I discovered internal validity risks while researching this study, including history and maturation. Historical circumstances may threaten this study, as students may have been exposed to things other than the intervention that may have contributed to the changes in reading levels rather than the intervention itself, as time proceeds from the pretest to the posttest. To limit the risk of internal validity posed by history, both groups should participate in the same activities, except for the treatments, during the research period.

Since the students who will engage in the research study will undergo continuing developmental processes at varying speeds during the study, maturation may pose a danger to internal validity. In this study, it is anticipated that the maturation will be comparable across the student participants because they share many of the same traits. The selection of pupils based on comparable qualities may help reduce the danger to maturation internal validity.

As Garcia-Perez (2012) identified, statistical conclusion validity occurs when a research study's conclusions are based on thorough data analysis. This generally means that appropriate statistical methods are used with proper small-sample behavior as being

logically capable of answering the research question. Additionally, Petursdottir and Carr (2018) stated that the validity of the statistical result might not be jeopardized by the research study since the statistical tests used are strict enough to offer the requisite statistical capacity to support the conclusion. The MAP evaluation, the dependent variable, was reliable. Between .70 and .85 for children in Grades 2–10, says the MAP technical manual .70 and .86 for students in Grades 2–10 (NWEA, 2011).

Ethical Procedures

As Garcia-Perez (2012) stated, research offers clear and organized information to ethical committees about their study's objective, methodology, and predicted dangers. Additionally, Olaniran and Baruwa (2020) posit that ethics has grown so important in academic endeavors that it prioritizes research and publishing while requiring researchers to protect the dignity of their subjects. Identities of student participants will be kept private and coded before being supplied to me for research. The administrator who provides the data will issue each student a four-digit ID number.

The administrator will provide, in soft, the DLL students' second grade reading scores before the 2015 deployment of the SFRS program and their 2017 MAP performance results following the 2015 implementation of the SFRS program. This data will be encrypted with a security password. Once I receive the data and the decrypting password, I will maintain it secret by storing it in a password-protected computer that is only accessible to me. The research study data will be stored securely and deleted after five years. Although the research will be conducted at the researcher's school, the pupils to be studied will be from different grade levels to alleviate ethical problems. Walden

University's Institutional Evaluation Board (IRB) will conduct a thorough review to ensure that all participants' human rights are upheld.

Summary

A comparing pre- and post-scores before and after intervention causal comparative research design will be conducted to assess if there was a difference in students' reading scores between the pretest and posttest following the implementation of the SFR program. The study design will help me determine if there were any variations in the reading levels of the group receiving intervention on the pretest and posttest. All third grade DLL struggling readers from Commonwealth schools in Riyadh, Saudi Arabia, will participate in the study. The study will use secondary data from historical documents from the 2015 and 2017 academic years. The next chapter will discuss data gathering and the findings of the study.

Chapter 4: Results

The purpose of this quantitative study was to compare the MAP RF scores of third grade DLL struggling readers before and after participation in the SFRS program at Commonwealth School for one school year. Studies of intervention components in literacy programs showed that boosting DLLs' L1 abilities may lead to their effective acquisition of a second language. Chapter 4 is divided into two sections representing the two research questions addressed in this study. The first section presents results regarding the difference in MAP RF score changes after third grade DLL female students participated in the SFRS program for one school year. The second section presents results regarding the difference in MAP RF score changes after third grade DLL male students participated in the SFRS program for one school year.

Data Collection

Archived data from third grade DLL struggling readers who completed one full year of the SFRS program in the school years 2015, 2016, and 2017 were used. The archived data included information on male and female second grade DLL struggling readers who progressed to third grade after completing 1 year of the SFRS program. There was a discrepancy in the quantity of data available because 210 students were designated to receive the built-in intervention plan in the SFRS program. I used all available data for the 142 second grade DLL struggling readers who received 1 year of instruction in the SFRS program. Out of the 142 learners, 73 were female and 69 were male.

Results

Analysis Results for Research Question 1

The first research question addressed the difference in MAP RF score changes after third grade DLL struggling reader female students participated in the SFRS program for 1 year. First, descriptive analysis of reading scores from examinations administered as pretest and MAP RF scores are presented. Statistical test results and the results from the repeated measures *t* test follow the descriptive analysis.

Descriptive Analysis

Data for the pretest scores for female students are presented in Table 3. The pretest reading scores were categorized into grades A (90%–100%), B (80%–89%), C (70%–79%), and D (60%–69%). The descriptive analysis revealed that most learners scored between 70% and 79%, with a mean of 75.55%. The mean, minimum, and maximum scores for each grade are also displayed.

Table 3

Descriptive Characteristics of Pretest Reading Scores for Third Grade DLL Struggling Reader Female Students

Score category	D	C	B	A
Frequency	16	36	17	4
Mean	65.69	73.53	85	93
Min	61	70	81	91
Max	69	77	89	98

Note. $n_1 = 73$, $M = 75.55$ $SD = 8.272$.

Table 4 summarizes these data into the categories pass (traditionally passing with a Grade A, B, or C) and fail (failing with a Grade D). From the total of 73 female students, 16 (22%) received an A, B, or C while 57 (78%) received a D.

Table 4

Descriptive Characteristics of the Reading Scores and MAP RF Scores of Third Grade DLL Struggling Reader Female Students Before and After Participation in the SFRS Program

Semester/pretest	Total <i>M</i>	Pass (A) <i>n</i> = 57 <i>M</i>	Fail (B) <i>n</i> = 16 <i>M</i>	Difference (A-B) <i>M</i> diff	Significance
Fall	175.81 (1.99)	175.14 (2.33)	178.19 (3.73)	-3.07*	0.064
Winter	185.42 (1.77)	184.54 (2.02)	188.56 (3.68)	-4.02**	0.034
Spring	191.52 (1.607)	191.40 (1.81)	191.94 (3.60)	-0.53**	0.026
Overall	184.25 (1.703)	183.70 (1.97)	186.23 (3.39)	-2.53**	0.041
Reading Score	75.55 (0.97)	78.32 (0.94)	65.69 (0.76)	12.63***	0.000

Note. In parenthesis are standard error values; * $p < .05$, ** $p < .01$, *** $p < .001$.

The results in Table 4 show that most third grade DLL struggling reader female students who participated in the SFRS program had high MAP RF scores during the spring semester (pooled Spring RIT $M = 191.52$, $SD = 13.73$). The learners who failed in the reading pretest had a significantly ($p < 0.1$) higher MAP RF score in fall ($p = 0.064$), winter ($p = 0.034$), and spring ($p = 0.026$) semesters than those who passed in the pretest. The failed students had a significantly ($p = 0.000$) higher MAP RF score than the students who passed after 1 year in the SFRS program.

Statistical Test Results

It was necessary to test whether the statistical assumptions were met before performing repeated measures t test analysis. I checked the assumptions about the measuring scale and sample size sufficiency and random sampling, data normality, and variance homogeneity. The repeated measures t test is based on the assumption that the

measurement scale is ordinal. Because the dependent variable was continuous, this assumption was met. There was no sampling because the entire population was used.

Table 5

Statistical Test Results

Variable	Shapiro-Wilk test	Kaiser-Meyer-Olkin (KMO)	Levene's test
Reading score	0.952 (0.083)	0.742	17.119 (0.421)
Fall	0.977 (0.208)	0.957	0.953 (0.332)
Winter	0.971 (0.087)	0.948	0.476 (0.492)
Spring	0.934 (0.141)	0.949	0.001 (0.980)
Overall	0.977 (0.210)	0.999	0.579 (0.449)

Note. In parenthesis are error probability values.

The Shapiro-Wilk test was used to test for normality of the data because of its superior ability to detect deviations from a normal distribution (see Wijethilake, 2018). As shown in Table 5, the Shapiro-Wilk test results were not significant (p values greater than 0.05), indicating that the data were normally distributed; thus, the assumption was met. The KMO measure was used to determine sample adequacy. The test compares partial correlations to zero-order correlations and is a statistical metric (Zhang & Wang, 2021).

Additionally, Table 5 shows that the KMO values for each variable were above 0.9 except for pretest data, which had a KMO value of 0.742 (i.e., KMO values greater than 0.5), indicating that all variables were highly adequate for repeated measures t test analysis. The presence of equal variance across observations was tested using Levene's test. According to Zang and Wang (2021), a p value greater than 0.05 indicates

homogeneity of variance. Table 5 shows that Levene's test statistics were greater than 0.05, implying that the data were suitable for repeated measures *t* test analysis.

MAP RF Score Changes as a Result of SFRS Program

To answer the research questions, I compared the second grade reading scores (pretest) to the MAP-RF standardized 2017 third grade test scores (posttest) of 73 DLL struggling reader female students following the SFRS program implementation for 1 year. The score changes were tested for significance with a repeated measures *t* test. Table 6 shows that second grade reading scores and third grade MAP RF scores for all the semesters were negligibly and negatively correlated ($r < 0.52, p < 0.05$). The repeated measures *t* test analysis revealed that the initial scores at the beginning of the year were $M = 100.260$ with $SD = 19.677$. The final scores for DLL struggling reader female students at the end of the year were $M = 115.973$ and $SD = 16.717$. The MAP RF score difference (15.73 points) was significant with $t = 59.275, df = 72, \text{ and } p = 0.000$.

Table 6

Repeated Measures t test for MAP RF Score Changes for the Third Grade DLL Struggling Reader Female Students After 1 Year Participation in the SFRS Program

Pair	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Pair 1 Fall-pretest	-0.109 (0.357)	100.260	19.677	43.534	72	0.000
Pair 2 Winter-pretest	-0.066 (0.582)	109.877	17.724	52.967	72	0.000
Pair 3 Spring-pretest	-0.099 (0.406)	115.973	16.717	59.275	72	0.000
Pair 4 Overall-pretest	-0.096 (0.418)	108.703	17.418	53.323	72	0.000

Note. In parenthesis are probability values for correlations; *r* is correlation coefficient,

and *p* is significance level at $p < 0.001$.

Analysis Results for Research Question 2

The second research question addressed how MAP RF scores changed after third grade DLL struggling reader male students participated in the SFRS program for 1 year. Descriptive analysis of reading scores from the pretest and MAP RF scores is presented. This is followed by statistical test results after which the findings for the repeated measures *t* test analysis are presented.

Descriptive Analysis

Descriptive statistics for DLL struggling reader male students on the pretest scores are presented in Table 7. The reading results on the pretest were categorized based on the four grades: A, B, C, and D. The descriptive analysis indicated that most of the students attained grade C (70%–79%), with a mean of 77.61%. Table 7 also shows the mean, minimum, and maximum score for each grade of the pretest reading score.

Table 7

Descriptive Analysis of Pretest Reading Scores for Third Grade DLL Struggling Reader Male Students

Score category	D	C	B	A
Frequency	7	37	23	2
Mean	65.57	74.68	84.82	91
Min	61	70	80	90
Max	68	79	89	92

Note. $n_2 = 69$, $M = 77.61$, $SD = 7.103$.

Table 8 shows a summary of third grade DLL struggling reader male students by status pass versus fail. Passed students are those with grades C and above in the pretest reading examination, while fail students are those who scored Grade D on the test. Sixty-two students (89.86%) were in the pass category, while seven (10.14%) were in the fail category. The results revealed that most third grade DLL struggling reader male students who participated in the SFRS program for 1 year had low and high MAP RF scores during the fall and spring semesters, respectively (pooled fall RIT $M = 172.78$, $SD = 7.103$, pooled spring RIT $M = 188.58$, $SD = 14.011$). The learners who passed in the reading pretest had a significantly ($p < 0.05$) higher MAP RF score in fall (173.13), winter (184.32), and spring (189.23) semesters than those who failed in the pretest. The passed students had a significantly ($p = 0.021$) higher MAP RF score (182.23) than the failed students (176.57) after 1 year in the SFRS program.

Table 8

Descriptive Characteristics of the Reading Scores and MAP RF Scores of Third Grade DLL Struggling Reader Male Students After Participation in the SFRS Program

Semester/pretest	Total <i>M</i>	Pass (A) <i>n</i> = 62 <i>M</i>	Fail (B) <i>n</i> = 7 <i>M</i>	Difference (A–B) <i>M</i> diff	Significance
Fall	172.78 (1.88)	173.13 (1.10)	169.71 (6.02)	3.42**	0.050
Winter	183.59 (1.88)	184.32 (2.03)	177.14 (4.82)	7.18**	0.021
Spring	183.59 (1.90)	189.23 (1.80)	182.86 (4.60)	6.37**	0.029
Overall	181.65 (1.71)	182.23 (1.83)	176.57 (4.82)	5.65**	0.035
Reading score	77.61 (0.86)	78.97 (0.77)	65.57 (1.04)	13.40***	0.000

Note. In parenthesis are standard error values; * $p < .05$, ** $p < .01$, *** $p < .001$

Statistical Test Results

Data were statistically tested before repeated measures *t* test analysis to verify that the model assumptions were met. I checked the assumptions on the measurement scale, random sampling, normality, sample size adequacy, and homogeneity. An ordinal scale is assumed when using the repeated measures *t* test. Because the dependent variable was continuous, this assumption was met. Additionally, the assumption that data for analysis are drawn from a random population sample was irrelevant because all of the population's data were used in the study.

The Shapiro-Wilk test was used to establish normality in this research because of its ability to discover deviations from a normal distribution (see Wijethilake, 2018). The

Shapiro-Wilk test p values varied from 0.175 to 0.808 (i.e., $p > 0.05$), suggesting that the data were normally distributed and, therefore, the assumption was satisfied (see Table 9).

Table 9

Statistical Test Results

Variable	Shapiro-Wilk test	Kaiser-Meyer-Olkin (KMO)	Levene's test
Reading score	0.975 (0.175)	0.730	5.644 (0.201)
Fall	0.988 (0.727)	0.942	0.057 (0.813)
Winter	0.981 (0.374)	0.954	0.461 (0.500)
Spring	0.989 (0.808)	0.923	0.216 (0.644)
Overall	0.985 (0.588)	1.000	0.020 (0.888)

Note. In parenthesis are probability values.

The KMO measure was used to determine sample adequacy in the sampling adequacy test. Table 9 reveals that each variable's KMO value was larger than 0.7, suggesting that all variables were sufficient for repeated measures t test analysis. Levene's test was used to determine whether there was equal variance among data. A p value larger than 0.05, according to Zang and Wang (2021), shows variance homogeneity. The Levene's test statistics were greater than 0.05 (see Table 9), indicating that the data were suitable for repeated measures t test analysis.

MAP RF Score Differences Associated With SFRS Program

The repeated measures t test was used to compare DLL struggling reader second grade reading scores (pretest) to MAP RF 2017 test scores (posttest) following the SFRS program deployment of 69 DLL struggling reader male pupils. The correlations between second grade reading scores and third grade MAP RF scores for all semesters were positive but negligible ($r < 0.2$, $p > 0.05$). The analysis revealed statistically significant

changes in MAP RF scores between the initial and the final examinations, as presented in Table 10. The initial scores at the beginning of the year were $M = 95.174$, $SD = 17.122$. The final scores at the end of the year were $M = 110.971$, $SD = 14.867$. The MAP RF score changes were statistically significant, $t = 62.003$, $df = 68$, and $p = 0.000$.

Table 10

Repeated Measures t test for MAP RF Score Changes the Third Grade DLL Struggling Reader Male Students After Participation in the SFRS Program

Pair	<i>r</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Pair 1 Fall-pretest	0.009 (0.939)	95.174	17.122	46.174	68	0.000
Pair 2 Winter-pretest	0.054 (0.660)	105.986	16.915	52.046	68	0.000
Pair 3 Spring-pretest	0.129 (0.290)	110.971	14.867	62.003	68	0.000
Pair 4 Overall-pretest	0.066 (0.592)	104.043	15.487	55.805	68	0.000

Note. In parenthesis are probability values for correlations; *r* is correlation coefficient,

and *p* is significance level at $p < 0.001$.

Chapter 5: Discussion, Conclusions, and Recommendations

The purpose of this study was to compare the MAP RF scores of third grade DLL struggling readers before and after participating in the SFRS program at Riyadh, Saudi Arabia, Commonwealth School for one school year. Several significant findings were revealed from an assessment of pretest and posttest data and progress tracking probe data. First, the intervention helped reading students, including both male and female students, support increased use of the SFRS program. Second, when learners were separated by reading score performance, those who passed and those who failed the pretest had higher MAP RF scores after 1 year of the SFRS program. Furthermore, low-performing failing students had considerably better MAP RF scores than low-performing students.

Results also revealed that male readers who passed the pretest improved their MAP RF scores compared to their counterparts who failed. The goal of this chapter is to analyze and understand the findings of this investigation. A framework of research questions is used to organize the findings. The chapter ends with a summary of the study's limitations and areas for future research.

Discussion

Research Question 1

The first of the two research questions addressed whether there was a difference in MAP RF scores for third grade DLL struggling reader female students before and after participating in the SFRS program for one school year. Descriptive analysis of pretest and MAP RF scores showed that DLL struggling reader female students demonstrated growth during the intervention on reading fluency, as indicated by a statistically significant

difference between low and high pretest reading performance (see Table 4). This finding is consistent with what other researchers found regarding intervention components in literacy programs: that context, child feedback, and educational intervention aspects contribute to the explanation of academic achievements (see Barber et al., 2020; Donnelly et al., 2020; Grover et al., 2020). In the current study, female students who failed in the pretest recorded significantly higher MAP RF scores after the SFRS program than those who passed in the pretest. This finding could be attributed to how teachers cater better to low-performing students and provide them with more structured learning activities (O'Neal & Zimmerman, 2019). This means teachers could spend more time on the students who need it most, leading to better academic performance for low-performing students. The same cannot be said for high-performing students, who had much more to gain from the intervention. As shown in Table 4, the DLL struggling readers performed progressively well as the school year continued, as demonstrated by positive MAP RF score changes when comparing the first and last semesters of the year. The repeated measures *t* test results further revealed DLL struggling readers achievement after 1 year of the SFRS intervention program. As shown in Table 6, the MAP RF score difference was statistically significant, indicating that third grade learners improved their reading fluency after being subjected to the SFRS program for 1 year. The findings are consistent with those of Wang and Lin (2019), who found that the SFRS program improved students' reading and writing abilities and enthusiasm to learn, resulting in improved performance. The current results also support Leal (2016), who found that

reading intervention programs in Taiwanese schools encouraged students' critical thinking, meaningful dialogues, and creative writing.

Research Question 2

The second question addressed whether DLL struggling reader male students' MAP RF scores differed between second and third grade after the SFRS program intervention for one year. At the end of Grade 2, the students were divided into passed and failed readers based on their reading fluency performance. As shown in Table 8, learners identified as passed readers had significantly higher MAP RF scores at the end of the third grade examination. This finding is in line with a growing body of evidence demonstrating the predictive value of educational intervention programs in raising students' academic achievement (Lathouras et al., 2019). The high MAP RF scores achieved by DLL struggling reader male students with high reading scores imply that these learners required less effort to enhance their fluency. Although students who failed the pretest still needed to focus on areas they were underperforming in, those who passed focused on areas that required more cognitive capacity. The observed results support the assertion by Wang and Lin (2019) that the SFRS program is important in predicting students' reading fluency achievement. The current findings also suggest that learners with weak reading fluency who have received educational assistance should be closely monitored. As Lathouras et al. (2019) suggested, individual students with lower MAP RF scores may be at risk of long-term reading difficulties after the intervention. Identifying these learners would be necessary because they would require additional intervention to enhance their reading fluency.

The repeated measures *t* test results revealed that the MAP RF scores for the first and final assessments changed substantially between second and third grade struggling reader DLLs after 1 year of SFRS training. The results showed that the scores at the end of DLL struggling readers' third grade were higher than those at the end of DLLs' second grade, indicating a statistically significant improvement in DLL struggling readers' reading fluency. These findings support the results reported by Erickson and Whorton-McDonald (2019), which identified that literacy achievement of students is supported by differentiating instruction, offering adequate scaffolds, and providing ample opportunities for cooperation. The SFSR program provides activities that support students' need for competence by including various frameworks (e.g., visuals, partner work, teacher modeling) and preferences for an extra challenge (e.g., independent work, open sort) that encourage DLLs to feel belonging and collaborate toward a common goal of increasing their reading skills. According to the findings, access to high-quality dual language teaching may be part of why bi/multilingual students outperform monolingual peers in reading growth rates, as Bialystok (2018) supported. Previous research had also shown positive correlations between dual language program participation and academic achievement (Howard et al., 2018).

General Discussion

Multilingual environments make the process of literacy acquisition more difficult. In accordance with the linguistic interdependence theory (Cummins, 1979), students learning to read in L2 may benefit from their L1 capabilities, such as content-area knowledge, literacy, and problem solving. I tested this hypothesis by examining the score

changes related to SFRS program interventions on the MAP RF scores of third grade DLL struggling readers at Commonwealth School in Riyadh, Saudi Arabia.

This study's overall outcomes showed a link between second grade reading scores and third grade DLL struggling readers' MAP RF scores after students participated in the SFRS program at Commonwealth School for one school year. These results were obtained after the students had been enrolled in the program for an entire academic year. Tables 6 and 10 show that the associations were present during the study. The reading fluency and literacy abilities in English shown by students on the pretest were good predictors of their results on the MAP RF at the end of the third grade. These findings expand the scope of earlier research by looking beyond the unidirectional L1-to-L2 links. Instead, the findings demonstrate that literacy abilities in several languages have relations with one another that go in both directions. The importance of bidirectionality in the linguistic interdependence hypothesis had been established. However, most of the research had concentrated on L1-to-L2 transfer.

Learning to read and write in more than one language requires careful attention to the content and the educational situation. It is common for literacy instruction to be provided only in the L2 in many multilingual contexts. Even if literacy instruction is provided in the L1, L2 skills are given more weight because of their perceived social and economic importance (Moon & Zimmerman, 2017; O'Neal & Zimmerman, 2019; Zimmerman et al., 2019). Regarding social mobility in Saudi Arabia, it is more important to speak Arabic than English, even though both are widely spoken. Because of this, in many schools Arabic is the primary language of teaching despite the English language

policy encouraging the use of L2 or English (Kim & Piper, 2019). Although the SFRS program, which is an interactive English language arts program, emphasizes a student-centered approach to reading, writing, listening, speaking, and thinking from grades kindergarten through fifth, this trend can be seen in the lower mean scores in reading status at the beginning of first grade (see Table 1).

The current study findings have several implications for reading literacy in foreign language acquisition and foreign language assessment. Saudi Arabia's Economic Vision 2030 established a baseline for L2 English reading proficiency. Time needed to finish tasks, whether in an assessment circumstance or the classroom, may now be estimated by educators and test creators. The second benefit of this research is that the positive correlation between L2 reading fluency and MAP RF scores after the intervention may be used to influence explicit fluency instruction in the L2 classroom. Fluency training may benefit a student's ability to do well on tests, and it can also increase their general grasp of the language, particularly in the early stages of learning. Finally, given the high correlation between L2 reading literacy and MAP RF scores after 1 year of the SFRS program, foreign language instructors should be aware of a student's L2 reading profile. Successful L2 acquisition can be facilitated by helping those with difficulty reading in their native language.

The favorable association between reading literacy and MAP RF scores among DLL struggling readers after an intervention adds to the body of data indicating the value of explicit and systematic education in literacy acquisition. Early reading skills in languages with alphabetic writing systems have long been linked to improving

phonological and letter knowledge (Kim et al., 2018). Furthermore, the current findings have practical implications for the language of instruction, a complex and critical policy issue in multilingual contexts such as Saudi Arabia.

In conjunction with previous research, the current results suggest that educational interventions, including explicit and systematic training in basic literacy skills, will increase students' literacy and reading skills. Cross-language transfer of reading literacy, phonological awareness, and letter knowledge in multilingual children has been shown by the present results. They also have significant pedagogical implications for multilingual literacy teaching. A pre-post design, in which a single assessment before and after the intervention is used to establish the average level of gain in reading abilities, is the most common way to investigate intervention effectiveness. This technique has proved useful in determining whether a given intervention program enhances reading ability, estimating effect sizes for intervention-driven reading skill development, and evaluating the effectiveness of intervention methods.

The dose-response connection between the quantity of intervention and progress in reading abilities is one issue this approach has left unresolved. Researchers may construct models that address critical issues regarding intervention dosage by performing experimental sessions throughout the intervention and asking the following questions: What is the appropriate level of intervention for a child? What is the limit of how little is too little? What is the best cost-effective method for gaining the most benefits? Is it worthwhile to invest in this intervention, and what should the suggested dosage be for each child? The current study results suggest that 45 minutes of training per day for 9 to

10 weeks will result in considerable progress in reading literacy based on the favorable DLL struggling readers' MAP RF score changes after the intervention. However, I only looked at the MAP RF score for one intervention program with a small, heterogeneous sample of learners and no active control condition. Therefore, it will be critical to expand this research to more diverse groups and additional intervention programs using a more controlled research design. Conducting follow-up studies, researchers can compare intervention dose-response curves of component reading abilities and connect those development curves with particular intervention strategies.

Conclusion

This research adds to the growing body of data supporting the need to deliver universal, high-quality, evidence-based early intervention to learners at risk of long-term reading challenges. The findings of this study suggested a link between the reading intervention and changes in MAP RF scores 1 year after the intervention. The results confirm worldwide research that favorable benefits are displayed immediately after the intervention but that more intervention is required for children to preserve these advantages. Therefore, the need to design, deliver, and monitor evidence-based treatments for students who do not respond to classroom-based interventions, particularly those at risk of future oral language and literacy challenges, is highlighted in this research.

Limitations and Recommendations for Future Research

Despite the relevance of the conclusions of this research, it had significant limitations. First, due to demographic considerations, this study's generalizability was

restricted. Although this research had a sufficient sample size ($N = 142$), I evaluated only one elementary school in the Saudi Arabian district of Riyadh.

There were no linguistic assessments before Grade 3 in the research. If K–2 measurements had been included for this grade range, I would have been able to analyze the relationships between reading abilities from kindergarten to third grade in the growth models. In an ideal scenario, a study of L2 reading literacy would have been in place from the moment a child entered the SFSR program until they exited, using the same metric to assess growth. I was constrained to use the available data because the research depended on the existing data set of third grade DLL struggling readers. On the other hand, reliability estimates and predictive validity of reading literacy assessments for students in the first 2 years of school are often below expectations (Mitchell & Alfuraih, 2017).

Additionally, the sample size of the study was a limitation. Single-case designs have been criticized because their conclusions may never apply to other groups (Soland & Sandilos, 2021). Direct and systematic replication is required to improve the external validity. Second and third grade DLL struggling readers' comparisons may have shown different findings from research aiming to analyze the growth of the same children across grade levels.

Further, I did not examine elements such as fluency or working memory, which significantly influence students' ability to learn to read (see Barry, 2019; Geel, 2016; Snow & Matthews, 2016). Although learning settings, behavior, and attendance have been extensively studied and have been shown to influence academic performance, these

factors were not considered in the present study (see Acosta et al., 2019; Bauer et al., 2017; Edyburn et al., 2017; Paige et al., 2019; Piasta, 2016; Williams & Lowrance-Faulhaber, 2018).

This research focuses on the reading performance scores of third grade pupils in a particular school district, one specific elementary school in Saudi Arabia's District of Riyadh. Larger-scale research duplicating this intervention will require careful evaluation of cross-sector academic institutions, differentiated degrees of socioeconomic disadvantage, and instructor factors (e.g., engagement, participation, years of experience). Future research should examine protective characteristics for students with identified reading difficulties, such as language background and disability-related funding. Finally, these findings should be interpreted in light of parental support and the home learning environment, which substantially influences student achievement (Williams et al., 2017; Piasta, 2016).

Another promising area for future study is longitudinal studies examining how individual children improve their reading fluency over time. No conclusions can be drawn regarding L2 reading progress since the present research includes pupils of various abilities. A longitudinal study might give reliable evidence for determining the best moment to start fluency instruction in the L2 classroom. Educators might use this information to create programs to help pupils improve their L2 reading skills. This can have a long-term influence on student's motivation and pleasure to read in a foreign language and their actual reading skills.

A primary suggestion of this study is to conduct future research to investigate the relative contributions of teaching language and teaching strategy to DLLs engaged in early education intervention programs and to develop better techniques for conducting research with DLLs. One concept is that the field adopts a unified taxonomy to describe and classify the many educational interventions investigated by researchers in DLLs.

References

- Acosta, J., Williams, J., III, & Hunt, B. (2019). Dual language program models and English language learners: An analysis of the literacy results from a 50/50 and a 90/10 model in two California schools. *Journal of Educational Issues*, 5(2), 1–12. <https://doi.org/10.5296/jei.v5i2.14747>
- Alanís, I. (2018). Enhancing collaborative learning: Activities and structures in a dual language preschool classroom. *AMAE Journal*, 12(1), 5–26. <https://doi.org/10.24974/amae.12.1.375>
- Allmnakrah, A., & Evers, C. (2020). The need for a fundamental shift in the Saudi education system: Implementing the Saudi Arabian economic vision 2030. *Research in Education*, 106(1), 22–40. <https://doi.org/10.1177/0034523719851534>
- Almoayidi, K. A. (2018). The effectiveness of using L1 in second language classrooms: A controversial issue. *Theory & Practice in Language Studies*, 8(4), 375–379. <https://doi.org/10.17507/tpls.0804.02>
- Al-Zahrani, N. O. A., & Rajab, H. (2017). Attitudes and perceptions of Saudi EFL teachers in implementing Kingdom of Saudi Arabia's vision 2030. *International Journal of English Language Education*, 5(1), 83-99. <http://doi.org/10.5296/ijele.v5i1.10733>
- Arif, M., & Abdullah, I. H. (2017). The impact of L1 metaphorical comprehension on L2 metaphorical comprehension of Iraqi EFL learners. *Advances in Language and Literary Studies*, 8(4), 8–13. <https://doi.org/10.7575/aiac.all.v.8n.4p.8>

- Baldwin, L. (2018). Internal and external validity and threats to validity. In L. Baldwin (Ed.), *Research concepts for the practitioner of educational leadership* (pp. 31–36). Brill. https://doi.org/10.1163/9789004365155_007
- Ball, C. R., & O'Connor, E. (2016). Predictive utility and classification accuracy of oral reading fluency and the measures of academic progress for the Wisconsin knowledge and concepts exam. *Assessment for Effective Intervention, 41*, 195–208. <https://doi.org/10.1177/1534508415620107>
- Barber, A. T., Cartwright, K. B., Stapleton, L. M., Lutz Klauda, S., Archer, C. J., & Smith, P. (2020). Direct and indirect effects of executive functions, reading engagement, and higher order strategic processes in the reading comprehension of Dual Language Learners and English Monolinguals. *Contemporary Educational Psychology, 61*, Article 101848. <https://doi.org/10.1016/j.cedpsych.2020.101848>
- Barry, A. (2019). Gender differences in academic achievement in Saudi Arabia: A wake-up call to educational leaders. *International Journal of Education Policy and Leadership, 15*(15). <https://doi.org/10.22230/ijep.2019v15n15a890>
- Bauer, E. B., Presiado, V., & Colomer, S. (2017). Writing through partnership: Fostering translanguaging in children who are emergent bilinguals. *Journal of Literacy Research, 49*(1), 10–37. <https://doi.org/10.1177/1086296x16683417>
- Beins, B. C. (2017). *Research method: A tool for life*. Cambridge University Press.
- Berry, T., Byrd, K. K., & Collins, K. (2009). *The effects of Reading Street on reading achievement: A focus on second year curriculum users. Final report*. Claremont Graduate University. <https://www.pearson.com/content/dam/one-dot-com/one->

dot-com/ped-blogs/wp-content/uploads/ReadingStreetReportRevised_Final.pdf

- Bialystok, E. (2018). Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*, 21(6), 666–679. <https://doi.org/10.1080/13670050.2016.1203859>
- Bibler, A. (2021). Dual language education and student achievement. *Education Finance and Policy*, 16(4), 634–658. https://doi.org/10.1162/edfp_a_00320
- Bickman, L., Rog, D. J., & Best, S. J. (2009). *The SAGE handbook of applied social research methods* (2nd edition.). SAGE.
- Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, 37(6), 1461–1476.
<https://doi.org/10.1017/S0142716416000060>
- Brincks, A., Montag, S., Howe, G. W., Huang, S., Siddique, J., Ahn, S., Sandler, I. N., Pantin, H., & Brown, C. H. (2018). Addressing methodologic challenges and minimizing threats to validity in synthesizing findings from individual-level data across longitudinal randomized trials. *Prevention Science: The Official Journal of the Society for Prevention Research*, 19(1), 60–73.
<https://doi.org/10.1007/s11121-017-0769-1>
- Burns, M. K., & Young, H. (2019). Test review: Measures of academic progress skills. *Journal of Psychoeducational Assessment*, 37(5), 665–668.
<https://doi.org/10.1007/s11121-017-0769-1>
- Carroll, J. M., Holliman, A. J., Weir, F., & Baroody, A. E. (2019). Literacy interest,

home literacy environment and emergent literacy skills in preschoolers. *Journal of Research in Reading*, 42(1), 150–161. <https://doi.org/10.1111/1467-9817.12255>

- Cohen, L., Manion, L., & Morrison, K. (2017). Validity and reliability. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research methods in education*, (8th ed., pp. 245–284). Routledge. <https://doi.org/10.4324/9781315456539-14>
- Citron, F. M. M., Michaelis, N., & Goldberg, A. E. (2020). Metaphorical language processing and amygdala activation in L1 and L2. *Neuropsychologia*, 140, Article 107381. <https://doi.org/10.1016/j.neuropsychologia.2020.107381>
- Collier, V. P., & Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale, longitudinal research. *Annual Review of Applied Linguistics*, 37, 203–217. <https://doi.org/10.1017/S0267190517000034>
- Cordray, D., Pion, G., Brandt, C., Molefe, A., Toby, M. (2012). *The impact of the Measures of Academic Progress (MAP) program on student reading achievement* (NCEE Report No. 2013-4000). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Cummins, J. (1978). Educational implications of mother tongue maintenance in minority-language groups. *Canadian Modern Language Review*, 34(3), 395–416. <https://doi.org/10.3138/cmlr.34.3.395>
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.

<https://doi.org/10.3102/00346543049002222>

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14(2), 175–187. <https://doi.org/10.2307/3586312>

Cummins, J. (1981). *The role of primary language development in promoting educational success for language minority students. Schooling and language minority students. A theoretical framework*. Evaluation, Dissemination and Assessment Center California State University, Los Angeles Los Angeles, California.
<https://files.eric.ed.gov/fulltext/ED249773.pdf>

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language Processing in Bilingual Children* (pp. 70-89). Cambridge University Press.
<https://doi.org/10.1017/cbo9780511620652.0006>

Cummins, J. (2000). Immersion education for the millennium: What we have learned from 30 years of research on second language immersion.
<https://carla.umn.edu/cobal/t/modules/strategies/immersion2000.pdf>

Cummins, J. (2016). Reflections on Cummins (1980), “The cross lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue.” *TESOL Quarterly*, 50(4), 940–944. <http://www.jstor.org/stable/44984725>

Cummins, J., Bismilla, V., Chow, P., Giampapa, F., Cohen, S., Leoni, L., Sandhu, P., & Sastri, P. (2005). Affirming identity in multilingual classrooms. *Educational Leadership*, 63(1), 38–43. <https://doi.org/10.1002/tesq.241>

- Cuticelli, M., Collier-Meek, M., & Coyne, M. (2016). Increasing the quality of tier 1 reading instruction: Using performance feedback to increase opportunities to respond during implementation of a core reading program. *Psychology in the Schools, 53*(1), 89–105. <https://doi.org/10.1002/pits.21884>
- Daller, M., & Ongun, Z. (2018). The threshold hypothesis revisited: Bilingual lexical knowledge and non-verbal IQ development. *International Journal of Bilingualism, 22*(6), 675–694. <https://doi.org/10.1177/1367006917690835>
- Denis, D. J. (2016). *Applied univariate, bivariate, and multivariate statistics*. Wiley.
- De Young, K. P., & Bottera, A. R. (2018). A summary of reporting guidelines and evaluation domains for using single-case experimental designs and recommendations for the study of eating disorders. *International Journal of Eating Disorders, 51*(7), 617–628. <https://doi.org/10.1002/eat.22887>
- Dillon, A. M., Hojeij, Z., Perkins, A., & Malkawi, R. (2020). Examining the text quality of English/Arabic dual language children's picture books. *International Journal of Bilingual Education and Bilingualism, 23*(8), 888–901. <https://doi.org/10.1080/13670050.2017.1415867>
- Donnelly, P. M., Huber, E., & Yeatman, J. D. (2020). Intensive summer intervention drives linear growth of reading skill in struggling readers. *Frontiers in Psychology, 10*, Article 1900. <https://doi.org/10.3389/fpsyg.2019.01900>
- Durán, L. K., Hartzheim, D., Lund, E. M., Simonsmeier, V., & Kohlmeier, T. L. (2016). Bilingual and home language interventions with young dual language learners: A research synthesis. *Language, Speech & Hearing Services in Schools, 47*(4), 347–

371. https://doi.org/10.1044/2016_lshss-15-0030

- Durgunoğlu, A. Y., Segers, E., & Broek, P. V. D. (2017). An updated review of cross-language transfer and its educational implications. In E. Segers, & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 167–182). <https://doi.org/10.1075/z.206.11dur>
- Edyburn, K. L., Quirk, M., Felix, E., Swami, S., Goldstein, A., Terzieva, A., & Scheller, J. (2017). Literacy screening among latino/a and dual language learner kindergarteners: predicting first grade reading achievement. *Literacy Research and Instruction, 56*(3), 250–267. <http://doi.org/10.1080/19388071.2017.1305470>
- Erickson, J. D., & Wharton-McDonald, R. (2019). Fostering autonomous motivation and early literacy skills. *The Reading Teacher, 72*(4), 475-483. <https://doi.org/10.1002/trtr.1750>
- Farley-Ripple, E. N., Jennings, A., & Jennings, A. B. (2021). Tools of the trade: a look at educators’ use of assessment systems. *School Effectiveness & School Improvement, 32*(1), 96–117. <https://doi.org/10.1080/09243453.2020.1777171>
- Follmer, D. J., Fang, S.-Y., Clariana, R. B., Meyer, B. J. F., & Li, P. (2018). What predicts adult readers’ understanding of STEM texts? *Reading & Writing, 31*(1), 185–214. <https://doi.org/10.1007/s11145-017-9781-x>
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., Wissel, S., National Center for Education Evaluation and Regional Assistance (ED), What Works Clearinghouse (ED), & Mathematica Policy

- Research, I. (2016). Foundational skills to support reading for understanding in kindergarten through 3rd grade. In Nebraska Department of Education (Ed.), *Educator's practice guide*. NCEE 2016-4008. National Center for Education Evaluation and Regional Assistance.
- Foorman, B. R., Herrera, S., & Dombek, J. (2018). The relative impact of aligning tier 2 intervention materials with classroom core reading materials in grades K–2. *Grantee Submission*, 118(3), 477–504. <https://doi.org/10.1086/696021>
- Galloway, E., Uccelli, P., Aguilar, G., & Barr, C. D. (2020). Exploring the cross-linguistic contribution of Spanish and English academic language skills to English text comprehension for middle-grade dual language learners. *AERA Open*, 6, 1–20. <https://doi.org/10.1177/2332858419892575>
- Gao, T., Zhao, J., Li, X., Mao, Y., Chen, Q., & Harrison, S. E. (2020). Impact of rapid reading skills training on reading rate and reading achievement among primary school students in China. *Educational Psychology*, 40(1), 42–61. <https://doi.org/10.1080/01443410.2019.1607257>
- Garcia, E. (2018). The classroom language context and English and Spanish vocabulary development among dual language learners attending head start. *Early Childhood Research Quarterly*, 42, 148–157. <https://doi.org/10.1016/j.ecresq.2017.09.005>
- García-Pérez M. A. (2012). Statistical conclusion validity: some common threats and simple remedies. *Frontiers in Psychology*, 3., 325. <https://doi.org/10.3389/fpsyg.2012.00325>
- Gatti, G. G. (2005). Scott Foresman Reading Street benchmark item-validation study

2005. <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/ped-blogs/wp-content/pdfs/r3-scott-foresman-reading-street-benchmark-item-validation-study-year-1.pdf>

- Geel, A. (2016). Separate or together? Women-only public spaces and participation of Saudi women in the public domain in Saudi Arabia. *Contemporary Islam*, 10, 357–378. <https://doi.org/10.1007/s11562-015-0350-2>
- Gillanders, C., Franco, X., Seidel, K., Castro, D. C., & Méndez, L. I. (2017). Young dual language learners' emergent writing development. *Early Child Development and Care*, 187(3–4), 371–382. <http://doi.org/10.4324/9781315108278-7>
- Gillanders, C. (2018). ¿Cómo lo escribo en inglés o en español? Writing in dual-language learners. *Reading Teacher*, 71(4), 421–430. <https://doi.org/10.1002/trtr.1635>
- Goodrich, J. M., Thayer, L., & Leiva, S. (2021). Evaluating achievement gaps between monolingual and multilingual students. *Educational Researcher*, 50(7), 429–441. <https://doi.org/10.3102/0013189X21999043>
- Grover, V., Rydland, V., Gustafsson, J.-E., & Snow, C. E. (2020). Shared book reading in preschool supports bilingual children's second-language learning: A cluster-randomized trial. *Child Development*, 91(6), 2192–2210. <https://doi.org/10.1111/cdev.13348>
- Guo, L. (2018). Modeling the relationship of metacognitive knowledge, L1 reading ability, L2 language proficiency and L2 reading. *Reading in a Foreign Language*, 30(2), 209–231. <https://files.eric.ed.gov/fulltext/EJ1194492.pdf>
- Gutiérrez, N., Jiménez, J. E., de León, S. C., & Seoane, R. C. (2020). Assessing

foundational reading skills in kindergarten: A curriculum-based measurement in Spanish. *Journal of Learning Disabilities*, 53(2), 145–159.

<http://doi.org/10.1177/0022219419893649>

Hartanto, A., Yang, H., & Yang, S. (2018). Bilingualism positively predicts mathematical competence: Evidence from two large-scale studies. *Learning and Individual Differences*, 61, 216–227. <https://doi.org/10.1016/j.lindif.2017.12.007>

Heppt, B., & Stanat, P. (2020). Development of academic language comprehension of German monolinguals and dual language learners. *Contemporary Educational Psychology*, 62, Article 101868. <https://doi.org/10.1016/j.cedpsych.2020.101868>

Hillier, A., Ryan, J., Donnelly, S. M., & Buckingham, A. (2019). Peer mentoring to prepare high school students with autism spectrum disorder for college. *Advances in Neurodevelopmental Disorders*, 3(4), 411–422. <https://doi.org/10.1007/s41252-019-00132-y>

Howard, E. R., Lindholm-Leary, K. J., Rogers, D., Olague, N., Medina, J., Kennedy, B., Sugarman, J., & Christian, D. (2018). *Guiding Principles for Dual Language Education* (3rd ed.). Center for Applied Linguistics.

Ijalba, E., Bustos, A., & Romero, S. (2020). Phonological orthographic deficits in developmental dyslexia in three Spanish-English bilingual students. *American Journal of Speech-Language Pathology*, 29(3), 1133–1151.

https://doi.org/10.1044/2020_AJSLP-19-00175

Jager, J., Putnick, D. L., & Bornstein, M. H. (2017). More than just convenient: The scientific merits of homogeneous convenience samples. *Monographs of the*

Society for Research in Child Development, 82(2), 13–30.

<https://doi.org/10.1111/mono.12296>

Ji, M., & Baek, S. (2019). Native Korean-speaking children learning to read in English: A structural analysis of L2-English literacy acquisition. *Journal of Psycholinguistic Research*, 48(2), 391–415. <https://doi.org/10.1007/s10936-018-9610-7>

Johnson, A., & Annenberg Institute for School Reform at Brown University (2020).

Within-year achievement gains for English learners. *EdWorkingPaper No. 20-337*. Annenberg Institute for School Reform at Brown University.

Khaghaninejad, M. S. (2020). Are reading comprehension ability and its strategies transferable from L1 to L2? Evidence from upper-intermediate EFL learners in Iran. *Southern African Linguistics & Applied Language Studies*, 38(4), 293–306.

<https://doi.org/10.2989/16073614.2020.1854796>

Kieffer, M. J., & Thompson, K. D. (2018). Hidden Progress of Multilingual Students on NAEP. *Educational Researcher*, 47(6), 391–398.

<https://doi.org/10.3102/0013189X18777740>

Kim, Y.-S. G., Cho, J.-R., & Park, S.-G. (2018). Unpacking direct and indirect relationships of short-term memory to word reading: Evidence from Korean-speaking children. *Journal of Learning Disabilities*, 51(5), 473–481.

<https://doi.org/10.1177/0022219417724817>

Kim, Y.-S. G., & Piper, B. (2019). Cross-language transfer of reading skills: An empirical investigation of bidirectionality and the influence of instructional environments. *Reading and Writing*, 32, 839–871. <https://doi.org/10.1007/s11145->

[018-9889-7](#)

- Klingbeil, D. A., Nelson, P. M., Van Norman, E. R., & Birr, C. (2017). Diagnostic accuracy of multivariate universal screening procedures for reading in upper elementary grades. *Remedial and Special Education, 38*(5), 308–320.
<https://doi.org/10.1177/0741932517697446>
- Ladnier-Hicks, J., McNeese, R. M., & Johnson, J. T. (2010). Third grade reading performance and teacher perceptions of the Scott Foresman Reading Street program in title I schools in south mobile county. *Journal of Curriculum & Instruction, 4*(2), 51–70. <https://doi.org/10.3776/joci.2010.v4n2p51-70>
- Larson, A. L., An, Z. G., Wood, C., Uchikoshi, Y., Cycyk, L. M., Scheffner Hammer, C., Escobar, K., & Roberts, K. (2020). Social validity in early language interventions for dual language learners: A systematic review of the literature. *Topics in Early Childhood Special Education, 40*(1), 39–51.
<https://doi.org/10.1177/0271121419901289>
- Lathouras, M., Westerveld, M. F., & Trembath, D. (2019). Longitudinal reading outcomes in response to a book-based, whole class intervention for students from diverse cultural, linguistic, and socio-economic backgrounds. *Australian Journal of Learning Difficulties, 24*(2), 147–161.
<https://doi.org/10.1080/19404158.2019.1640755>
- Leal, P. (2016). Connecting reading and writing using children’s literature in the university L2 classroom. *Reading in a Foreign Language, 27*, 199–218.
<https://files.eric.ed.gov/fulltext/EJ1078427.pdf>

- Lindahl, K. M., & Sayer, P. (2018). Early EFL instruction and L1 literacy. *Colombian Applied Linguistics Journal*, 20(2), 180–190.
<https://doi.org/10.14483/22487085.12900>
- Lonigan, C. J., & Burgess, S. R. (2017). Dimensionality of reading skills with elementary-school-age children. *Scientific Studies of Reading*, 21, 239–253.
<https://doi.org/10.1080/10888438.2017.1285918>
- Lucas, C., Hood, P., & Coyle, D. (2021). Blossoming in English: Preschool children's emergent literacy skills in English. *Journal of Research in Childhood Education*, 35(3), 477–502. <https://doi.org/10.1080/02568543.2020.1742256>
- MacSwan, J. (2017). A multilingual perspective on translanguaging. *American Educational Research Journal*, 54, 167–201.
<https://doi.org/10.3102/0002831216683935>
- Maghsoudi, M. (2021). Contributions of motivation to read in L2, proficiency and L1 reading strategy awareness to L2 reading. *Reading & Writing Quarterly: Overcoming Learning Difficulties*.
<https://doi.org/10.1080/10573569.2021.1931591>
- Makhlouf, A. M. E. S. (2021). Saudi Schools' Openness to Change in Light of the 2030 Vision. *American Journal of Educational Research*, 9(1), 52-60.
<https://doi.org/10.12691/education-9-1-6>
- Maluch, J. T., & Sachse, K. A. (2020). Reading in developing L2 learners: The interrelated factors of speed, comprehension, and efficiency across proficiency levels. *TESL-EJ*, 24(1), 1–17. <https://files.eric.ed.gov/fulltext/EJ1257221.pdf>

- Merriman, N. A., Sexton, E., McCabe, G., Walsh, M. E., Rohde, D., Gorman, A., Jeffares, I., Donnelly, N.-A., Pender, N., Williams, D. J., Horgan, F., Doyle, F., Wren, M.-A., Bennett, K. E., & Hickey, A. (2019). Addressing cognitive impairment following stroke: systematic review and meta-analysis of non-randomized controlled studies of psychological interventions. *BMJ Open*, *9*(2), Article e024429. <https://doi.org/10.1136/bmjopen-2018-024429>
- Ministry of Education. (2020). *Education and vision 2030* (n.d.). <https://www.my.gov.sa/wps/portal/snp/aboutksa/EducationInKSA>
- Mitchell, B., & Alfuraih, A. (2017). English language teaching in the Kingdom of Saudi Arabia: Past, present, and beyond. *Mediterranean Journal of Social Sciences*, *8*(2), 317 – 325. <https://doi.org/10.5901/mjss.2017.v8n2p317>
- Mondt, K., Struys, E., Van den Noort, M., Balériaux, D., Metens, T., Paquier, P., Van De Craen, P., Bosch, P., & Denolin, V. (2011). Neural differences in bilingual children's arithmetic processing depending on language of instruction. *Mind, Brain & Education*, *5*(2), 79–88. <https://doi.org/10.1111/j.1751-228X.2011.01113.x>
- Moon, A. L., Wold, C. M., & Francom, G. M. (2017). Enhancing reading comprehension with student-centered iPad applications. *TechTrends*, *61*(2), 187–194. <http://doi.org/10.1007/s11528-016-0153-1>
- Murphy, A. F., & Torff, B. (2019). Teachers' beliefs about rigor of curriculum for English language learners. *Educational Forum*, *83*(1), 90–101. <https://doi.org/10.1007/s11528-016-0153-1>

- Northwest Evaluation Association. (2011). *Technical manual for measures of academic progress and measures of academic progress for primary grades*.
[https://www.richland2.org/RichlandDistrict/media/Richland-District/AdvancED/Standard 5/5.1/5-1-NWEA-Technical-Manual-for-MAP-and-MPG.pdf](https://www.richland2.org/RichlandDistrict/media/Richland-District/AdvancED/Standard%205/5.1/5-1-NWEA-Technical-Manual-for-MAP-and-MPG.pdf)
- Northwest Evaluation Association. (2016). Linking the North Carolina EOG Assessments to NWEA MAP Tests. *Northwest Evaluation Association*.
- Northwest Evaluation Association. (2020). Not all assessment data is equal: Why validity and reliability matter. Retrieved from: <https://www.nwea.org/resource-center/resource/not-all-assessment-data-is-equal-why-validity-and-reliability-matter/>
- OECD. (2016). *PISA 2015 results (Volume 1): Excellence and equity in education*. PISA. OECD.
- Olaniran, S. O., & Baruwa, I. B. (2020). Ethical considerations in adult and community education research in Nigeria: issues and perspectives. *International Journal for Educational Integrity*, 16(1), 1–10. <https://doi.org/10.1007/s40979-020-00057-3>
- O’Neal, C. R., Boyars, M. Y., & Riley, L. W. (2019). Dual language learners’ grit, engagement, and literacy achievement in elementary school. *School Psychology International*, 40(6), 598–623. <https://doi.org/10.1177/0143034319875176>
- Paige, D. D., Smith, G. S., Rasinski, T. V., Rupley, W. H., Magpuri-Lavell, T., & Nichols, W. D. (2019). A path analytic model linking foundational skills to Grade 3 state reading achievement. *Journal of Educational Research*, 112(1), 110–120.

<https://doi.org/10.1080/00220671.2018.1445609>

Park, M., Zong, J., & Batalova, J. (2018). *Growing superdiversity among young U.S. dual language learners and its implications*. Migration Policy Institute.

<https://www.migrationpolicy.org/research/growing-superdiversity-among-young-usdual-language-learners-and-its-implications>

Peltier, C., Flores, M. M., & Strickland, T. K. (2021). Practical use of single-case research designs when testing mathematics interventions for students with learning disabilities. *Learning Disability Quarterly*. Advance online publication.

<https://doi.org/10.1177/07319487211010318>

Petursdottir, A. I., & Carr, J. E. (2018). Applying the taxonomy of validity threats from mainstream research design to single-case experiments in applied behavior analysis. *Behavior Analysis in Practice*, *11*(3), 228–240.

<https://doi.org/10.1007/s40617-018-00294-6>

Piasta, S. B. (2016). Current understandings of what works to support the development of emergent literacy in early childhood classrooms. *Child Development Perspectives*, *10*(4), 234–239. <https://doi.org/10.1111/cdep.121888>

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

<https://doi.org/10.3102/0034654315584685>

Puglisi, M. L., Hulme, C., Hamilton, L. G., & Snowling, M. J. (2017). The home literacy

environment is a correlate, but perhaps not a cause, of variations in children's language and literacy development. *Scientific Studies of Reading*, 21(6), 498–514.

<http://doi.org/10.1080/10888438.2017.1346660>

Raikes, H. H., White, L., Green, S., Burchinal, M., Kainz, K., Horm, D., Bingham, G., Cobo-Lewis, A., Clair, L. St., Greenfield, D., & Estraich, J. (2019). Use of the home language in preschool classrooms and first- and second-language development among dual-language learners. *Early Childhood Research Quarterly*, 47, 145–158. <https://doi.org/10.1016/j.ecresq.2018.06.012>

Relyea, J. E., & Amendum, S. J. (2020). English reading growth in Spanish-speaking bilingual students: Moderating effect of English proficiency on cross-linguistic influence. *Child Development*, 91(4), 1150–1165.

<https://doi.org/10.1111/cdev.13288>

Ríos, C., & Castellón, C. (2018). *Bilingual literacy development: Trends and critical issues* (EJ1188732) ERIC. <https://eric.ed.gov/?id=EJ1188732>

Sarisahin, S. (2020). Reading comprehension strategies for students with learning disabilities who are emergent bilingual. *Intervention in School and Clinic*, 56(1),

3–12. <http://dx.doi.org/10.1177/1053451220910731>

Saudi Arabia's Vision for 2030. (2017). *Education*.

https://www.vision2030.gov.sa/media/rc0b5oy1/saudi_vision203.pdf

SAVVAS. (2021). *Reading Street 2013 Common Core*.

<https://mysavvastraining.com/products/reading-street-common-core>

Sierens, S., Slembrouck, S., Van Gorp, K., Agirdag, O., & Van Avermaet, P. (2019).

Linguistic interdependence of receptive vocabulary skills in emergent bilingual preschool children: Exploring a factor-dependent approach. *Applied Psycholinguistics*, 40(5), 1269–1297.

<https://doi.org/10.1017/S01427164190002500>

Simon-Cerejido, G., & Méndez, L. I. (2020). Similarities and differences in the lexical-grammatical relation of young dual language learners with and without specific language impairment. *Clinical Linguistics & Phonetics*, 34(1/2), 92–109.

<http://dx.doi.org/10.1080/02699206.2019.1611926>

Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *Future of Children*, 26(2), 57–74. <https://doi.org/10.1353/foc.2016.0012>

Soland, J., & Sandilos, L. E. (2021). English language learners, self-efficacy, and the achievement gap: Understanding the relationship between academic and social-emotional growth. *Journal of Education for Students Placed at Risk*, 26(1), 20–44. <https://doi.org/10.1080/10824669.2020.1787171>

Soto, X., Olszewski, A., & Goldstein, H. (2019). A systematic review of phonological awareness interventions for Latino children in early and primary grades. *Journal of Early Intervention*, 41(4), 340–365. <http://doi.org/10.1177/1053815119856067>

Soto, X. T., Crucet-Choi, A., & Goldstein, H. (2020). Effects of a supplemental Spanish phonological awareness intervention on Latinx preschoolers' dual language emergent literacy skills. *American Journal of Speech-Language Pathology*, 29(3), 1283–1300. https://doi.org/10.1044/2020_AJSLP-20-00029

Spies, T. G., Lara-Alecio, R., Tong, F., Irby, B. J., Garza, T., & Huerta, M. (2018). The

effects of developing English language and literacy on Spanish reading comprehension. *Journal of Educational Research*, 111(5), 517–529.

<https://doi.org/10.1080/00220671.2017.1306686>

Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion programs on student achievement: Evidence from lottery data. *American Educational Research Journal*, 54(1), 282–306. <http://doi.org/10.3102/0002831216634463>

Thomas, A. S., & January, S.-A. A. (2021). Evaluating the criterion validity and classification accuracy of universal screening measures in reading. *Assessment for Effective Intervention*, 46(2), 110–120. <http://doi.org/10.1177/1534508419857232>

UNESCO. (2017). *Education for sustainable development goals: Learning objectives*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000247444>

Van Norman, E. R., & Parker, D. C. (2018). A comparison of common and novel curriculum-based measurement of reading decision rules to predict spring performance for students receiving supplemental interventions. *Assessment for Effective Intervention*, 43(2), 110–120. <http://doi.org/10.1177/1534508417728695>

van Staden, A. (2016). Proficient reading and in English as a second language: Considering evidence-based strategies from the national reading panel. *EDULEARN16 Proceedings 8th International Conference on Education and New Learning Technologies July 4th-6th, 2016 — Barcelona, Spain* Edited by L. Gómez Chova, A. López Martínez, I. Candel Torres IATED Academy. (pp. 8952-8957). <https://doi.org/10.21125/edulearn.2016.0095>

- Wang, H., & Lin, M.-F. (2019). Linking reading and writing with picture books: A literacy buddy approach in rural Taiwan. *TESOL Journal*, *10*: e434.
<http://doi.org/10.1002/tesj.434>
- Wawire, B. A., & Kim, Y.-S. G. (2018). Cross-language transfer of phonological awareness and letter knowledge: Causal evidence and nature of transfer. *Grantee Submission*, *22*(6), 443–461. <http://doi.org/10.1080/10888438.2018.1474882>
- Wijethilake, C. (2018). Proactive sustainability strategy and corporate sustainability performance: The mediating effect of sustainability control systems. *Journal of Environmental Management*, *196*, 569–582.
<https://doi.org/10.1016/j.jenvman.2017.03.057>
- Wilkerson, S. B., Shannon, L. C., & Herman, T. (2006). The efficacy study on scott foresman's reading street program: Year one.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.183.2515&rep=rep1&type=pdf>
- Wilkerson, S. B., Shannon, L. C., & Herman, T. (2007). *The efficacy study of Pearson Education's Reading Street program: Year two*.
<https://www.pearson.com/content/dam/one-dot-com/one-dot-com/ped-blogs/wp-content/pdfs/r3-scott-foresman-reading-street-benchmark-item-validation-study-year-1.pdf>
- Willard, J. A., Kohl, K., Bihler, L.-M., Agache, A., & Leyendecker, B. (2020). Family literacy activities and their interplay with family and preschool language environments: Links to gains in dual language learners' German vocabulary.

Early Education and Development, 32(2), 189–208.

<https://doi.org/10.1080/10409289.2020.1744404>

Williams, C., & Lowrance-Faulhaber, E. (2018). Writing in young bilingual children:

Review of research. *Journal of Second Language Writing*, 42, 58–69.

<https://doi.org/10.1016/j.jslw.2018.10.012>

Williams, K. J., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). A synthesis of reading and spelling interventions and their effects on spelling outcomes for students with learning disabilities. *Journal of Learning Disabilities*, 50(3), 286–297.

<https://doi.org/10.1177/00222194060390060501>

Zhang, H., & Wang, H. (2021). A more powerful test of equality of high-dimensional two-sample means. *Computational Statistics and Data Analysis*, 164, Article

107318. <https://doi.org/10.1016/j.csda.2021.107318>

Zimmerman, B. S., Rasinski, T. V., Was, C. A., Rawson, K. A., Dunlosky, J., Kruse, S.

D., & Nikbakht, E. (2019). Enhancing outcomes for struggling readers: Empirical analysis of the fluency development lesson. *Reading Psychology*, 40(1), 70–94.

<https://doi.org/10.1080/02702711.2018.1555365>