Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners

---

A Dissertation

Presented to

The Faculty of the School of Education and

Human Development

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Isabel Vargas, M.T.

May 2024

University of Virginia
School of Education and Human Development Registrar
Office of Admissions and Student Affairs

Ehd-registrar@virginia.edu
Ridley Hall 102D
417 Emmet Street
Charlottesville, VA 22903

Dissertation Approval Form

Student Full Name: Vargas,Obeida Isabel

Department Name: Curriculum, Instruction & Special Education

Degree Program: Education (PhD) Special Education

Date of Defense: 3/25/2024

This doctoral dissertation has been approved by the Graduate Faculty of the School of Education and Human Development in partial fulfillment for the degree of Doctor of Philosophy.

Approved Title of Doctoral Dissertation:

Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners

	Name	Department/University	Signature
Chair	Emily Solari	CISE, EHD, UVA	DocuSigned by: Emily Solari A91C0ABD03D549F...
Co-Chair (if applicable)			
Committee Member	Colby Hall	CISE, EHD, UVA	DocuSigned by: Colby Hall 267A6A1E886D4E3...
Committee Member	Tisha Hayes	CISE, EHD, UVA	DocuSigned by: Tisha Hayes A49CEA07756B4D7...
Committee Member	Jamie DeCoster	EDLF, EHD, UVA	DocuSigned by: Jamie DeCoster 1CE72F207EC9443...
Committee Member			
Committee Member			
Committee Member			
Student	Isabel Vargas	Curriculum, Instruction & Special Education	DocuSigned by: Isabel Vargas A6F8B9CF49D0493...

Dedication

I dedicate this dissertation to my fiancé Chris, you are my rock, and to my family, los amo muchísimo.

Acknowledgements

I am immensely grateful to all of those who have encouraged, supported, and guided me during my doctoral journey. Firstly, I want to express immense gratitude to my advisor and chair, Dr. Emily Solari. Your mentorship, encouragement, and generosity has made what originally felt impossible, possible. Your vast wealth of knowledge and willingness to provide opportunities for collaboration and dissemination have been invaluable. I am in constant awe of your tireless dedication to advocating for science-based reading instruction for all children, and you inspire me to do the same.

To all my professors at UVA, thank you for sharing your expertise and for challenging me to be a better scholar. Additionally, I want to thank my committee members, Dr. Colby Hall, Dr. Tisha Hayes, and Dr. Jamie Decoster. Thank you for taking the time to provide thoughtful questions and insightful feedback throughout these past four years. Colby, I am especially indebted to you for your unwavering support and constant encouragement at all times of the day. I cannot thank you enough for all your responses to countless emails, texts, and phone calls. I have cherished your mentorship, and it has truly helped shape me as a scholar. Further, I want to thank the entire RADR lab for all their support, guidance, and collaboration. Sam, Alisha, Katie, Katlynn, and Steph, I would not have made it across that finish line without you all.

I also want to thank my unbelievably brilliant, and equally as kind, CISE cohort. You all have played a huge part of my experience, and I am so thankful to have spent these last few years surrounded by such wonderful people. An extra special thank you to Dr. Cassidi Richmond. I genuinely would not be here if it was not for our friendship and your willingness to tackle things together every step of the way. I could not imagine going on this journey with anyone but you.

I am eternally grateful to the principals and teachers who opened their schools and classrooms to welcome me and my research. Because of you, I was able to conduct meaningful and rewarding field-based scholarship that provided me with a daily reminder of my love for students and teaching. Further, thank you to my VLP and Equity Center teams for your guidance and support throughout my research. Working at VLP and at the Equity Center has been one of the biggest highlights of my time as a doctoral student.

To my Charlottesville friends who act like family, I cannot say thank you enough. Ben, this whole adventure started with you calling me and asking, "Do you trust me?" Even though it may not have always seemed like it, I am so glad I did. I am eternally grateful for your friendship, willingness to help open doors to new opportunities, and encouragement. Lucy, you are the older sister I never had. You are the truest definition of a loving and dependable friend. Whether it is needing to get our work done or needing to go get a drink, you are always there. Your kindness, dedication, and hard work inspire me every day. ¡Wepa! Gloria, thank you for your friendship and for giving me a home away from home. You have welcomed me into your family, and I cannot express my gratitude enough.

Finally, I want to give a special thanks to my family for their unconditional love and support. Mom, you have kept my refrigerator and freezer full through seemingly endless weeks of reading and writing. You consistently pick up my daily FaceTime calls to keep me company during short breaks and always give me something to laugh about. I could not have made it through this program without your empanadillas and constant words of wisdom. Papi, thank you for the "good genes" and for adulting for me at times so I could focus on my research. I still want to be just like you when I grow up. Gabyding, thank you for always sharing an apple tart with me, for always being up for a much-needed sister date, and for being the best auntie to Looney. I

am so thankful to have had you so close by these last few years, and I will always want to be a package deal with you no matter where life takes us. Joshi and Marissa, thank you for your support and for giving me the sweetest niece imaginable. Pictures of Little P have gotten me through plenty of tough days. Lastly, Christopher, this achievement is *our* achievement. You have been my loudest and proudest cheerleader, my teammate, and my very best friend. Your encouragement and support have kept me grounded throughout these past four years, and it is because of you that I am where I am today. Thank you for loving me so beautifully and for supporting my dreams. The best is yet to come, future husband.

Funding Acknowledgement

The contents of this work were developed under a grant from the U.S. Department of Education, #H325D190048. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. Additionally, the study reported in Manuscript Three was supported by the Innovative, Developmental, Exploratory Award from the School of Education and Human Development from the University of Virginia.

Abstract

This three-manuscript dissertation explores science-based reading instruction for multilingual learners. Utilizing a latent profile analysis, the degree to which distinct learner profiles emerged was examined in Manuscript One using code-based and language-based measures administered in the beginning of first grade. Participants included 11,803 English learners and 34,129 non-English learners. Three early literacy profiles emerged for English learners while four profiles emerged for non-English learners. Both sets of profiles could be identified based on the severity of students' difficulties with component skills rather than the specificity of their difficulties. Manuscript Two, a systematic review, summarizes the current state of knowledge regarding the use and characteristics of reading interventions that contain instruction in language comprehension for ELs in grades K-5 with or at risk for reading difficulties. This study synthesizes research that utilized an experimental or quasi-experimental group design conducted between 2000 and 2023. Nine studies were included, and results from this systematic review provides promising evidence for the use of reading interventions that include instruction in subcomponent skills that contribute to language comprehension for elementary-aged ELs with or at risk for reading difficulties. Lastly, utilizing a convergent mixed methods design, Manuscript Three evaluates the acceptability, appropriateness, and feasibility of a seven-week listening comprehension and vocabulary intervention implemented by teachers for elementary-aged ELs and non-ELs. Findings from this study suggest that the intervention is acceptable, appropriate, and feasible when implemented in elementary school practice settings. These findings indicate promise for the successful implementation of the intervention by teachers in diverse classrooms that may ultimately impact students' language and literacy outcomes.

Contents

Linking Document……..……..……..……..……..……..……..…… 1

Manuscript 1: Examining the Heterogeneous Early Literacy Profiles of First-Grade Students Who Are English Learners……..... 16

Introduction…… 18

Literature Review……..... 20

Conceptual Framework…….. 20

Method……..…… 23

Results……..……..…… 30

Discussion and Implications……..……..……..... 35

References……..……..……..…… 43

Figures and Tables……..……..……..…….. 50

Manuscript 2: Language Comprehension Interventions for English Learners: A Research Synthesis……... 58

Introduction…… 60

Literature Review……..... 64

Conceptual Framework…….. 64

Method……..…… 72

Results……..……..…… 77

Discussion and Implications……..……..……..... 100

References……..……..……..…… 106

Figures and Tables……..……..……..……... 115

Manuscript 3: Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners…… 120

Introduction…… 121

Literature Review……..... 124

Conceptual Framework…….. 124

Method……..…… 138

Results……..……..…… 147

Discussion and Implications……..……..……...….. 159
References……..……..……..…… 173
Figures and Tables……..……..…….…… 182

Linking Document

Science-Based Reading Instruction for Elementary Students Who Are English Learners

Isabel Vargas

Department of Curriculum, Instruction, and Special Education

School of Education and Human Development, University of Virginia

Science-Based Reading Instruction for Elementary Students Who Are English Learners

Over the past few decades, the English learner (ELs) population has grown drastically across U.S. public schools. Currently, there are over 5 million ELs enrolled in the U.S. education system, accounting for 10% of the total K-12 student population (U.S. Department of Education, 2017). Recent federal- and state-level education legislation has encouraged and incentivized policy makers, school leaders, and practitioners to meet the instructional needs of the increasingly diverse student population by selecting and implementing interventions and practices that have research-based evidence for improving student literacy outcomes (Snyder et al., 2017). However, national assessment data has demonstrated that there are large and persistent discrepancies in reading achievement between ELs and their non-English learner (non-ELs) peers. Long-term trend data from the National Assessment of Educational Progress (NAEP) demonstrate that students currently classified as ELs ("current ELs") consistently score significantly lower than non-ELs on the NAEP literacy assessments in fourth, eighth, and twelfth grade (U.S. Department of Education, 2022).

These national data trends, combined with the exponential growth the nation has seen in this population of students over the past few decades, highlight the essential need to continue to identify effective instructional practices that improve ELs' early literacy outcomes. With the simple view of reading (SVR) as a conceptual framework, this set of studies explores three aspects of reading development and instruction for elementary student who are ELs: (a) the heterogenous nature of ELs' early literacy profiles; (b) the characteristics and effects of reading interventions that have language comprehension instruction on reading outcomes for elementary-age ELs with or at risk for reading difficulties; and (c) the feasibility of an oral language and listening comprehension intervention for elementary-age ELs and non-ELs.

Conceptual Framework

The SVR, an empirically validated framework, posits that reading comprehension is the product of two broad skill categories: code-based skills that contribute to word reading (e.g., PA, alphabet and phonics knowledge, decoding) and language-based skills that contribute to linguistic comprehension (e.g., vocabulary knowledge, knowledge of syntax/language structure; Gough & Tunmer, 1986; Hoover & Gough, 1990). The SVR states that both code-based and language-based skills contribute separately to students' reading development, and proficiency in both is necessary for successful reading comprehension. This framework holds that poor reading comprehension are a result of one of three conditions: (a) adequate code-based but weak language-based skills, (b) adequate language-based but weak code-based skills, or (c) weak code-based and language-based skills (Hoover & Gough, 1990).

The SVR describes code-based skills, which are also referred to as code-focused and code-related skills, as efficient word recognition or "the ability to rapidly derive a representation from printed input that allows access to the appropriate entry in the mental lexicon, and thus, the retrieval of semantic information at the word level" (Hoover & Gough, 1990, p. 130). In other words, it is the ability to decode words quickly and accurately. The SVR describes language-based skills, which are also referred to as meaning-based skills, language comprehension, listening comprehension, and linguistic comprehension, as "the process by which, given lexical (i.e., word) information, sentences and discourses are interpreted" (Gough & Tunmer, 1986, p. 7). It is the ability to derive meaning from both written and oral language. A student with the ability to decode isolated words quickly and accurately will not comprehend text if they are unable to derive meaning from the words they read (Gough & Tunmer, 1986). Similarly, a student who has a strong understanding of language and language structures will not be able to

successfully comprehend a text if they are unable to decode the words in the text (Gough & Tunmer, 1986). Therefore, it is essential that students develop both code-based and language-based skills early in the primary grades to support reading comprehension.

The SVR has been empirically validated for both non-ELs (Gough & Tunmer, 1986; Lonigan et al., 2018a) and ELs (Grimm et al., 2018; Hoover & Gough, 1990; Joshi et al., 2015). Hoover and Gough's (1990) seminal study provided initial support for the SVR by tracking the reading development of a sample of English and Spanish bilingual students through the early elementary grades. Their findings illustrated that code-based skills and language-based explained significant and unique variance in students' reading comprehension.

Importantly, research has also demonstrated that there are some differences in the ways in which the SVR explains reading comprehension for ELs compared to non-ELs. For example, the relative contribution of language comprehension to reading comprehension increases over time as texts become more complex and students become more proficient decoders and can focus more on gaining meaning from text (Adolf et al., 2010). Though linguistic comprehension has been found to explain more variance in reading comprehension for all students as they progress beyond the primary grades, the effects of linguistic comprehension are on average more responsible for reading comprehension difficulties for ELs than for non-ELs (Cho et al., 2019).

It is important to note that though the same reading component skills appear to be important for both groups of readers, research suggests that there is heterogeneity in early literacy profiles for both ELs and non-ELs (e.g., Cabell et al., 2011; Miciak et al., 2022). Students within each group vary in their development of code-based and language-based component skills. This heterogeneity in these early literacy profiles suggests that instruction should be data-driven and tailored to meet students' unique instructional needs.

Heterogeneity of Early Literacy Profiles

Recent research investigating early literacy development suggests there is within-group variability in ELs' early reading component skill performance (Ford et al., 2013; Gonzalez et al., 2016; Lesaux & Kieffer, 2010; Lonigan et al., 2018b; Miciak et al., 2022; Ren et al., 2019; Solari et al., 2022) similar to their non-EL peers (Cabell et al., 2011; Cabell et al., 2013; Grimm et al., 2018; Norwalk et al., 2012; Ozernov-Palchik et al., 2017). This heterogeneity is evident when using students' performance on code-based tasks (i.e., phonological awareness, decoding), as well as when in conjunction with language-based tasks (i.e., vocabulary). Additionally, this heterogeneity has been demonstrated when also considering the influence of ELs' home language on literacy development (i.e., using measures in both English and students' home language; Gonzalez et al., 2016; Solari et al., 2022). Within-group variability in ELs' early reading skills would be expected as ELs represent considerable cultural and linguistic diversity who differ considerably in their prior exposure to English (U.S. Department of Education, 2017). Further, ELs possess varying levels of home language and literacy proficiency when entering school, which impacts their English literacy development (Goldenberg, 2020).

The heterogeneity in ELs' foundational reading skills suggests that educators should not group ELs based solely on their limited English proficiency. Treating them as a homogenous group of students who benefit from the same types of instructional interventions neglects their linguistic diversity. Instead, to improve ELs' foundational literacy skill development and learning outcomes, educators should utilize both code-based and language-based assessments that are valid and reliable for linguistically diverse populations, and then use these results to design instruction that is specifically aligned with each student's unique instructional needs.

Language Comprehension Instruction

While research has demonstrated that there is heterogeneity is ELs' language-based skills (e.g., Ford et al., 2013; Gonzalez et al., 2016; Lesaux & Kieffer, 2010; Lonigan et al., 2018b), by definition, ELs are not yet proficient in English. They are developing foundational literacy skills while simultaneously developing their English language skills (Goldenberg, 2020). Although language comprehension can be developed in part through exposure to language in the environment, research has demonstrated that explicit instruction can support, and is often required, to develop requisite language skills (Connor et al., 2011; Silverman et al., 2022). This is especially true for linguistically diverse learners. Explicit vocabulary instruction is essential and should be strongly emphasized for all ELs as language-based skills, such as vocabulary knowledge, have been found to be strong predictors of later reading achievement for both ELs and non-ELs (Lindsey et al., 2013).

Language-based skills rely on various components: students' background knowledge; the breadth and depth of their vocabulary; their understanding of language structures including syntax and semantics, verbal reasoning; their understanding of morphology and syntax; and their linguistic knowledge (Scarborough, 2001). As ELs acquire these component skills, their foundational English language knowledge can be expected to impact their reading comprehension especially as students get older, the relative contribution of language comprehension to reading comprehension increases, and texts become more complex (Adolf et al., 2010; Cho et al., 2019).

Several meta-analyses have recently examined the effects of language comprehension interventions for elementary-age students (Elleman et al., 2009; Fitton et al., 2018; Goodwin & Ahn, 2013; Larson et al., 2018; Marulis & Neuman, 2010; Rodge et al., 2019; Silverman et al.,

2020; Stahl & Fairbanks, 1986). These syntheses focused on individual components of language comprehension instruction separately, such as examining the effects of vocabulary interventions (Elleman et al., 2009; Marulis & Neuman, 2010; Stahl & Fairbanks, 1986), specific language-based interventions such shared book reading interventions (Fitton et al., 2018), or morphology interventions (Bowers et al., 2010; Goodwin & Ahn, 2013). Though there are varying results, most of these syntheses and meta-analyses found that language comprehension interventions have positive effects on measures aligned with the intervention and smaller, yet still positive effects, on measures of general language knowledge. With the increasing size and diversity of the EL population, one topic that requires further exploration is the characteristics of language comprehension interventions for ELs with or at risk for reading difficulties.

Dissertation Overview

Manuscript 1: *Examining the Heterogeneous Early Literacy Profiles of First-Grade Students Who Are English Learners*

In the first manuscript, I examined the heterogeneity of early literacy profiles of English learners and non-English learners. Utilizing a latent profile analysis, the degree to which distinct learner profiles emerged was examined using code-based and language-based measures administered in the beginning of first grade. Participants included 11,803 ELs and 34,129 non-ELs. Three early literacy profiles emerged for ELs while four profiles emerged for non-ELs. Both sets of profiles could be identified based on the severity of students' difficulties with component skills rather than the specificity of their difficulties. This study extended the literature by taking the resulting profiles in both samples and utilizing them to predict performance on a measure of broad reading comprehension administered at the end of first and second grade. Results indicated that the profile that was associated with the greatest success on the later

measures of reading comprehension for both samples included the strongest performance on measures of both code-related and language-related skills. Results highlight the heterogeneity of early literacy skills within the EL and non-EL populations and demonstrate the importance of designing instruction that addresses the severity of a student's skill deficit. This manuscript was published in *Reading and Writing* in May, 2023.

Manuscript 2: *Language Comprehension Interventions for English Learners: A Research Synthesis*

In the second manuscript, I led a systematic literature review to examine the characteristics of instruction targeting language comprehension (e.g., vocabulary, syntax, morphology, and listening comprehension) for ELs with or at risk for reading difficulties in grades K through 5. This study synthesized research that utilized an experimental or quasi-experimental group design conducted between 2000 and 2023. Previous syntheses of language comprehension research have focused on the components of language comprehension instruction separately such as examining the effects of vocabulary interventions (Elleman et al., 2009; Stahl & Fairbanks, 1986) or morphology interventions (Bowers et al., 2010; Goodwin & Ahn, 2013) for students. In contrast to prior syntheses and meta-analyses, this synthesis extends the research by examining intervention characteristics that include various approaches and components of language comprehension instruction for ELs specifically with or at risk for reading difficulties. Additionally, this synthesis examines studies that include ELs from various language backgrounds across the United States. With the increasing size and unique cultural and linguistic characteristics of the EL population, identifying effective language comprehension interventions is essential to improve students' literacy outcomes and to inform policy and practice.

Manuscript 3: *Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners*

Lastly, for the third study, I evaluated the acceptability, appropriateness, and feasibility of a seven-week listening comprehension and vocabulary intervention for elementary-aged ELs and non-ELs. The intervention, Building Vocabulary and Early Reading Strategies (BVERS; Solari & Ciancio, 2014), is considered supplemental to the school's typical English language arts (ELA) curriculum and was not used as a substitute or replacement for core instruction. BVERS targets listening comprehension strategies and vocabulary through explicit instruction and the use of preselected children's storybooks to provide opportunities for instruction in one of the following four comprehension skills: (1) direct recall of information from texts; (2) making connections to background knowledge; (3) narrative retell; and (4) making inferences. A total of 10 teachers, 8 general K-2 education teachers, 1 English as a Second Language (ESOL) teacher, and 1 reading interventionist, participated in the study and implemented the BVERS for 30 minutes a day, four days a week, for 7 weeks to a total of 193 students, including ELs and non-ELs, at 2 schools. Following the intervention, teachers answered questions from a 16-item measure of acceptability, appropriateness, and feasibility and participated in a 45-minute semi-structured focus group interview to discuss their experience implementing the intervention. Fidelity data was also collected through observations and coded using a fidelity checklist to measure intervention adherence. A convergent mixed methods design merging both quantitative and qualitative data from teachers and students was used to comprehensively examine the intervention's implementation process. Importantly, the findings from this study suggest that BVERS is acceptable, appropriate, and feasible when implemented by teachers in elementary school

practice settings. These findings indicate promise for the successful implementation of BVERS in diverse classrooms that may ultimately impact students' language and literacy outcomes.

References

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

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144-179.

Cabell, S. Q., Justice, L. M., Konold, T., & McGinty, A. S. (2011). Profiles of emergent literacy skills among preschool children who are at risk for academic difficulties. *Early Childhood Research Quarterly*, *26*, 1-14.

Cabell, S. Q., Justice, L. M., Logan, J. A. R., & Konold, T. R. (2013). Emergent literacy profiles among prekindergarten children from low-SES backgrounds: Longitudinal considerations. *Early Childhood Research Quarterly*, *28*, 608-620.

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology*, *111*(6), 982-1000.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P. S., Bayraktar, A., Crowe, E. C., & Schatschneider, C. (2011). Testing the impact of child characteristics× instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, *46*(3), 189-221.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712-751.

Ford, K. L., Cabell, S. Q., Konold, T. R., Invernizzi, M., & Gartland, L. B. (2013). Diversity among Spanish-speaking English language learners: profiles of early literacy skills in kindergarten. *Reading and Writing*, *26*(6), 889-912.

Goldenberg, C. (2020). Reading wars, reading science, and English learners. *Reading Research Quarterly*, *55*(51), S131-S144.

Gonzalez, J., Pollard-Durodola, S., Saenz, L., Soares, D., Davis, H., Resendez, N., & Zhu, L. (2016). Spanish and English early literacy profiles of preschool Latino English language learner children. *Early Education and Development*, *27*(4). 513-531.

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257-285.

Gough, P. B., & Tunmer, P. B. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Grimm, R. P., Solari, E. J., McIntyre, N. S., & Denton, C. A. (2018). Early reading skill profiles in typically developing and at-risk first grade readers to inform targeted early reading instruction. *Journal of School Psychology*, *69*, 111-126.

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

Joshi, R. M., Ji, X. R., Breznitz, Z., Amiel, M., & Yulia, A. (2015). Validation of the simple view of reading in Hebrew—A Semitic language. *Scientific Studies of Reading*, *19*(3), 243-252.

Larson, A. L., Cycyk, L. M., Carta, J. J., Hammer, C. S., Baralt, M., Uchikoshi, Y., An, Z. G., & Wood, C. (2020). A systematic review of language-focused interventions for young children from culturally and linguistically diverse backgrounds. *Early Childhood Research Quarterly*, *50*, 157-178.

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

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology*, *95*(3), 482-494.

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018a). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education, 39(5)*, 260–273.

Lonigan, C. J., Goodrich, J. M., & Farver, J. M. (2018b). Identifying differences in early literacy skills across subgroups of language-minority children: A latent profile analysis. *Developmental psychology*, *54*(4), 1-48.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300-335.

Miciak, J., Ahmed, Y., Capin, P., & Francis, D. J. (2022). The reading profiles of late elementary English learners with and without risk for dyslexia. *Annals of Dyslexia*, *72*, 276-300.

Norwalk, K. E., DiPerna, J. C., Lei, P., & Wu, Q. (2012). Examining early literacy skill differences among children in head start via latent profile analysis. *School Psychology Quarterly*, *27*(3), 170-183.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D. E., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental Science*, *20*(5), 1-18.

Ren, L., Hu, B. Y., & Wu, Z. (2019). Profiles of literacy skills among Chinese preschoolers: Antecedents and consequences of profile membership. Learning and Individual Differences, 69, 22-32.

Rogde, K., Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, *15*(4).

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

Silverman, R. D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K–5 students' language and literacy outcomes. *Reading Research Quarterly, 55*, S207-S233.

Snyder, E., Witmer, S. E., & Schmitt, H. (2017). English language learners and reading instruction: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 136-145.

Solari, E. J., Zucker, T. A., & Grimm, R. P. (2022). An examination of the heterogenous nature of literacy development in Spanish-speaking preschool children. *Early Education and Development*, 1-17.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72-110.

U.S. Department of Education. (2017). *Our nation's English learners: What are their characteristics?* National Center for Education Statistics, Institute of Education Sciences.

U.S. Department of Education. (2022). National Assessment of Educational Progress (NAEP), 2020 and 2022 Long-Term Trend (LTT) Reading and Mathematics Assessments. National Center for Education Statistics, Institute of Education Sciences.

Manuscript 1

Examining the Heterogeneous Early Literacy Profiles of First-Grade Students Who Are English Learners

Isabel Vargas, Mia C. Daucourt, Sara A. Hart, Emily J Solari

Author Note

Isabel Vargas, M.T., Department of Curriculum, Instruction, and Special Education, University of Virginia School of Education and Human Development.

Mia C. Daucourt, M.S., Department of Psychology, Florida State University.

Sara A. Hart, Ph.D., Professor, Department of Psychology, Florida Center for Reading Research, Florida State University.

Emily J. Solari, Ph.D., Professor, Reading Education Program, Department of Curriculum, Instruction, and Special Education, Virginia Literacy Partnerships / Director, University of Virginia School of Education and Human Development.

Abstract

This study examined the heterogeneity of early literacy profiles of English learners and non-English learners. Utilizing a latent profile analysis, the degree to which distinct learner profiles emerged was examined using code-based and language-based measures administered in the beginning of first grade. Participants included 11,803 English learners and 34,129 non-English learners. Three early literacy profiles emerged for English learners while four profiles emerged for non-English learners. Both sets of profiles can be identified based on the severity of students' difficulties with component skills rather than the specificity of their difficulties. Resulting profiles in both samples were then utilized to predict performance on a measure of broad reading comprehension administered at the end of first and second grade. Results indicated that the profile that was associated with the greatest success on the later measures of reading comprehension for both samples included the strongest performance on measures of both code-related and language-related skills. Results highlight the heterogeneity of early literacy skills within the English learner and non-English learner populations and demonstrate the importance of designing instruction that addresses the severity of a student's skill deficit.

Keywords: English learners, reading, multilingual, early literacy

Examining the Heterogeneous Early Literacy Profiles of First-Grade Students Who Are English Learners

English learners (EL) constitute one of the largest and fastest growing student populations in public schools across the United States (U.S.; National Center for Educational Statistics [NCES], 2022). Currently, there are over 5 million EL enrolled in U.S. schools, accounting for 10% of the total K-12 student population (U.S. Department of Education, 2017). Nationally, the number of EL enrolled in schools grew 28% from 2000 to 2016, with 43 states seeing an increase in EL enrollment during this time (NCES, 2022). Florida, where the data analyzed in the present study were collected, is ranked fourteenth among states with high numbers of EL; still, about 1 in 10 public school students in Florida has the EL designation (NCES, 2022).

Historically, national assessment data has suggested that there are large and persistent discrepancies in reading achievement between EL and their peers who do not have the EL designation (non-EL). For instance, long-term trend data from the National Assessment of Educational Progress (NAEP) demonstrate that EL consistently score significantly lower than non-EL on the NAEP literacy assessments in fourth grade (U.S. Department of Education, 2022), suggesting that schools have not effectively met the needs of EL. However, it is important to point out that EL are a subgroup of the broader population of multilingual students that also includes “former EL” (i.e., multilingual students who once had the EL designation but lost this designation upon achieving grade-level English language proficiency standards) and “never EL” (i.e., multilingual students who are classified as proficient in English upon school entry; Kieffer & Thompson, 2018, p. 391). Focusing exclusively on the current-EL subgroup can obscure the progress that schools, educators, and families make in moving multilingual students toward

English proficiency and academic achievement (Kieffer & Thompson, 2018). Recent research has argued that EL NAEP reading scores may be misleading because NAEP literacy scores only include students currently classified as EL ("current EL") (Goodrich et al., 2021). When considering current EL and former EL, known as the "ever EL" group, NAEP scores demonstrate substantially improved reading achievement for EL and a decrease in the national achievement gaps between EL and non-EL students from 2003-2015 and (Kieffer & Thompson, 2018). Though the US educational system is progressing to meet the needs of linguistically diverse learners, we must continue to identify approaches to instruction that are effective in improving academic outcomes for EL.

EL are often characterized only by their limited English proficiency and treated by educators as a homogenous group of students who benefit from the same type of instructional intervention (Ford, 2013). However, research suggests that EL exhibit similar within-population variance in early literacy skills to that demonstrated by their non-EL peers (Lesaux & Kieffer, 2010; Gonzalez et al., 2016; Lonigan et al., 2018; Miciak et al., 2022; Solari et al., 2022a). Treating EL as a homogeneous population fails to acknowledge this diversity and assumes that a "one size fits all" model of instruction will be equally effective for all EL. To improve foundational literacy skill development and learning outcomes for EL, educators need to better understand within-EL differences so they can design instruction that is aligned with EL instructional needs.

Reading achievement disparities between EL and non-EL are driven by a combination of factors, including structural inequities that create an association between poverty and EL status, thus reducing educational opportunities and access to resources that impact learning (Siegel et al., 2001). They also reflect the critical contribution of language knowledge to literacy; by

definition, EL are not yet proficient in English, and having less English language knowledge would be expected to impact reading comprehension (Cho, 2019). The focus of the present study is on early literacy skills, including language comprehension skills as well as code-focused skills that facilitate word reading. Though schools and teachers may not be able to directly address broad systemic inequities, school personnel can be equipped to provide evidence-based instructional services to teach early literacy skills more efficiently and improve outcomes for students. To provide such instruction effectively, it is important to understand how early literacy profiles of EL may differ from their non-EL peers. This study examines the early literacy profiles of first grade EL and non-EL separately to better understand similarities and differences between the two groups and, ultimately, inform instructional practices. The study also examines how the early literacy profiles of the two groups are related to reading achievement longitudinally.

Early Literacy Skills and Their Contribution to Later Reading Achievement

Early literacy knowledge and skills like phonological awareness (PA), alphabet knowledge, vocabulary knowledge, and knowledge of syntax are important precursors to students' later reading achievement (National Early Literacy Panel [NELP], 2008). The simple view of reading (SVR), an empirically validated framework that applies to both EL and non-EL, posits that reading comprehension is the product of two broad skill categories: skills that contribute to word reading (e.g., PA, alphabet and phonics knowledge, decoding) and skills that contribute to linguistic comprehension (e.g., vocabulary knowledge, knowledge of syntax/language structure; Gough & Tunmer, 1986; Hoover & Gough, 1990). Proficiency in both word reading (i.e., code-based skills) and linguistic comprehension (i.e., language-based skills) is necessary for successful reading comprehension.

Research suggests that the same early indicators of reading risk predict later reading performance to a similar degree for EL and their non-EL peers (Gottardo et al., 2008; Grimm et al., 2018; Jared et al., 2010; Kieffer & Vukovic, 2012; Lesaux et al., 2007; Lindsey et al., 2003). For example, early PA is the dominant predictor of word reading for both populations of students (Gottardo et al., 2008; Jared et al., 2010; Lindsey et al., 2003). English vocabulary knowledge similarly predicts English reading comprehension for both EL and non-EL (Cho et al., 2019). Importantly, although the same reading component skills appear to be important for both groups of readers, research suggests that there is heterogeneity in learner profiles within groups (e.g., Cabell et al., 2011; Miciak et al., 2022) and nuances may also exist in terms of the degree to which component skills contribute to reading comprehension for EL compared to non-EL (Cho et al., 2019).

Heterogeneity of Early Literacy Profiles Among Non-EL

Studies investigating early literacy development in non-EL suggest there is within-group variability in reading component skill performance (Cabell et al., 2011; Cabell et al., 2013; Grimm et al., 2018; Norwalk et al., 2012; Ozernov-Palchik et al., 2017). Previous research using students' performance on code-based tasks (i.e., PA, decoding), sometimes in conjunction with language-based tasks (i.e., vocabulary), consistently derives three to six distinct latent profiles within samples of non-EL.

For example, Cabell et al. (2011) found that among preschool students who were at risk for academic difficulties and grew up in households experiencing economic disadvantage, five fall early literacy profiles emerged using four oral language measures (i.e., expressive and receptive grammar, expressive and receptive vocabulary) and four code-related measures (i.e., print concepts, alphabet knowledge, name writing, and rhyme): highest emergent literacy;

average oral language, strength in alphabet knowledge; high average oral language, weakness in alphabet knowledge; low average oral language, broad code-related weaknesses; and lowest oral language, broad code-related weaknesses. Profiles were differentiated by large variability in students' performance on code-based skills. The majority of students (63%) demonstrated average oral language skills and a small cluster of students (16%) demonstrated average oral language skills and alphabet knowledge. These profiles showed predictive validity when compared to mid-year teacher ratings of emergent literacy as well as end-of-year kindergarten literacy performance. Although there was some movement in profile membership, 65 percent of students remained in a similar profile (i.e., their original profile or a directly adjacent one) throughout their preschool year (Cabell et al., 2013). Performance on code-based measures also differentiated profiles in a study conducted by Ozernov-Palchik and colleagues (2017), where the authors derived six latent profiles within a sample of preschool and kindergarten-aged students based on code-based measures (i.e., PA, rapid automatic naming, and letter knowledge) and verbal short-term memory measures: high performers; average performers; low average performers; RAN deficit; PA deficit; and double-deficit (both RAN and PA).

In support of both the Cabell et al. (2011) and Ozernov-Palchik et al. (2017) studies, Grimm et al. (2018) also found that early literacy profiles were differentiated by code-based measures for both at-risk and not at-risk non-EL first grade students. Specifically, authors used code-based measures (PA, decoding), language-based measures (comprehension passages, Clinical Evaluation of Language Fundamentals, Fourth Edition; Qualitative Reading Inventory-5), and oral reading fluency to derive two distinct profiles for at-risk students and three profiles for not-at-risk students. Grimm and colleagues noted that phonological awareness (elision and blending words subtests of the Comprehensive Test of Phonological Processing) and decoding

(letter-word identification and word attack subtests of the Woodcock-Johnson) were best at differentiating the profiles of students. While oral reading fluency did distinguish the higher achieving profiles, this was not true when examining lower achieving profiles. Additionally, authors found that profile membership was predictive of measures of reading comprehension.

Heterogeneity of Early Literacy Profiles Among EL

While there is less research investigating latent profiles for EL, the research that exists suggests similar heterogeneity in EL early literacy development (Ford et al., 2013; Gonzalez et al., 2016; Lesaux & Kieffer, 2010; Lonigan et al., 2018; Miciak et al., 2022; Ren et al., 2019; Solari et al., 2022a). This heterogeneity is evident both when studies use only English code-based and language-based measures and when they consider the influence of EL first language (L1) on literacy development (i.e., using measures in both English and students' L1). As an example of a study that used English-language measures and focused on code-based skills alone, Ford et al. (2013) found that within a large sample of Spanish-speaking EL in kindergarten, four distinct literacy profiles emerged based on English measures of PA, alphabet knowledge, and phonetic spelling: highest early literacy skills; average phonological awareness and phonetic spelling, strength in alphabet knowledge; average phonological awareness, weakness in alphabet knowledge and phonetic spelling; lowest literacy skills. The profiles with stronger code-based skills were associated with stronger performance on later measures of word reading and skills that contribute to word reading. While most EL fell into the lowest early literacy skills profile, 71% demonstrated at or above average PA and 43% demonstrated at or above average phonetic spelling skills.

Several studies employed both English-language and L1 measures and explored performance on both code-based and language-based measures. For example, Gonzalez et al.

(2016) demonstrated heterogeneity in the early literacy development of EL by deriving four profiles using Spanish and English oral language measures as well as measures of Spanish PA and letter knowledge in a sample of Spanish-speaking preschoolers who lived in households experiencing economic disadvantage. The most prevalent group among the profiles demonstrated mixed performance in English and Spanish oral language skills and a weakness in Spanish code-related skills. Authors found that students' early oral language ability, regardless of language dominance, was significantly associated with later English listening comprehension, a subcomponent skill of reading comprehension, at the end of preschool.

Like Gonzalez et al. (2016), Solari et al. (2022a) derived four profiles using both code-based and language-based measures in English and Spanish in a sample of Spanish-speaking preschoolers enrolled in a dual language program: low English and Spanish; low Spanish and average English; average English and Spanish; and above average English and Spanish. The authors found that the majority of EL performed below the EL sample average on measures of Spanish literacy skills and within the average range on English literacy measures. The authors also found that students' fall literacy profiles predicted spring literacy skills.

Despite this emerging body of evidence demonstrating the heterogenous nature of early literacy development in EL, many studies have focused on code-based skills alone, and few studies have linked profiles to later reading comprehension outcomes. Tying early literacy profile performance to comprehension measures is essential, as reading comprehension is the ultimate goal of reading. Examining how learner profiles are related to reading comprehension longitudinally will better inform the field of the patterns of subcomponent skills associated with stronger reading comprehension and help guide instruction.

Current Study

This study builds on previous research that has described the heterogeneous nature of EL early literacy skills. Utilizing a latent profile analysis, the degree to which distinct learner profiles emerge when using code-based (DIBELS LNF, PSF, NWF, ORF) and language-based measures (PPVT) measured in the fall of first grade was investigated. The current study extends the literature by examining how early literacy profiles from the start of first grade (i.e., the fall semester) were related to a measure of broad reading comprehension assessed by a state standardized assessment at the end of first and second grade (i.e., in the spring semester). The following research questions were addressed:

1. How many EL early literacy profiles emerge at the start of first grade, and which measures best distinguish profiles?
2. How many non-EL early literacy profiles emerge at the start of first grade and how do they compare to the early literacy profiles that emerge within the EL student sample?
3. How are EL first grade fall early literacy profiles related to performance on a measure of broad reading comprehension in the spring of first and second grade?

Method

Participants

The present study used data from Florida's Progress Monitoring and Reporting Network (PMRN), a statewide educational database housing standardized assessment data for all Florida public school students in kindergarten through twelfth grade. First grade reading data came from the fall and spring of the 2006-2007 school year, and second grade reading data came from the spring of the 2007-2008 school year. This study focused on typical literacy development for EL

and non-EL; students with exceptionalities were not included except for students who are orthopedically impaired or gifted.

In Florida, students are classified as an EL during the school enrollment process based on the Home Language Survey (HLS) filled out by their guardian. The HLS contains three main questions: a) Is a language other than English used in the home? b) Does the student have a first language other than English? and c) Does the student most frequently speaks a language other than English? Students who indicated "yes" for any item on the HLS are required to take an aural/oral English language proficiency test completed within 20 days of the HLS completion date. If a "yes" response is given only to the first (a) question, the student is not required to be placed in an EL program pending the results of the English language proficiency test. If a "yes" response is given to the second (b) and/or third (c) question, the student is placed in the ESOL program pending the results of the English proficiency test. In first grade, students are considered Non-English Speaker or Limited English-Speaking and placed in the school's ESOL instructional program if they score below the 51st percentile on the aural/oral English language proficiency test. The operationalization of EL status in this study included any student who had previously been or was currently enrolled in an EL class at any point in their school career (i.e., ever EL). The final sample included 11,803 EL and 34,129 non-EL. The demographic makeup of the final EL and non-EL samples in terms of sex and ethnicity is reported in Table 1.

Procedure and Measures

All reading-related assessments were administered as a part of regular school attendance, with the measures used to create the latent profiles administered in the fall semester of first grade and the Standardized Achievement Test (SAT10) administered in the spring semester of each school year. The measures used to create the latent profiles included the Peabody Picture

Vocabulary Test (PPVT-III, Dunn & Dunn, 1997) and four Dynamic Indicators of Basic Early Literacy Skills (DIBELS, Good & Kaminski, 2002) subtests: Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF).

Socioeconomic Status (SES)

SES was operationalized as student eligibility to receive free or reduced-price lunch (FRPL) and was drawn from the PMRN. FRPL status is a factor coded as 0/1 that indicates whether students are eligible for subsidized lunch. Free lunch is offered to students from households with an income at or below 130 percent of the poverty income threshold, and reduced-price lunch is available for students from households with an income in the range of 130-185 percent of the poverty threshold (Kena et al., 2015).

Peabody Picture Vocabulary Test (PPVT-III, Dunn & Dunn, 1997)

The PPVT was designed for both children and adults and has been validated for EL and non-EL (Dunn et al., 2005). The PPVT (Cronbach's alpha = .81-.95 in early elementary school students; Goriot et al, 2021) is widely used to capture students' receptive vocabulary skills. During the assessment, students are shown a series of four pictures and asked to point to the picture that illustrates a word (noun, verb, or adjective) that the test administrator states aloud (e.g., "Show me 'horse'").

Dynamic Indicators of Basic Early Literacy Skills (DIBELS, Good & Kaminski, 2002)

The DIBELS is a set of seven reading measures given in kindergarten through sixth grade and administered individually, which are used for benchmarking (i.e., tracking a student's progress in relation to other students) and progress monitoring (i.e., tracking a student's progress over time relative to an initial level of performance). Alternate-form reliability coefficients range

from .73-.97 across measures and grade level (Good & Kaminski, 2011). Not all measures are administered in all the grades; as students become proficient in foundational reading skills, some of the measures are designed to be phased out and other measures focused on more complex skills introduced. Because the latent profiles were created with first grade assessments, only four out of the seven DIBELS reading measures available were included in the analyses.

Letter Naming Fluency (LNF) is a measure of print awareness that measures a student's ability to rapidly (i.e., within one minute) identify upper- and lowercase letters arranged in a random order on a page.

Phoneme Segmentation Fluency (PSF) measures phonemic awareness, focusing on a student's ability to segment as many orally-presented words into phonemes as possible within one minute. For example, if the teacher orally presents the word "cat," students should respond with /c/ /a/ /t/.

Nonsense Word Fluency (NWF) measures knowledge of letter-sound correspondences and the ability to blend letters into words. Students are presented with a page of three-letter, consonant-vowel-consonant pseudowords (e.g., sog) and asked to read as many words as possible within one minute. Students get credit for every word read correctly or for correctly saying a sound represented by a letter even when they are unable to read the whole pseudoword.

Oral Reading Fluency (ORF) is a measure of decoding and fluency that measures a student's reading accuracy and fluency reading connected text. Students are asked to read three passages aloud for one minute each. Errors are scored when words are omitted or substituted and when hesitations last more than three seconds, but words self-corrected within three seconds are scored as accurate (Good & Kaminski, 2002). For the current analysis, the total number of words read correctly in one minute was used based on an average of all three passages.

Stanford Achievement Test (10th edition; Harcourt Brace, 2003)

The SAT-10 is a widely used standardized measure of reading comprehension that assesses initial understanding, interpretation, critical analysis, and awareness and usage of reading strategies. This untimed test is group-administered by classroom teachers in all Florida Reading First schools and scored by the test publisher. Students read a series of literary, informational, and functional text passages, which are followed by multiple choice questions that total to 54 items. The alpha coefficient for SAT-10 on a nation-wide representative sample of students was .88.

Data Analyses

Prior to creating the latent profiles, PPVT scores and each DIBELS measure (LNF, PSF, NWF, ORF) were residualized on FRPL status, race/ethnicity and gender to control for confounding. The resulting residualized outcome variables were converted to z-scores (i.e., standardized) based on the whole sample, then separated into EL and non-EL children in order to run a separate latent profile analysis (LPA) for each group. Given that students were nested within schools, standard errors were adjusted for nesting of children within schools using the type=COMPLEX command in Mplus. Finally, the LPA was conducted using a three-step approach known as the BCH method so that latent profiles were created based on Grade 1 DIBELS and PPVT assessments at the same time that Grade 1 end-of-year SAT10 reading scores and Grade 2 SAT10 reading scores were predicted (Asparouhov and Muthen, 2014). The BCH method is a three-step approach that starts off as a typical LPA in the first step, computes classification errors for each profile in the second step, and finally, uses the inverse logits of those individual-level error rates as weights in the to predict an additional outcome in the third step (Bolck et al., 2004). One advantage of this approach is that it does not require equal

variances across profiles, so variances could vary between profiles (Nylund-Gibson et al., 2019). The LPA analyses were conducted in Mplus version 7.4 (Muthén & Muthén, 1998-2017).

Latent profile analysis (LPA) operates under the assumption that the observed sample is a combination of individuals from different latent profiles, with individuals who have similar observed scores on a set of selected measures assumed to come from the same probability distributions and thus, belonging to the same profile (Vermunt & Magidson, 2002). LPA is conducted using a systematic model comparison approach to determine the best model based on a balance of parsimony, model fit, and the interpretability of the profiles for a series of models that increase by one profile (i.e., subgroup) at a time. When an increase in model fit is not statistically significant, it is an indication that that the more parsimonious model, with one fewer profile, should be chosen. Starting with a two-profile solution, the four DIBELS measures (LNF, PSF, NWF, ORF) and PPVT scores were used to systematically test solutions with increasing numbers of profiles. The fit statistics used to evaluate the model were the log-likelihood ratio test, entropy, Akaike information criterion (AIC), Bayesian information criterion (BIC), sample-size-adjusted BIC (SABIC), and the Lo-Mendell-Rubin adjusted likelihood ratio test (LMR-LRT) test and its accompanying *p*-values. For the log likelihood ratio test and entropy, higher values signify a better fit. For AIC, BIC, and SABIC, smaller values are indicative of better fit. For the LMR-LRT test, a higher value and a statistically significant *p*-value indicate a better fit.

Results

Descriptive statistics, including means (presented as sample-specific z-scores residualized on SES, race/ethnicity and gender), sample sizes, standard deviations, skew and kurtosis values for all achievement measures are presented in Table 2. The partial correlations, controlling for SES, among the reading assessments are presented in Table 3. The partial

correlations for EL are presented above the diagonal and the partial correlations for non-EL are presented below the diagonal.

LPA of EL

Table 4 presents all profile solutions tested for the LNF, PSF, NWF, ORF, and PPVT scores of the EL sample. However, when comparing the achievement patterns on the reading component skills between the model with four profiles and the model with three profiles, indeed, Profile 2 in the four-component model (the profile that was added when moving from three profiles to four) did not reveal patterns among the component skills that provided any additional information to meaningfully distinguish Profile 2 as an important profile to include in the modeling. As such, the three-profile solution appears to be the most translatable and parsimonious in terms of providing unique, meaningful information on reading component skills of EL in a real-world context, so the three-profile solution was chosen as the final model to represent EL. The figures comparing the two best-fitting models for EL are presented in Figure 1. The estimated mean scores for the three profiles of EL on the DIBELS and PPVT measures are reported in Table 5.

As shown on the top part of Figure 1, one of the profiles in the EL group, Profile 1, performed more than one standard deviation below the sample average on all assessments, except for ORF, which was about .60 SD below average. The next EL group, Profile 2, performed *just* below the sample average on all assessments, ranging from .21 to .37 below average, apart from PSF, which was almost at the sample average, just .08 SD below. The third and final EL group, Profile 3, demonstrated performance a full standard deviation above the sample average on all assessments, except for LNF and PSF which had mean scores .80 and .89 SD above the sample average. The three profiles were consistent across all measures, with scores

on all reading and vocabulary tests following the same rank order for each profile (i.e., one profile had the highest scores on all assessments, one profile had the second-highest score on all assessments, and so on); however, there were some interesting patterns that emerged to distinguish the component abilities demonstrated by each profile. The highest-performing profile (Profile 3), which represented approximately 15% of the sample, had especially high NWF and PPVT scores (1.31 and 1.22 SD above average, respectively), ORF scores just over 1 SD above average (1.03) and LNF scores just 1/10 SD below 1 SD above average (.89), and PSF scores ~ 1/5 SD short of being 1 SD above average (.81), indicating that they had exceptionally strong NWF and PPVT (vocabulary) skills. Going forward, this profile will be referred to as the "High Nonsense Word Fluency and Vocabulary Skills" profile. The lowest-performing profile (Profile 1), which included approximately 8.3% of the sample, had especially low PSF scores (~2 SD below average), quite low LNF scores (1.65 SD below average), NWF and PPVT scores that were about equal (~1.3 SD below average), and comparably high ORF scores (< 1 SD below average). Going forward, this profile will be referred to as the "Very Low Phonological Skills" profile. The majority of the sample (~77%) scored just below average and about the same across all assessments (~0.1 to ~0.4 SD below average), with phonemic awareness (PSF) showing the highest performance. This group will be referred to as the "Mostly Equal Component Skills" profile. Overall, results show that the small, high-performing group demonstrated high nonsense word fluency and vocabulary ability, whereas the small low-performing group had especially low phonemic awareness but demonstrated ORF performance that was almost equal to the majority of the EL sample.

LPA of Non-EL

Table 6 presents all profile solutions tested for the LNF, PSF, NWF, ORF, and PPVT scores of the L1 sample. The fit indices indicated that the four-profile model was the best fit for non-EL. The estimated mean scores for the four profiles of non-EL on the DIBELS and PPVT measures are reported in Table 8 and also presented in Figure 2.

As shown in Figure 2, two of the profiles in the non-EL group demonstrated performance above the sample average and the other two performed below the sample average on almost all assessments. The highest-performing profile (Profile 4), which represented approximately 9% of the sample, had especially high ORF and NWF scores (~2.5 SD above average and ~2 SD above average, respectively), LNF scores just over 1 SD above average (1.04), PPVT scores just below 1 SD above average (.79), and showed the lowest performance on the PSF assessment (about .35 SD above average). From this point on, this group will be referred to as the "Very High Oral Reading Fluency Skills" profile. The lowest-performing profile (Profile 1), which included approximately 9% of the sample, had especially low PSF scores (~2 SD below average), and all other subtest scores were about equal (0.6-0.8 SD below average), except for PPVT scores, which were slightly higher than other tests (.4 SD below average). Going forward, this group will be referred to as the "Very Low Phonological Skills Non-EL" profile. The majority of the sample (~60%) belonged to Profile 2 that had scores above average on the PSF, average on the PPVT, and below average on the LNF, NWF, and ORF measures, with the lowest performance on ORF (almost ~.5 SD below average). This group will be referred to as the "Poor Letter Naming and Reading Fluency Skills" profile. The final 21% of the sample represented the second-highest performing profile, with just above average scores on all five tests, ranging from approximately 0.3 to 0.6 SD above average and will be referred to as the "Mostly Equal Component Skills Non-

EL" profile. Thus, about one fifth of non-EL children demonstrated approximately equal just-above-average performance across all five component reading skills, and the majority of the sample showed average performance with a fluency deficit. Both the small high-performing group and the small low-performing group demonstrated a relative phonological deficit compared to the other reading components. The high-performing group also demonstrated high fluency ability, whereas the small low-performing group demonstrated relatively high vocabulary skills compared to the other component tests.

Predicting SAT-10 Grade 1 and 2 Scores Using the BCH Method

After the latent profiles were created using only the five component skill tests, the BCH weights were used to train a model that predicted first- and second-grade SAT-10 scores. The mean SAT-10 scores for each profile are presented in Table 8 for EL and Table 9 for non-EL. As expected, the results showed that ELs scored lower than non-ELs on the SAT-10 in both first and second grade. The BCH-weighted LPA predicting ELs' first-grade SAT-10 scores included 11,803 EL and 429 schools. The results showed that all profiles performed statistically significantly differently from all other profiles, with Profile 3 ("High Nonsense Word Fluency and Vocabulary Skills" profile) scoring higher than both other profiles, Profile 2 ("Mostly Equal Component Skills" profile) scoring the next highest, with higher scores than Profile 1 and lower score than Profile 3, and Profile 1 ("Very Low Phonological Skills" profile) scoring the lowest, with lower scores than Profiles 2 and 3. The BCH-weighted LPA predicting non-ELs' SAT-10 scores included 34,128 non-EL and 579 schools. The results showed that all profiles performed statistically significantly differently from all other profiles, with Profile 4 ("Very High Oral Reading Fluency Skills" profile) scoring higher than all other profiles, and Profile 3 ("Mostly Equal Component Skills Non-EL" profile) scoring the next highest, with higher scores than

Profiles 1 and 2 and lower scores than Profile 4. Profile 2 (“Poor Letter Naming and Reading Fluency Skills” profile) scored higher than Profile 1 and lower than Profiles 3 and 4, and Profile 1 (“Very Low Phonological Skills Non-EL"profile) achieved the lowest scores, with lower scores than Profiles 2, 3, and 4.

Discussion

This study investigated the English reading profiles of EL and non-EL. The first goal was to determine whether distinct latent profiles could be identified in EL and non-EL samples using student performance on code-based (DIBELS LNF, PSF, NWF, ORF) and meaning-based measures (vocabulary, PPVT) in English measured in the fall of first grade. Next, the non-EL profiles were visually inspected to determine how the EL and non-EL profiles compared in terms of level of performance and patterns of subcomponent skills that made up each profile. Finally, in order to determine how the early skills profiles were related to later literacy achievement, analyses were run to determine the relationship between the profiles and performance on a measure of broad reading comprehension administered in the spring of first and second grade. This study builds on prior research examining heterogeneity in early literacy profiles of EL and extends the literature in important ways. First, it uses both code-based and language-based measures with a very large sample of EL. In contrast to EL in samples examined in most previous research, EL in the present sample spoke a variety of L1 languages (i.e., not only Spanish). Additionally, the current study links early literacy profiles to a measure of broad reading comprehension longitudinally. In the following sections, key findings are highlighted and situated within the existing research with EL.

Heterogeneous Reading Profiles

Consistent with previous research (Lesaux & Kieffer, 2010; Ford et al., 2013; Gonzalez et al., 2016; Lonigan et al., 2018; Ren et al., 2019; Miciak et al., 2022; Solari et al., 2022a), the current study reveals heterogeneity in EL early literacy skill development, just as it demonstrates heterogeneity in skill development for non-EL peers. Three literacy skill profiles emerged among first-grade EL while four distinct literacy skill profiles emerged for first-grade non-EL. Our person-centered approach identified specific patterns of performance that would not have been revealed had this study only utilized composite mean scores, highlighting important differences between EL and non-EL students. In addition, had the analysis only used composite scores without examining the relations between the variables, it would have relied on unjustified assumptions about the characteristics of the variables in the profile analysis.

The two sets of profiles, illustrated in Figure 1, suggest the presence of low-, medium-, and high-performing groups within both a sample of non-EL students and a sample of EL students in Grades 1, which was unknown prior to conducting the analysis. These results suggest the presence of a stronger, overall literacy factor underlying performance on all the measures. However, the differences between the literacy profiles are not consistent across the measures, suggesting that the skills represented by the measures may emerge at different points of the developmental trajectory for literacy or that some measures we used may have been less sensitive to knowledge or skill attained earlier in development.

The heterogeneity in early literacy profiles suggests that grouping EL and designing literacy instruction solely based on their EL status and English language proficiency is inappropriate. To improve students' foundational literacy skill development and learning outcomes, educators should design instruction that is aligned with EL instructional needs.

Importantly, the findings of this study suggest that profiles can be identified based on the severity of students' difficulties with component skills rather than the specificity of their difficulties. In other words, the profiles differ in students' level of performance on each component skill rather than being differentiated by a specific weakness in one component skill relative to adequate or strong performance in another (Miciak et al., 2022).

These results have important educational implications. Because the profiles differ in severity, interventions can be adjusted to meet the instructional needs of students in different reading profiles by adjusting dosage, rather than instructional foci, where students with more severe deficits receive greater dosage while maintaining the same instructional focus (Capin et al., 2021). Research has demonstrated that higher dosage reading interventions yield statistically significant higher effect sizes for elementary students with or at risk for reading difficulties and disabilities, supporting recommendations for educators to adjust instructional intensity by increasing dosage in order to accelerate reading gains (Hall et al., 2022; Vaughn et al., 2003). First grade students within the two lowest literacy profiles for both the EL and non-EL sample require the greatest instructional dosage across all early literacy skills.

Both EL and Non-EL students in our sample scored below the national average on the PPVT, with EL scoring almost 17 points lower and non-EL students scoring just under 4 points lower**.** Additionally, the overwhelming majority of EL (85%) had English vocabulary scores below the sample average. This aligns with intuitive expectations, given that EL are still developing English proficiency, including vocabulary knowledge. Still, it is an important result to acknowledge because vocabulary knowledge is an essential subcomponent skill of linguistic comprehension and can also aid in word identification. Both receptive and expressive vocabulary are predictive of later reading and should be an instructional focus for teachers of EL (Wise,

2007). EL who have developed strong conversational fluency in social English are commonly misidentified by educators as having strong English proficiency and vocabulary knowledge. However, there is a distinction between conversational fluency and a students' ability to understand and express academic language, concepts, and ideas orally and in written form (Cummins, 1979).

These findings suggest that targeted explicit vocabulary instruction is essential and should be strongly emphasized for all EL, as they are all still developing English proficiency. This is particularly important as EL progress into upper elementary grades. As students encounter more complex texts, there are differences in the relative contributions of word reading and linguistic comprehension to reading comprehension for all students and beyond the primary grades, vocabulary and other linguistic comprehension variables make a larger contribution to reading comprehension for all students; however, research suggests this is particularly true for EL (Cho, 2019).

While mean scores tended to be higher for non-EL as compared to EL, the profiles that emerged were remarkably similar across the two groups. The biggest mean difference between the two groups was in vocabulary knowledge, which was to be expected given that EL are still developing English proficiency. However, within both the EL and the non-EL samples, scores on all reading component skill tests followed the same rank order for each profile, where one profile had the highest scores on all assessments, another profile had the second-highest score on all assessments, and this patterned continued.

The similarities across the profiles for EL and non-EL should be considered in the context of existing longitudinal studies that suggests that the same early reading skills predict later reading achievement difficulties to a similar extent for EL and non-EL (e.g., Kieffer & Vukovic,

2012; Lesaux et al., 2007). Further, large bodies of research suggest that similar interventions are effective in improving early literacy skills for both EL and non-EL (e.g., Hall et al., 2022; Richards-Tutor et al., 2016; Solari et al., 2022b; Vaughn et al., 2006). Relative to non-EL, a greater number of EL in the sample had difficulties with both sets of reading skills (i.e., with those associated with decoding development and those associated with meaning making or comprehension), such that they would benefit from instructional intervention addressing both code-based and language-based skills. EL may need particularly intensive support when it comes to English vocabulary development. That said, both EL and non-EL fall into learner profiles that are differentiated on the basis of severity rather than specificity: on average, learner profiles in both groups will likely benefit from similar instructional approaches (i.e., approaches that involve similar components) delivered with varying intensity.

Relationship to Reading Comprehension

The third research question investigated the relation between first grade fall early literacy profiles and performance on a measure of broad reading comprehension in the spring of first and second grade. Both EL and non-EL literacy profiles in the fall of first grade were associated with later reading comprehension in the spring of first and second grade. The profile that was associated with the greatest success on later measures of reading comprehension for both EL and their non-EL peers included the strongest performance on measures of both code-related and language-related skills. This aligns with the simple view of reading, which posits that reading comprehension is the product of word reading (i.e., code-based skills) and linguistic comprehension (i.e., language-based skills), such that both are necessary for successful reading comprehension.

While the majority of the non-EL sample (91%) showed at or above average performance with a fluency deficit, the majority of the EL sample (85%) scored just below or substantially below average across all assessments. These disparities highlight the need to identify approaches to instruction that are effective in improving academic outcomes for EL. Intervention research demonstrates that young EL at risk for reading difficulties benefit from the same explicit, systematic approach to foundational skills instruction that non-EL with reading difficulties benefit from (Richards-Tutor et al., 2016; Vaughn et al., 2006). It is imperative that educators are guided by scientifically-based reading research and implement evidence-based literacy instruction to impact reading outcomes for all students, but specifically for linguistically diverse students.

Limitations and Future Directions

There are limitations of the current study that should be noted. EL represent a diverse population: while all EL have an L1 other than English, they vary in terms of country of origin, family economic advantage, age of initial exposure to English, English language proficiency, L1 proficiency, amount of exposure to literacy instruction in English and the L1, and home literacy experiences in any language. This study used only English measures of early literacy skills; L1 data, including languages spoken, are not available in the dataset, and the dataset does not include information on home language environment and program of instruction (e.g., English only instruction, dual language instruction). This limited the research questions to a sole focus on English language and literacy development and did not allow for an investigation of the role of the L1 and cross-linguistic contributors to literacy development in a bilingual sample. Future studies should include information about L1 language use at home and school as well as measures of early literacy skills in both English and students' L1, so that student early literacy

profiles account for a more comprehensive linguistic and literary skill set that includes the entirety of the students' linguistic repertoire and literacy skills that enable instructors to make more fully informed appropriate instructional decisions. Although there are a number of studies that have addressed the contribution of these factors with Spanish-speaking ELs (Gonzalez et al., 2016; Lonigan et al., 2018; Solari et al., 2022a), there is a dearth of research investigating the impact of these predictors on language and literacy for speakers of other L1 languages (e.g., Arabic, Chinese, and Vietnamese).

An additional limitation is that this study included only one measure of language comprehension. The PPVT provided a measure of receptive vocabulary; however, this study failed to account for additional components of language comprehension, including students' expressive vocabulary, knowledge of language structures, verbal reasoning, and literacy knowledge. As previously noted, measuring linguistic comprehension is particularly important as students progress beyond the primary grades, when vocabulary and other linguistic comprehension skills make a larger contribution to reading comprehension. Future research should include additional language-based measures to examine how components of language may affect student membership in literacy profiles.

Finally, this study's analysis did not examine learner profiles among multilingual students broadly. This study only examined profiles for first-grade students designated "EL" in the PMRN data set (i.e., any student who had previously been or was currently enrolled in an EL class at any point in their school career). It is important to point out that EL are a subgroup of a broader population of multilingual first graders that includes multilingual students who are classified as proficient in English upon school entry (Kieffer & Thompson, 2018). Multilingual students in the US are a broad and diverse group. In order to provide the most effective and efficient early

literacy instruction for multilingual students, it would be beneficial for future research to explore learner profiles for additional categories of multilingual students, distinguishing between current EL, multilingual students who previously had an EL designation but exited the program (i.e., former EL), and multilingual students who were classified as proficient in English upon school entry.

Conclusion

This study examined the early literacy skills and reading comprehension performance of 11,803 EL and 34,129 non-EL. Utilizing measures of early literacy collected at the beginning of first grade, analyses identified multiple distinct literacy profiles for each subgroup of students. Severity, rather than specificity, distinguished both the EL and non-EL student learner profiles, which has important implications for the planning and implementation of effective and efficient early literacy interventions for EL. These profiles predicted performance on a measure of broad reading comprehension at the end of first and second grade. These results add to a growing body of studies that demonstrate the heterogeneity of early literacy skills within the EL population and highlight the importance of designing instruction that addresses the severity of a student's skill deficit.

References

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural equation modeling: A multidisciplinary Journal*, *21*(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*(1), 3-27. https://doi.org/10.1093/pan/mph001

Cabell, S. Q., Justice, L. M., Konold, T., & McGinty, A. S. (2011). Profiles of emergent literacy skills among preschool children who are at risk for academic difficulties. *Early Childhood Research Quarterly*, *26*, 1-14. https://doi.org/10.1016/j.ecresq.2010.05.003

Cabell, S. Q., Justice, L. M., Logan, J. A. R., & Konold, T. R. (2013). Emergent literacy profiles among prekindergarten children from low-SES backgrounds: Longitudinal considerations. *Early Childhood Research Quarterly*, *28*, 608-620. http://dx.doi.org/10.1016/j.ecresq.2013.03.007

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, No. 19, 121-129.

Capin, P., Cho, E., Miciak, J., Roberts, G., & Vaughn, S. (2021). Examining the reading and cognitive profiles of students with significant reading comprehension difficulties. *Learning Disability Quarterly*, *44*(3), 183–196. https://doi.org/10.1177/0731948721989973

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and

mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111(6),* 982-1000. https://doi.org/10.1037/edu0000332

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test--Third Edition (PPVT-III) [Database record]. PsycTESTS.

Ford, K. L., Cabell, S. Q., Konold, T. R., Invernizzi, M., & Gartland, L. B. (2013). Diversity among Spanish-speaking English language learners: profiles of early literacy skills in kindergarten. *Reading and Writing*, *26*(6), 889-912. https://doi.org/10.1007/s11145-012-9397-0

Gonzalez, J., Pollard-Durodola, S., Saenz, L., Soares, D., Davis, H., Resendez, N., & Zhu, L. (2016). Spanish and English early literacy profiles of preschool Latino English language learner children. *Early Education and Development*, *27*(4). 513-531. https://doi.org/10.1080/10409289.2015.1077038

Good, R. H., & Kaminski, R. A. (2002). DIBELS oral reading fluency passages for first through third grades. (Technical report no. 10). Eugene, OR: University of Oregon

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

Goriot, C., Van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2021). Using the Peabody Picture Vocabulary test in L2 children and adolescents: effects of L1. *International Journal of Bilingual Education and Bilingualism*, *24*(4), 546-568. https://doi.org/10.1080/13670050.2018.1494131

Gottardo, A., Collins, P., Baciu, I., & Gebotys, R. (2008). Predictors of grade 2 word reading and

vocabulary learning from grade 1 variables in Spanish-speaking children: Similarities and differences. *Learning Disabilities Research and Practice, 23(1),* 11-24. https://doi.org/10.1111/j.1540-5826.2007.00259.x

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

Grimm, R. P., Solari, E. J., McIntyre, N. S., & Denton, C. A. (2018). Early reading skill profiles in typically developing and at-risk first grade readers to inform targeted early reading instruction. *Journal of School Psychology, 69*, 111-126. https://doi.org/10.1016/j.jsp.2018.05.009

Hall, C., Dahl-Leonard, K., Cho, E., Solari, E. J., Capin, P., Conner, C. L., ... & Kehoe, K. F. (2022). Forty Years of Reading Intervention Research for Elementary Students with or at Risk for Dyslexia: A Systematic Review and Meta-Analysis. *Reading Research Quarterly,* 1-28. https://doi.org/10.1002/rrq.477

Harcourt, B. (2003). *Stanford Achievement Test – Tenth Edition: Technical data report*. San Antonio, TX: Author.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2,* 127-160. https://doi.org/10.1007/BF00401799

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2010). Early predictors of biliteracy development in children in French immersion: A 4-year longitudinal study. *Journal of Educational Psychology, 103(1),* 119-139. https://doi.org/10.1037/a0021284

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., & Velez, E. D., Nachazel, T., Dziuba, A., Smith, W., Nelson, V., Robles-Villalba, V., Soo, W., & Ballard,

D. (2015). The Condition of Education 2015. NCES 2015-144. National Center for Education Statistics.

Kieffer, M. J., & Vukovic, R. K. (2012). Components and context: Exploring sources of reading difficulties for language minority learners and native English speakers in urban schools. *Journal of Learning Disabilities, 45(5)*, 433-452. https://doi.org/ 10.1177/0022219411432683

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

Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology*, *99*(4), 821-834. https://10.1037/0022-0663.99.4.821

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*(3), 596-632. https://doi.org/10.3102/0002831209355469

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology, 95(3)*, 482-494. https://doi.org/10.1037/0022-0663.95.3.482

Lonigan, C. J., Goodrich, J. M., & Farver, J. M. (2018). Identifying differences in early literacy skills across subgroups of language-minority children: A latent profile analysis. *Developmental psychology*, *54*(4), 1-48. https://dx.doi.org/10.1037/dev0000477

Miciak, J., Ahmed, Y., Capin, P., & Francis, D. J. (2022). The reading profiles of late elementary English learners with and without risk for dyslexia. *Annals of Dyslexia*, *72*, 276-300. https://doi.org/10.1007/s11881-022-00254-4

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition.

Norwalk, K. E., DiPerna, J. C., Lei, P., & Wu, Q. (2012). Examining early literacy skill differences among children in head start via latent profile analysis. *School Psychology Quarterly*, *27*(3), 170-183. https://doi.org/10.1037/spq0000003

National Center for Education Statistics. (2022). English Learners in Public Schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved [date], from https://nces.ed.gov/programs/coe/indicator/cgf

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel. Executive Summary.* Washington, DC: National Institute for Literacy.

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. S*tructural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 967-985.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D. E., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental Science*, *20*(5), 1-18. https://doi.org/10.1111/desc.12471

Ren, L., Hu, B. Y., & Wu, Z. (2019). Profiles of literacy skills among Chinese preschoolers: Antecedents and consequences of profile membership. Learning and Individual Differences, 69, 22-32. https://doi.org/10.1016/j.lindif.2018.11.008

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The

effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children*, *82(2),* 144-169. https://doi.org/10.1177/0014402915585483

Siegel, P., Martin, E., & Bruno, R. (2001). Language use and linguistic isolation: Historical data and methodological issues. *Washington, DC: US Census Bureau*.

Solari, E. J., Zucker, T. A., & Grimm, R. P. (2022a). An examination of the heterogenous nature of literacy development in Spanish-speaking preschool children. *Early Education and Development*, 1-17. https://doi.org/10.1080/10409289.2021.1981200

Solari, E. J., Kehoe, K. F., Hall, C., Vargas, I., Dahl-Leonard, K., Richmond, C. L., Henry, A. R., Cook, L., Hayes, L., Conner, C., & Cho, E. (2022b). Effectiveness of Interventions for English Learners with Word Reading Difficulties: A Research Synthesis. *Learning Disabilities Research & Practice,* 1-17. https://doi.org/10.1111/ldrp.12286

U.S. Department of Education. (2017). *Our nation's English learners: What are their characteristics?* National Center for Education Statistics, Institute of Education Sciences.

U.S. Department of Education. (2022). *National Assessment of Educational Progress (NAEP), 2020 and 2022 Long-Term Trend (LTT) Reading and Mathematics Assessments*. National Center for Education Statistics, Institute of Education Sciences.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, *69(4)*, 391–409.

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E., & Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal, 107(2),* 153-180.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. *Applied latent class analysis*, *11*(89-106), 60.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech Language and Hearing Research*, *50(4)*, 1093-1109. http://doi.org/10.1044/1092-4388(2007/076)

Table 1

Demographic Composition of EL and non-EL

Demographic	% EL	% Non-EL
Sex		
Female	49.06	50.79
Male	50.94	49.21
Ethnicity		
White	3.13	34.08
Black or African American	17.55	46.41
Asian	2.53	1.06
Hispanic	75.63	12.99
Alaskan or Pacific Islander	0.13	0.40
Two or More Racial Identities	1.03	5.07

Note. EL = English learners; non-EL = non-English learners; N = 11,803 for EL and N = 34,129 for non-EL

Table 2

Raw Scores for Component Reading Skills of EL and Non-EL

EL									
Measure	Mean	SD	Min	Max	Skew	Kurt.	1st Q	Median	3rd Q
LNF	44.22	18.05	0	110	-.17	.05	33	45	56
PSF	35.82	18.83	0	121	-.16	-.61	22	38	49
NWF	34.53	22.28	0	192	1.09	2.84	19	33	46
ORF	13	16.72	0	135	2.56	8.25	2	8	16
PPVT	83.01	13.38	40	157	.25	.45	74	81	92
Non-EL									
Measure	Mean	SD	Min	Max	Skew	Kurt.	1st Q	Median	3rd Q
LNF	51.24	15.52	0	150	.11	.28	40	50	61
PSF	41.33	16.55	0	138	-.38	.22	33	43	52
NWF	40.47	22.94	0	219	1.46	4.22	26	38	50
ORF	19.47	22.64	0	205	2.21	5.93	5	12	25
PPVT	96.12	13.92	40	160	.05	-.32	86	97	106
Overall									
Measure	Mean	SD	Min	Max	Skew	Kurt.	1st Q	Median	3rd Q
LNF	49.44	16.50	0	150	-.07	.37	39	49	60
PSF	39.91	17.33	0	138	-.35	-.05	31	42	51
NWF	38.94	22.92	0	219	1.35	3.89	24	37	49
ORF	17.81	21.47	0	205	2.33	6.67	4	10	22
PPVT	92.75	14.93	40	160	.05	-.25	81	93	104

Note. N = 11,803 for EL, N = 34,129 for Non-EL, N = 45,932 for overall sample; Kurt = kurtosis, 1st Q = first quartile, 3rd Q = 3rd quartile. *SD* = standard deviation; Min = minimum value; Max = maximum value. DIBELS = Dynamic Indicators of Basic Early Literacy Skills; LNF = Letter Naming Fluency subtest; PSF = Phoneme Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; ORF = Oral Reading Fluency subtest; PPVT= Peabody Picture Vocabulary Test 3rd Edition; The PPVT is a norm-referenced test with a mean of 100 and an SD of 15.

Table 3

Partial correlations among reading assessments for EL and non-EL

Variable	1	2	3	4	5	6	7
1. G1 DIBELS LNF	-	.46	.69	.55	.32	.54	.40
2. G1 DIBELS PSF	.36	-	.47	.31	.29	.38	.30
3. G1 DIBELS NWF	.64	.37	-	.72	.29	.58	.42
4. G1 DIBELS ORF	.54	.22	.75	-	.29	.61	.42
5. G1 PPVT	.21	.24	.26	.29	-	.40	.37
6. G1 SAT10	.50	.30	.55	.60	.40	-	.63
7. G2 SAT10	.34	.21	.39	.42	.43	.62	-

Note. The values presented here reflect partial correlations controlling for free or reduced-price lunch status. The values above the diagonal correspond to EL, and the values below the diagonal correspond to non-EL. For EL, n = 11,802 for LNF, PSF, NWF, ORF, and PPVT, n = 7,194 for first and second grade SAT10. For non-EL students, n = 34,128 for LNF, PSF, NWF, ORF, and PPVT, n = 21,442 for first grade SAT10, n = 21,441 for second grade SAT10. All p-values are statistically significant at $p < .0001$

Table 4

Latent profile analysis fit statistics for EL

Profiles	*DF*	-2LL	AIC	BIC	BICC	Entropy	LMR LRT statistic	LMR LRT *p*-value
2	31	-67137.43	134336.86	134565.52	134467.01	0.84	11409.45	<.0001
3	42	-66214.21	132512.43	132822.22	132688.75	0.85	1828.70	<.0001
4	53	-65516.34	131138.67	131529.61	131361.18	0.77	1382.35	<.0001
5	64	-65201.53	130531.06	131003.13	130799.75	0.77	623.57	.1962
6	65	-64909.98	129969.95	130523.16	130284.82	0.72	587.56	.1426

Note. -2LL = -2 Log Likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICC = Sample Size-Corrected Bayesian Information Criterion; LMR LRT statistic/*p*-value = Lo-Mendell-Ruben Likelihood Ratio Test statistic and associated statistical significance.

Table 5

Latent profile analysis results for EL

Profile	*N*	% of sample	LNF mean score (*SE*)	PSF mean score (*SE*)	NWF mean score (*SE*)	ORF mean score (*SE*)	PPVT mean score (*SE*)
1	985	8.34	-1.65 (.13)	-1.96 (.03)	-1.32 (.04)	-0.61 (.01)	-1.21 (.07)
2	9077	76.90	-0.21 (.02)	-0.08 (.04)	-0.25 (.02)	-0.36 (.01)	-0.37 (.02)
3	1741	14.75	0.89 (.04)	0.81 (.04)	1.31 (.06)	1.03 (.04)	1.22 (.05)

Note. LNF = Letter Naming Fluency subtest; PSF = Phoneme Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; ORF = Oral Reading Fluency subtest; *SE* = standard error.

Table 6

Latent profile analysis fit statistics for non-EL

Profiles	*DF*	-2LL	AIC	BIC	BICC	Entropy	LMR LRT test statistic	LMR LRT *p*-value
2	31	-193781.44	387624.88	387886.45	387787.94	0.86	33690.03	<.0001
3	42	-191240.19	382564.39	382918.78	382785.31	0.74	5038.61	<.0001
4	53	-190186.7 8	380479.56	380926.77	380758.33	0.76	2088.64	<.0001
5	64	-189565. 80	379259.59	379799.62	379596.22	0.73	1920.83	<.0001
6	75	-189233.17	378616.34	379249.19	379010.84	0.66	1570.61	<.0001

Note. -2LL = -2 Log Likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICC = Sample Size-Corrected Bayesian Information Criterion; LMR LRT statistic/*p*-value = Lo-Mendell-Ruben Likelihood Ratio Test statistic and associated statistical significance.

Table 7

Latent profile analysis results for non-EL

Profile	N	% of sample	LNF mean score (*SE*)	PSF mean score (*SE*)	NWF mean score (*SE*)	ORF mean score (*SE*)	PPVT mean score (*SE*)
1	3168	9.28	-0.37 (.03)	-1.76 (.02)	-0.80 (.02)	-0.61 (.01)	-0.41 (.03)
2	20769	60.85	-0.20 (.01)	0.17 (.02)	0.30 (.01)	-0.46 (.01)	0.03 (.01)
3	7170	21.01	0.56 (.02)	0.31 (.02)	0.43 (.02)	0.48 (.04)	0.29 (.02)
4	3022	8.86	1.02 (.02)	0.35 (.02)	1.90 (.05)	2.55 (.05)	0.79 (.03)

Note. Means are sample-derived z-scores. LNF = Letter Naming Fluency subtest; PSF = Phoneme Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; ORF = Oral Reading Fluency subtest; *SE* = standard error.

Table 8

Grade 1 and 2 SAT-10 Scores for EL

Outcome	Profile	Mean (*SE*)
Grade 1 SAT-10	1	-1.17 (.03)
	2	--0.29 (.02)
	3	1.00 (.03)
Grade 2 SAT-10	1	-1.03 (.07)
	2	-0.17 (.02)
	3	0.72 (.03)

Note. SAT10 = Standard Achievement Test 10th Edition; *SE* = standard error. All *p*-values < .0001 for pairwise comparisons between profiles.

Table 9

Grade 1 and 2 SAT-10 Scores for Non-EL

Outcome	Profile	Mean (*SE*)
Grade 1 SAT-10	1	-0.87 (.03)
	2	-0.36 (.01)
	3	0.89 (.02)
	4	1.27 (.02)
Grade 2 SAT-10	1	-.60 (.04)
	2	-0.18 (.02)
	3	0.54 (.02)
	4	1.02 (.02)

Note. SAT10 = Standard Achievement Test 10th Edition. *SE* = standard error. All *p*-values < .0001 for pairwise comparisons between profiles.

Figure 1

Three-and Four-Profile LPA Results for EL

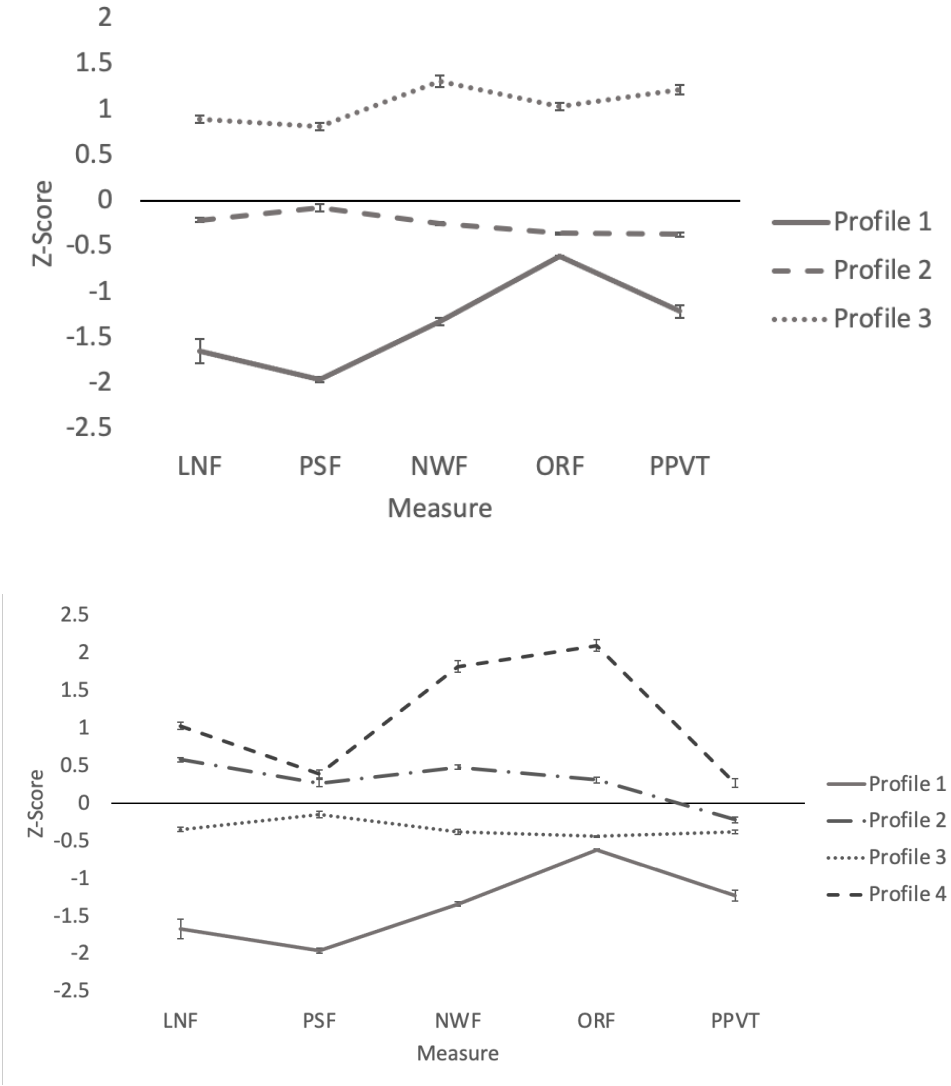
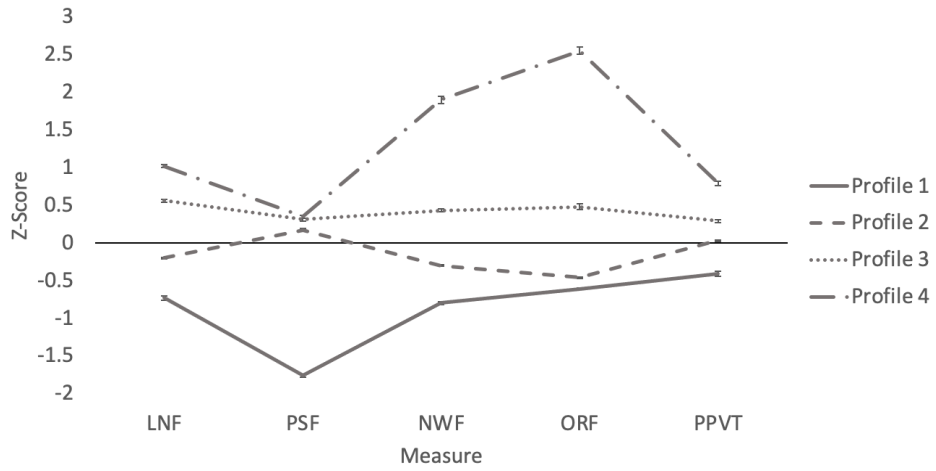


Figure 2

Four-Profile LPA Results for non-EL



Manuscript 2

Language Comprehension Interventions for English Learners:

A Research Synthesis

Isabel Vargas and Emily J. Solari

Author Note

Isabel Vargas, M.T., Department of Curriculum, Instruction, and Special Education, University of Virginia School of Education and Human Development.

Emily J. Solari, Ph.D., Professor, Reading Education Program, Department of Curriculum, Instruction, and Special Education, Virginia Literacy Partnerships / Director, University of Virginia School of Education and Human Development.

Abstract

This systematic review summarizes the current state of knowledge regarding the use and characteristics of reading interventions that contain instruction in language comprehension for ELs in grades K-5 with or at risk for reading difficulties. This study synthesizes research that utilized an experimental or quasi-experimental group design conducted between 2000 and 2023 and extends the research by examining intervention characteristics that include various approaches and components of language comprehension instruction for ELs with or at risk for reading difficulties. Nine studies met all inclusion criteria and were grouped into two categories: (1) interventions implemented with students in the early primary grades (K-1), and (2) interventions implemented with students beyond Grade 1 (grades 2-5). Narrative syntheses of each group of studies are provided. The results from this systematic review provide promising evidence for the use of reading interventions that include instruction in subcomponent skills that contribute to language comprehension for elementary-aged ELs with or at risk for reading difficulties. However, these results should be interpreted with caution due to the small number of included studies and the large variability in participant, intervention, and study design characteristics.

Keywords: English learners, at risk, elementary, language comprehension

Language Comprehension Interventions for English Learners: A Research Synthesis

Over the past few decades, a large body of empirical research conducted around the world has sought to inform how proficient reading develops and identify the most effective ways to teach literacy skills. Within federal and state-level education legislation, there has been growing momentum to encourage the use of scientifically based research to guide instructional practice (Snyder et al., 2017). The Every Student Succeeds Act of 2015 (ESSA) requires all public schools receiving federal funding to select and implement evidence-based interventions. Additionally, since 2013, 32 states and the District of Columbia have passed laws or enacted policies related to evidence-based reading instruction (Schwartz, 2022). Though these recent laws and policies have encouraged policy makers, school leaders, and practitioners to select and implement interventions and practices that have a demonstrated scientific record of being effective at improving student literacy outcomes, interventions can be differentially effective for different populations of students (Connor et al., 2011), and it is essential for policy makers and educators to consider the characteristics of the student for which an intervention is intended when implementing reading instruction.

Research has suggested that intervention effects may differ for English learners (ELs) when compared to their peers who do not have the EL designation (non-ELs; Foorman et al., 2018). ELs must develop the same foundational knowledge and skills that proficient English speakers develop to successfully read in English (Gillon, 2018; Goldenberg, 2020). However, ELs must develop these foundational skills while simultaneously developing their English language skills (Goldenberg, 2020), posing unique challenges for this population of students. ELs are one of the largest and fastest growing student groups in public schools across the United States (U.S.; National Center for Educational Statistics [NCES], 2022). Nationally, the number

of ELs enrolled in schools grew 28% from 2000 to 2016, with 43 states seeing an increase in EL enrollment during this time (NCES, 2022). Currently, there are over five million ELs enrolled in U.S. schools, accounting for 10% of the total K-12 student population (U.S. Department of Education, 2017). Though Spanish-speaking ELs make up over 76% of the total EL population in the nation, there are over 400 different languages spoken in schools (U.S. Department of Education, 2017). This student group represents considerable cultural and linguistic diversity: ELs differ considerably in their exposure to English as well as their home language and literacy proficiency when entering school (U.S. Department of Education, 2017).

National assessment data has demonstrated that there are large and persistent discrepancies in reading achievement between ELs and non-ELs. Long-term trend data from the National Assessment of Educational Progress (NAEP) demonstrate that students currently classified as ELs ("current ELs") consistently score significantly lower than non-ELs on the NAEP literacy assessments in fourth, eighth, and twelfth grade (U.S. Department of Education, 2022). However, it is important to point out that these results are not necessarily surprising. By definition, current ELs are not yet proficient in English, and having less English language knowledge and skills would be expected to impact reading comprehension.

Current ELs are a subgroup of the broader population of multilingual students which also include "former ELs", multilingual students who once had the EL designation but lost this designation upon achieving grade-level English language proficiency standards, and "never ELs", multilingual students who were classified as proficient in English upon school entry (Kieffer & Thompson, 2018). Focusing exclusively on current ELs obscures the progress that schools, educators, and families have made in helping the entire multilingual student population progress towards English proficiency and advanced academic achievement (Kieffer &

Thompson, 2018). When considering both current ELs and former ELs, a group of students referred to as “ever ELs”, NAEP scores from 2003 to 2015 demonstrate substantially improved reading achievement for ever ELs and a decrease in national achievement gaps between ever ELs and non-ELs (Kieffer & Thompson, 2018). However, due to the exponential growth the nation has seen in this population of students over the past few decades, along with their substantial projected population growth in the coming years, it is essential to continue to identify effective instructional practices that improve ever ELs’ early literacy outcomes.

One model for understanding reading development and comprehension for both ELs (Grimm et al., 2018; Hoover & Gough, 1990; Joshi et al., 2015) and non-ELs (Gough & Tunmer, 1986) is the simple view of reading (SVR). This empirically validated framework posits that reading comprehension is the product of two broad skill categories: code-based skills that contribute to word reading (e.g., PA, alphabet and phonics knowledge, decoding) and language-based skills that contribute to language comprehension (e.g., vocabulary knowledge, knowledge of syntax/language structure; Gough & Tunmer, 1986; Hoover & Gough, 1990). Both code-based and language-based skills contribute separately to students’ reading development, and proficiency in both is necessary for successful reading comprehension. Skilled readers typically have stronger abilities in both sets of skills while poorer readers have weaker skills in one or both components, increasing their risk for developing later reading difficulties (Hoover & Gough, 1990; Solari et al., 2022).

Research has shown some differences in how the SVR explains reading comprehension for ELs compared to non-ELs. The relative contribution of language comprehension to reading comprehension increases over time as texts become more complex and students become more proficient decoders and can focus more on gaining meaning from text (Adolf et al., 2010).

Though language comprehension has been found to explain more variance in reading comprehension for all students as they progress beyond third grade, research suggests that weak language comprehension skills are a greater source of reading comprehension difficulties for ELs than for non-ELs (e.g., August & Shanahan, 2006; Cho et al., 2019).

While language comprehension is well established to be an essential component of successful reading comprehension (Gough & Tunmer, 1986; Grimm et al., 2018; Hoover & Gough, 1990; Joshi et al., 2014), explicit language comprehension instruction in the early elementary grades oftentimes receives little attention as it competes with word-reading skills for limited instructional time (Silverman et al., 2020). However, research suggests that language comprehension develops cumulatively over time, and develops alongside decoding (Cunningham et al., 1997; Silverman et al., 2020), suggesting that explicit instruction in both is essential even in the earliest primary grades. Further, though language comprehension can be developed in part through exposure to language in the environment, explicit instruction can support, and is often required, to develop requisite language skills (Connor et al., 2011; Silverman et al., 2020). For example, explicit vocabulary instruction alone or in combination with implicit instruction during storybook reading has demonstrated a stronger effect on word learning for students in pre-K and kindergarten when compared to implicit exposure alone (Marulis & Neuman, 2010). This suggests that language comprehension instruction should be robust in the elementary grades to equip students with the language skills required to comprehend texts over time. This is particularly important for linguistically diverse learners with lower levels of English proficiency as they progress through school and language comprehension increasingly contributes towards reading comprehension.

Building on previous reviews of research, this study reports on the characteristics of interventions with language comprehension instruction (e.g., instruction in vocabulary, morphology, listening comprehension, knowledge of syntax/language structure) in elementary school for ELs with or at risk for reading difficulties. Previous syntheses of language comprehension research have oftentimes focused on the individual components of language comprehension instruction separately such as examining the effects of vocabulary interventions (Elleman et al., 2009; Stahl & Fairbanks, 1986) or morphology interventions (Bowers et al., 2010; Goodwin & Ahn, 2013). In contrast to these syntheses and meta-analyses, this study concentrated on examining the characteristics of various approaches and components of language comprehension instruction to help inform research, policy, and practice on literacy instruction for ELs with or at risk for reading difficulties.

Prior Syntheses and Meta-Analyses of Language Comprehension Instruction

Language comprehension is "the process by which, given lexical (i.e., word) information, sentences and discourses are interpreted" (Gough & Tunmer, 1986, p. 7). It is the ability to derive meaning from both written and oral language. Language comprehension relies on students' background knowledge, the breadth and depth of their vocabulary, understanding of language structures including syntax and semantics, verbal reasoning, understanding of morphology and syntax, and linguistic knowledge (Scarborough, 2001). Research has found that the interaction between students' language comprehension and code-related skills in the primary grades predicted their reading comprehension in third grade for both ELs and non-ELs (Kieffer & Vukovic, 2012). Further, components language comprehension, such as vocabulary knowledge and listening comprehension, are strong predictors of later reading achievement for both ELs and

non-ELs (Lindsey et al., 2003). Therefore, it is imperative to evaluate the effects of various approaches to teaching the foundational skills needed to build students' language comprehension.

Previous meta-analyses have examined the effects of language comprehension interventions for elementary-age students (Elleman et al., 2009; Fitton et al., 2018; Goodwin & Ahn, 2013; Larson et al., 2018; Marulis & Neuman, 2010; Rodge et al., 2019; Silverman et al., 2020; Stahl & Fairbanks, 1986). These syntheses focused on individual components of language comprehension instruction separately, such as examining the effects of vocabulary interventions (Elleman et al., 2009; Marulis & Neuman, 2010; Stahl & Fairbanks, 1986), listening comprehension interventions such as shared book reading (Fitton et al., 2018), or morphology interventions (Bowers et al., 2010; Goodwin & Ahn, 2013).

In their seminal meta-analysis, Stahl and Fairbanks (1986) found a statistically significant mean effect size of .97 (p <.01) for vocabulary instruction on reading comprehension outcomes containing taught words and a smaller effect of .30 (p <.01) for global measures of reading comprehension across 52 studies that included students from Grade 2 through college. Further, they found that providing definitional and contextual information of a word (i.e., the student is provided with a definition and exposed to the word in context) was the most effective vocabulary teaching method.

Elleman et al. (2009) meta-analyzed 37 vocabulary interventions in Grades pre-K-12. The authors examined the impact of vocabulary on reading comprehension outcomes and found that vocabulary instruction was associated with an increase in students' ability to comprehend text with an overall effect size of .50 (p <.01) on custom measures developed by the research team. However, vocabulary instruction was found to be less effective on standardized measures with an overall mean effect size of .10 (p =.08). Further, the authors found that interventions with higher

levels of discussion were associated with greater effects custom measures of vocabulary. Additionally, Elleman and colleagues found that vocabulary interventions had greater effects for students with reading difficulties than for students without reading difficulties. Though both Elleman et al. (2009) and Stahl and Fairbanks (1986) both found generally positive effects from vocabulary interventions, these effects were stronger on custom measures aligned with the intervention than on measure of global vocabulary knowledge.

Marulis and Neuman (2010) meta-analyzed the effects of vocabulary interventions on the receptive and expressive language of preschool and kindergarten aged children. The 64 studies produced an overall effect size of .88 (p <.001) on word learning in pre-K and kindergarten. In contrast with Elleman et al. (2009), the authors found that there was no significant difference between gains on vocabulary measures for students at risk and not at risk for learning difficulties. The effects the authors reported were larger than those found by Elleman et al. (2009). However, the authors noted that this difference could be due to the selection of studies, as Elleman and colleagues only included one study with pre-K-aged children and focused on print-related interventions.

One language-based intervention approach commonly used with young children is known as shared book reading. Fitton et al. (2018) examined the impact of shared book reading on language and literacy outcomes among ELs 12-years old or younger. The authors meta-analyzed 54 empirical studies published between January 1981 and April 2017. They found that the overall combined effect size for the impact of shared book reading on ELs' language and literacy outcomes was positive and moderate, 0.28 (p <.001). These results were robust to design characteristics and were only moderated by children's special education status: children with developmental disorders exhibited smaller amounts of growth from the shared reading

interventions. These results slightly differ to results from prior meta-analyses that have examined the effects of shared reading interventions with non-ELs (e.g., Lonigan et al., 2008; U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2015). These prior meta-analyses have found mixed effects for shared reading interventions on non-ELs' language comprehension and composite language skills, with positive effects on oral language outcomes, print knowledge, and writing, but no significant effects on alphabet knowledge, cognitive ability, phonological awareness, or reading readiness (Lonigan et al., 2008).

Most recently, Silverman et al. (2020) meta-analyzed studies published between January 2010 and January 2020 that examined the impact of language comprehension interventions in U.S. elementary schools. The authors examined peer-reviewed studies that utilized quasi-experimental or experimental designs with at least one treatment and one control or comparison condition. The authors examined the effects of language comprehension interventions across measures of language comprehension (syntax, vocabulary/semantics, or morphology), listening comprehension, and reading comprehension. Silverman and colleagues found positive and significant effects on custom vocabulary measures aligned with the intervention ($g = 1.27$, $p < .01$), custom listening comprehension measures ($g = 0.10$, $p < .01$), and custom reading comprehension measures ($g = 0.19$, $p < .01$), but did not find that these results generalized to standardized measures of these outcomes.

Additionally, the Silverman et al. (2020) investigated whether effects varied depending on participant characteristics. When examining the intervention effects for ELs, the authors also found positive vocabulary and reading comprehension effects, and that the interventions were equally or more effective for ELs compared with non-ELs. Further, results suggested that interventions that include more than one component of language (i.e., vocabulary and semantics,

morphology, and/or syntax) may be advantageous. However, unlike previous research that has suggested that combining decoding and language comprehension instruction is beneficial (e.g., Rosenthal & Ehri, 2008), Silverman and colleagues found that studies that included decoding instruction did not show differential effects. Further, the authors found that language comprehension interventions did not demonstrate significant effects on decoding. However, the authors advise that these results should be interpreted with caution due to the relatively small number of studies that included decoding instruction or decoding outcomes. Much more research is needed to understand the relationship between decoding and language comprehension instruction and outcomes.

Prior Syntheses and Meta-Analyses of Reading Interventions for ELs

In recent decades, there has been an increase in research conducted examining the effects of reading interventions for elementary ELs. Various meta-analyses and syntheses conducted during this time have examined and identified instructional methods and practices associated with improving reading outcomes for this student population and how these effects differ across participant and intervention characteristics (e.g., Cheung & Slavin, 2012; Graham et al., 2022; Li et al., 2021; Ludwig et al., 2019; Roberts et al., 2022; Snyder et al., 2017). Additionally, given that interventions can be differentially effective for different populations of students (Connor et al., 2011), recent research has also examined how intervention effects may differ for typically developing ELs compared to ELs with or at risk for reading difficulties (Baker et al., 2018; Boon & Barbetta, 2017; Richards-Tutor et al., 2016; Solari et al., 2022).

For example, Cheung and Slavin (2012) conducted a systematic review to identify the most effective reading programs and approaches for improving English reading outcomes for Spanish-dominant ELs in grades K-6, holding constant the language of instruction. The authors

examined whole-school and whole-class interventions along with small group and one-to-one supplemental interventions. Each of the 22 studies included at least one measure of reading comprehension, with many studies also including additional word-based and language-based outcome measures. Cheung and Slavin found a small overall effect size of .23 (p <.01); however, the authors note that there was substantial variation between studies. Instructional approaches that used cooperative learning, where ELs had extensive opportunities to use their developing language skills in meaningful contexts, and small group instruction produced larger effects than other approaches.

Richards-Tutor et al. (2016) examined the effectiveness of reading interventions for English learners in grades K-12 with or at risk for a learning disability. The authors reviewed 12 experimental studies published in peer-reviewed journals from 2000 to 2012. They found positive effect sizes on standardized and researcher-developed measures of phonological awareness (PA), phonics/word reading, fluency, reading comprehension, and listening comprehension but not for vocabulary. The authors note that the seven studies conducted in kindergarten and Grade 1 had multi-component interventions that focused on improving foundational skills, such as PA and phonics, and produced larger (moderate-to-large) and more consistent effect sizes than other interventions that included older students and primarily targeted vocabulary and comprehension. Richards-Tutor and colleagues also found that the effects of interventions for older ELs in Grade 4 and above were minimal.

Ludwig et al. (2019) extended Richards-Tutor et al.'s (2016) findings by examining the effects of reading interventions in 26 experimental and quasi-experimental studies with ELs regardless of whether they had or were at risk for reading difficulties. Consistent with Richards-Tutor, their findings suggest that reading interventions have positive effects on ELs' reading

outcomes, though they report larger effects than Richards-Tutor. While Richards-Tutor noted the largest effect on PA outcomes, Ludwig et al. noted largest effects on reading accuracy. However, consistent with Richard-Tutor et al.'s findings, they reported the smallest mean effect on reading comprehension outcomes.

Boon & Barbetta (2017) conducted a systematic review of the literature on reading interventions for ELs with learning disabilities (LD) in Grades K-5. Across the nine studies, outcome measures included standardized, curriculum-based, and research developed measures. The authors identified a series of reading interventions (i.e., computer-based constant time delay, graphic organizers, peer-tutoring, repeated reading with a vocabulary component) that were associated with positive effects on early reading outcomes for ELs with LD. Consistent with Cheung and Slavin (2012), the largest effect sizes on reading comprehension were identified for peer-tutoring interventions where students had multiple opportunities to use their developing language skills in meaningful contexts.

Most recently, Solari et al. (2022) examined the last forty years of reading intervention research for ELs. The participants were from multiple language backgrounds in grades K-5 and demonstrated word reading difficulties rather than broadly defined reading or language difficulties. The authors meta-analyzed effects in 10 experimental and quasi-experimental group design studies as well as 7 single case experimental design studies conducted between 1980 and 2020. Solari and colleagues included peer-reviewed articles and grey literature in which the intervention included instruction in at least one of following components: PA, phonics, decoding, encoding (spelling), word reading accuracy or fluency, passage reading accuracy, or fluency. The results supported the notion that explicit and systematic instruction in subcomponent literacy skills is effective for ELs who have reading difficulties, similar to the impact of reading

interventions for ELs without reading difficulties and non-ELs (Ludwig et al., 2019). The analysis revealed a statistically significant overall mean weighted effect of .31 (p <.01) on combined reading outcomes in group design studies, larger than those reported by Cheung and Slavin (2012); however, Solari et al. attribute this to their focus on strictly ELs with or at risk for word reading difficulties. Further, Solari and colleagues provide corroborating evidence that interventions that emphasize both code- and meaning- related skills are more effective for improving reading comprehension outcomes for ELs with word reading difficulties than interventions that focus on one set of skills in isolation (Al Otaiba et al., 2023; Baker et al., 2012).

Current Study

Though previous reviews of research have examined the effects of reading interventions that have language comprehension instruction for non-ELs or students without risk for reading difficulties, one topic that requires further exploration is the impact of these interventions on language and literacy outcomes for ELs with or at risk for reading difficulties. With the increasing size and unique cultural and linguistic characteristics of the EL population, identifying effective language comprehension interventions is essential to improve students' literacy outcomes and to help inform policy and practice. The current study builds on previous research that has described the effects of reading interventions for ELs by examining the characteristics of language comprehension instruction in elementary schools for ELs from various language backgrounds with or at risk for reading difficulties. The current study synthesizes research published in peer-reviewed journals that utilized experimental and quasi-experimental group design studies conducted between January 2000 and November 2023. The research question addressed is: What are the characteristics (i.e., study design, participant characteristics,

intervention characteristics, and control group characteristics) of included studies that examine the effects of reading interventions that have language comprehension instruction for ELs with or at risk for reading difficulties?

Method

Search Procedures

The review began by searching research articles published in English between January 1, 2000 and November 30, 2023, following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberatti et al., 2009). Studies included in this review were identified through a four-step process: (a) identification of studies through electronic databases; (b) identification of studies through other sources; (c) abstract screening; and (d) full-text review.

Identification of Studies through Electronic Databases

First, a search of electronic databases ERIC and PsycINFO was conducted with a combination of the following Boolean search terms, specifying that abstracts must include one of the following population identifiers (English learner*, English language learner*, language minority, emergent bilingual, multilingual learner*, limited English proficien*, English as an additional language, English to speakers of other language*, English as a second language, or dual language learner*) and a term that described reading interventions (reading interven*, reading instruction, language comprehension, listening comprehension, language comprehension, vocabulary, syntax, semant*, pragmat*, morph*, background knowledge, oral language, summar*, main idea, infer*, predict*, paraphrase, think aloud*, think-aloud*, reading, literacy, reading skill*, reading comprehension). This search yielded 7,975 studies, 1,212 of

which were duplicates and were removed prior to the abstract screening process. Figure 1 represents the search procedure and inclusion results at each stage of the search process.

Inclusion Criteria

To investigate the characteristics of interventions with instruction in language comprehension for ELs in K-5 with or at risk for reading difficulties, studies were included if they met the following criteria:

1. The study was published on or after January 1, 2000, and before November 30, 2023 and was set in the United States.
2. Participants were in grades K–5 or the mean age of participants was 5.0–11.0 years old.
3. At least 50% of study participants were ELs who were identified by the study's authors as having or being at risk for reading difficulties or results were disaggregated for a subset of students meeting this description. Studies were excluded if participants had intellectual disabilities, autism spectrum disorders, or sensory disabilities.
4. Instruction occurred over more than one session.
5. The intervention included instruction in at least one of the following components: vocabulary; syntax, semantics; morphological awareness; and listening comprehension; reading comprehension (e.g., summarizing, inference making, narrative retell); or oral language.
6. Instruction was delivered primarily in English (i.e., greater than 50% of instruction was delivered in English).
7. The setting was participants' schools (including after-school and summer interventions).

8. The study reported intervention effects on at least one of the following outcomes in English: language comprehension (syntax, vocabulary/semantics, or morphology), listening comprehension, or reading comprehension.
9. The study had an experimental or quasi-experimental group design with at least 15 participants per group and investigated the effects of a researcher-manipulated intervention.
10. There was an acceptable treatment-comparison contrast, including (a) business-as-usual (BAU) classroom reading instruction; (b) BAU exposure to nonreading instruction or no instruction; (c) a researcher-manipulated nonreading intervention (e.g., math); or (d) a research-manipulated intervention that did not include training in language comprehension. A group design study must have contrasted the outcomes of treated ELs to a similar group of ELs in an acceptable comparison condition.

Screening

The electronic database search yielded 6,763 unique articles. After a one-hour long training and achieving ≥90% with the lead author on a practice set of ten abstracts, a team of four reviewers, all doctoral students in a reading or special education program, used the Covidence systematic review software (Covidence, 2023) to screen each abstract. Studies were recommended for full-text review if the study met the inclusion criteria. If the abstract did not provide sufficient information to determine eligibility status, the study was recommended for full-text review. Each abstract was independently screened by two reviewers. Reviewers achieved 99% overall agreement during the screening process, and all disagreements were reviewed and resolved by the lead author. Abstract screening resulted in 81 texts that were reviewed in full for study eligibility.

Identification of Studies through Other Sources

After the electronic database search, a hand search was conducted for articles published within the past five years in the three journals that published the greatest number of included studies: *Learning Disability Quarterly, Reading Research Quarterly, and Exceptional Children*. This hand search included articles published on or after January 1, 2017, and before November 30, 2023. During this hand search, 1 additional article was identified for full-text review. Additionally, an ancestral search of relevant meta-analyses and syntheses that were conducted in the past five years was also conducted: Rodge et al., 2019; Larson et al., 2020; Silverman et al., 2020; Roberts et al, 2022; Solari et al., 2022. This ancestral search yielded an additional 42 studies. After an initial review of abstracts, 30 studies were reviewed in full for study eligibility.

Full-Text Review

Together, the search of electronic databases and ancestral search yielded a total of 111 studies that required full-text review. Prior to beginning full-text review, six reviewers, all doctoral students in either a reading or special education program, participated in a 3-hour training and were required to achieve ≥90% reliability with the first author when reviewing a practice set of three articles. During full-text review, 101 studies were excluded because they did not meet at least one eligibility criterion applied in the following order: the study was not accessible (n = 1); the study was conducted outside of the United States (n = 6); enrolled students were not identified as being with or at-risk for reading or language difficulties (n = 48); enrolled students were not identified as ELs or data was not disaggregated for ELs (n = 6); intervention did not address language comprehension(n = 6); intervention was implemented for fewer than two sessions, was not conducted in a school setting, or was conducted in a language other than English (n = 3); study did not utilize an experimental or quasi-experimental research design or

did not enroll at least 15 participants per group (n = 25); study did not include an eligible outcome measure (n = 1); and study was not published in a peer-reviewed journal (n = 6). Each text was independently reviewed by at least two members of the research team. Reviewers achieved 84% overall agreement during full-text review, and, and the lead author reviewed and resolved any disagreements. Of the 111 studies reviewed, a total of 9 studies met all inclusion criteria.

Data Analysis

Coder Training

Each of the 12 studies was independently coded by the first author and another member of the research team. Four coders, all of whom had participated in the full-text review process, participated in a two-hour training with the lead author and independently coded three practice articles to achieve a minimum of 90% interrater agreement with the first author on each set of characteristics coded for. All disagreements were reviewed and resolved by the lead author.

Coding Procedures

Once the studies were identified as being eligible, the research team summarized the information in the corpus of studies utilizing a detailed coding protocol. Information recorded included (a) study design, (b) participant characteristics, (c) intervention characteristics, and (d) control group characteristics. When examining study design, information was recorded for two types of experimental designs: (1) randomized controlled trail with at least two groups, and (2) quasi-experimental design with at least two groups. Quasi-experimental studies were included only if authors reported no statistically significant differences between groups at pretest on outcomes of interest or if the authors reported posttest means that were adjusted for pretest performance. Additionally, information regarding fidelity of implementation was also recorded.

Participant characteristics coded for included grade level(s) of participants, eligibility for free or reduced-price lunch, and language learner status. Additionally, this synthesis examined how language learner status was determined in each study and the first language of EL participants. Interrater agreement while coding for study design and participant characteristics was 93.10%.

For intervention characteristics, coders recorded the duration (i.e., total weeks), dosage (i.e., minutes per week), and frequency (i.e., times per week) during which the intervention was implemented. Further, coders recorded whether instruction was delivered in a one-on-one, small group, or whole class setting as well as the position of the individual(s) delivering the instruction (i.e., classroom teacher, special education teacher, bilingual teacher, ESOL teacher, reading specialist, paraprofessional or member of the research team). Additionally, the research team coded for incorporation of specific instructional components in two broad categories: (a) instruction in language-based skills that contribute to language comprehension (e.g., vocabulary, syntax, semantics, morphology, listening comprehension, reading comprehension, and oral language), and (b) instruction in code-based skills that contribute to word reading (e.g., phonological awareness, decoding, encoding, rote reading, reading fluency). Interrater agreement while coding for intervention characteristics was 92.10%. Comparison or control conditions were noted as being BAU or an alternative acceptable treatment. Interrater agreement while coding for the comparison condition characteristics was 100%.

Results

The purpose of this review is to summarize the current state of knowledge regarding the use and characteristics of reading interventions that contain instruction in language comprehension for elementary-aged ELs with or at risk for reading difficulties. First, the study

and intervention characteristics for all nine included studies are summarized below to better understand the broader context for this corpus of studies. Next, given the difference in the contribution of language comprehension to reading comprehension as students get older, studies were grouped into two categories: (1) interventions implemented with students in the early primary grades (K-1), and (2) interventions implemented with students beyond Grade 1 (grades 2-5). Narrative syntheses of each group of studies are provided below.

Study Features

Of the nine studies that met the selection criteria, four were published in the last 5 years of the of the 23-year span included in the current review (2000-2023), and none of the studies were published prior to 2006. Table 1 summarizes methodological characteristics (i.e., number of participants, EL determination, reading risk determination) and participant characteristics (i.e., grade level and ELs' first language) of the included studies.

Sample Characteristics

In all, 2,008 students were included with a wide range of 48 to 1,166 participants per study. One study included students in kindergarten (Neuman & Kaefer, 2018); five studies included students in Grade 1 (Baker et al., 2016; Filippini et al, 2012; Neuman & Kaefer, 2018; Vaughn et al., 2006a; Vaughn et al., 2006b); two studies included students in Grade 2 (Denton et al., 2014; Jacob et al., 2016); two studies included students in Grade 3 (Denton et al., 2014; Jacob et al., 2016); 4 studies included students in Grade 4 (Denton et al., 2014 ; Jacob et al., 2016; Toste et al., 2009; Wanzek & Roberts, 2012); and 3 studies included students in Grade 5 (Denton et al., 2014; Toste et al., 2009; Wanzek & Roberts, 2012). Table 2 summarizes the approximate number of total participants in each grade level. Grade 4 had the most students included while kindergarten had the least.

ELs comprised the entire student sample in five of the studies (Baker et al., 2016; Denton et al., 2014; Filippini et al., 2012; Vaughn et al., 2006a; Vaughn et al. 2006b). In the remaining four studies, ELs made up a portion of the sample and student outcomes were disaggregated for the subpopulation of students (Jacob et al., 2016; Neuman & Kaefer, 2018; Toste et al., 2009; Wanzek & Roberts, 2012). Six of the studies included Spanish-dominant ELs (Baker et al., 2006; Denton et al., 2014; Filippini et al., 2012; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012) while the remaining three studies did not report students' first language. EL status was determined by students' schools in four studies (Filippini et al., 2012; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012) and by parent report in one study (Baker et al, 2016). The remaining four studies did not report how EL status was established. In three of the studies, 50% or more of the participants experienced economic disadvantage, measured as the percentage of the participating students eligible to receive free or reduced-price school lunch (Jacob et al., 2016; Neuman & Kaefer, 2018; Wanzek & Roberts 2012). The remainder of the studies did not provide information about the socioeconomic status of the participant sample.

The method used to identify reading or language risk status varied by study. Six studies included either a norm- or criterion-referenced outcome measure as part of determining student reading risk (Filippini et al., 2012; Neuman & Kaefer, 2018; Toste et al., 2009; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012); four studies included either a teacher or district referral (Denton et al., 2014; Jacob et al., 2016; Toste et al., 2009; Wanzek & Roberts, 2012); two studies included a researcher created word-reading measure (Vaughn et al., 2006a; Vaughn et al., 2006b); and two studies utilized unspecified assessments (Baker et al., 2016; Jacob et al., 2016).

Intervention Characteristics

The inclusion criteria for this review required interventions to include instruction in a subcomponent skill that contributes to language comprehension. Eight interventions included instruction in vocabulary (Baker et al., 2016; Denton et al., 2014; Filippini et al., 2012; Jacob et al., 2016; Neuman & Kaefer, 2018; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012); one intervention included instruction in semantic relations (Filippini et al., 2012); three interventions included instruction in morphology (Filippini et al., 2012; Toste et al., 2008; Wanzek & Roberts, 2012); six interventions included instruction in listening (i.e. language) comprehension (Baker et al., 2016; Filippini et al., 2012; Jacob et al., 2016; Neuman & Kaefer, 2018; Vaughn et al., 2006a; Vaughn et al., 2006b); and six interventions provided instruction in reading comprehension (Baker et al., 2016; Denton et al., 2014; Jacob et al., 2016; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012). Though it was not a requirement for this review, all the interventions also included instruction in at least one subcomponent skill that contributes to word reading. Six interventions included instruction in phonological awareness (Baker et al., 2016; Denton et al., 2014; Filippini et al., 2012; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012); eight interventions included instruction in decoding (Baker et al., 2016; Denton et al., 2014; Filippini et al., 2012; Jacob et al., 2016; Toste et al., 2008; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012); four interventions included instruction in rote reading (Denton et al., 2014; Filippini et al., 2012; Vaughn et al., 2006a; Vaughn et al., 2006b); five interventions included instruction in encoding (Baker et al., 2016; Toste et al., 2008; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012); and seven interventions included instruction in reading connected-text (Baker et

al., 2016; Denton et al., 2014; Jacob et al., 2016; Toste et al., 2008; Vaughn et al., 2006a; Vaughn et al., 2006b; Wanzek & Roberts, 2012).

The length of the interventions ranged anywhere from 8 to 28 weeks. Most were implemented in small groups except for one intervention, which was delivered whole class (Neuman & Kaefer, 2018), and one was implemented in a one-on-one setting (Jacob et al., 2016). One intervention was implemented by certified teachers already working at the school (Neuman & Kaefer, 2018); one by certified teachers hired by the research team (Wanzek & Roberts, 2012); two by certified bilingual teachers hired by the research team (Vaughn et al., 2006a; Vaughn et al., 2006b); one by a combination of certified bilingual teachers and instructional assistants (Baker et al., 2016); two by university undergraduate or graduate students (Denton et al., 2014; Filippini et al., 2012); one by community volunteers (Jacob et al., 2016); and one by tutors hired by the research team (Toste et al., 2009).

All the included studies utilized an experimental or quasi-experimental design. Each study utilized a BAU comparison condition except for one study which utilized a researcher manipulated reading treatment that did not include instruction in language comprehension as the comparison condition (Filippini et al., 2012). Fidelity of implementation was monitored in all nine studies. Table 3 summarizes the intervention characteristics for each study.

Interventions with Students in K – Grade 1

Five studies examined the effects of reading interventions that included instruction in language comprehension on language and literacy outcomes for ELs at-risk for reading difficulties in grades K-1 (Baker et al., 2016; Filippini et al., 2012; Neuman & Kaefer, 2018; Vaughn et al., 2006a; Vaughn et al., 2006b). Vaughn et al. (2006a) examined the effects of Proactive Reading, a comprehensive literacy intervention that utilizes direct instruction

principles, used as a supplementary intervention for Spanish-dominant ELs at-risk for reading difficulties in Grade 1. To increase the likelihood that the students participating in the study represented true reading difficulties and were not a result of inadequate reading instruction, the intervention was implemented in schools that had an established record as being effective for bilingual students based on high student performance (at least 80% pass rate) on state-wide reading assessments.

To determine whether a student was at-risk for reading difficulties and eligible for the study, the authors first administered a series of four measures in English and Spanish: (a) the Letter-Word Identification (LWID) subtest from the Woodcock Johnson Language Proficiency Battery in English and Spanish, and (b) a researcher-created word-reading list in English and Spanish. Students were deemed eligible for the study if they scored below the 25th percentile for Grade 1 in both the Spanish and English LWID subtests and were unable to read more than one word from the researcher created word list. A total of 96 students who received core literacy instruction in English were deemed eligible and participated in the study. On average, they demonstrated oral language scores more than three standard deviations below the normative sample.

Fifty students were assigned to the comparison condition and 46 students were assigned to the treatment condition. Students in the comparison condition received the school's BAU core curriculum, Language Enrichment at one school and McGraw Hill Reading Series at the second school. Students in the treatment condition received Proactive Reading. Eight bilingual teachers provided the intervention in small groups of three to five students in the intervention condition for 50 minutes a day, 5 days a week, for a total of 28 weeks. This instruction was supplementary to the students' core reading curriculum.

Each Proactive Reading lesson utilized a direct instruction approach and consisted of 6 to 10 short activities focused on providing explicit and systematic instruction in phonemic awareness, letter knowledge, word recognition, text fluency, and comprehension strategies including narrative retell. Vaughn et al. (2006a) also embedded three to eight language support activities with scripts and visuals into each lesson. Additionally, the intervention emphasized oracy, vocabulary, and listening comprehension development by integrating these skills into the reading instruction as well as dedicating 10 minutes per lesson to the development of these skills. The lessons incorporated research-based instructional practices for ELs, such as the use of facial expressions, gestures, and extended opportunities to respond. Each lesson followed a quick pace and included teacher modeling, group practice, and individual student practice.

Results indicated that students who received the Proactive Reading intervention did not show significant differences in measures of letter naming, letter word identification, nonword repetition, oral language, and measures of fluency based on reading of connected text compared to students in the comparison condition. However, students in the intervention condition outperformed students in the comparison condition on measures of rapid letter naming, phonological awareness (letter sound identification), decoding (word attack), word reading fluency (word reading efficiency), and spelling. Students in both the intervention and comparison condition had a moderate increase on measures of oral language and a strong increase in comprehension. The authors suggest that a possible justification for the comparison condition's strong reading comprehension performance was because the current study provided each school's most struggling readers a supplemental intervention. This therefore may have freed up school resources and allowed for schools to provide students in the comparison condition with increased learning opportunities in smaller groups with increased attention from school personnel.

Vaughn et al. (2006b) replicated this study to examine the effects of Proactive Reading with a different sample of Spanish-dominant ELs in Grade 1 with varying oral language and literacy profiles. A total of 48 students were randomly assigned to condition with 24 students receiving the intervention and 24 students receiving BAU literacy instruction. Students across both groups demonstrated low English and Spanish decoding skills with a relative strength in English, and low oral language performance in both languages. Students in this study demonstrated higher English oracy and literacy scores at the beginning of first grade than the participants in Vaughn et al. (2006a), but still had relatively low language and literacy skills.

Results indicated that, like the students in the previous study (Vaughn et al., 2006a), students who received the Proactive Reading intervention outperformed students in the control condition on some measures of early reading skills including measures of English rapid letter naming, phonological awareness (letter-sound identification), and decoding (word attack). In contrast to Vaughn et al. (2006a), the authors also found that students who received the Proactive Reading intervention outperformed students in the control condition in passage comprehension. Interestingly, students in both conditions improved approximately 10 standard score points on a composite measure of oral language, although performance was still low.

The combined results of these studies suggest that ELs who are at-risk for reading difficulties and who demonstrate low levels of language and literacy skills can benefit from explicit and systematic instruction similar to non-ELs at risk for reading difficulties (Mathes et al., 2005). Importantly, students who received the intervention demonstrated growth in reading comprehension and significantly greater growth than ELs in the comparison condition. This suggests that engaging in language listening and discussion can positively impact students' comprehension. Importantly, though both students in the comparison and control conditions

demonstrated growth in language abilities and students in the intervention condition made significantly more growth on some language subtests (Vaughn et al., 200b), the rates of progress for both groups were very low, likely affecting their reading and comprehension of text. However, it is important to note that this may be in part due to students having low language and literacy level prior to receiving the intervention in both studies.

Filippini et al. (2012) examined the effects of a supplemental vocabulary focused intervention on decoding outcomes for predominantly Spanish-speaking ELs in Grade 1. Students were administered the Dynamic Indicators of Basic Early Literacy Skills (DIBELS, Good & Kaminski, 2002) Nonsense Word Fluency subtest to determine their at-risk status. Students who scored below 35 sounds per minute were considered higher risk. Though 35 is within the "emerging" range on the official DIBELS benchmarks, Filippini et al. (2012) over selected higher risk students to reduce the possibility of false negatives at-risk identification.

Five classrooms were randomly assigned to one of two conditions: (a) Vocab+, or (b) phonological awareness comparison. Ninety-seven students within the classrooms were then placed into homogeneous groups based on their scores on the Nonsense Word Fluency subtest. The groups of students in the classrooms assigned to the Vocab+ condition were then randomly assigned to one of two Vocab+ conditions: (a) Vocab+ semantic relations focus, (b) Vocab+ morphological focus. Trained undergraduate and graduate students hired by the research team implemented the supplemental interventions to small groups of students in classrooms for 15 minutes a day, 4 days a week, for a total of 8 weeks.

In all three conditions, students received fast-paced explicit and systematic instruction that included clear instructional procedures for multiple levels of scaffolding as well as various opportunities to respond and receive reinforcement and corrective feedback. Students in all three

condition also listened to the same science focused expository text read aloud each week and received instruction in phonological awareness and decoding, including phonics. However, between each condition, the proportion of time spent on developing students' phonological awareness and phonics skills varied. Students in the phonological awareness comparison condition received only phonological awareness and decoding instruction for the entirety of the 15-minute supplemental intervention. Students in the Vocab+ interventions spent only 30% of the time focused on phonological awareness and phonics and spent 70% of the time on vocabulary instruction including direct instruction in word meanings with images. Students in the Vocab+ semantic relations condition focused on the meaning of words and their relationship among other words based on their meaning. Students in the Vocab+ morphology intervention focused on the meaning of morphemes and word formation rules.

ELs at higher risk for reading difficulties demonstrated greater growth on nonsense word fluency than students at low risk for reading difficulties regardless of instructional condition. For higher-risk ELs, the Vocab+ morphological awareness intervention had a greater effect on nonsense word fluency than the Vocab+ semantic relations intervention and phonological awareness intervention; however, these group differences were not found to be statistically significant. Additionally, students who received either of the Vocab+ interventions, including both high- and low-risk ELs, demonstrated higher outcomes on a researcher created target word vocabulary measure than ELs in the phonological awareness comparison condition, though these differences were not found to be statistically significant.

The findings from this study suggest that the supplementary, explicit intervention focused on vocabulary development, phonological awareness, and decoding has the potential to improve reading outcomes for both high- and low-risk EL students. Importantly, students who spent 70%

of the instructional time on vocabulary and 30% of the time on phonological awareness demonstrated gains in measures of decoding just as large as students who spent 100% of the instructional time focusing on phonological awareness exclusively. These findings provide support that teaching vocabulary alongside phonological awareness early and before students demonstrate comprehension difficulties, especially for at-risk students, can help improve both students' vocabulary and decoding skills.

Baker et al. (2016) examined the effects of a researcher developed supplemental reading intervention that emphasized vocabulary and academic language to support the transition from bilingual instruction to English only instruction for Spanish-speaking ELs in Grade 1 at risk for reading difficulties receiving Tier 1 bilingual instruction. To determine if students were at risk for reading difficulties and therefore eligible for the study, Baker et al. (2016) first identified students who scored below benchmark performance in pseudoword reading and oral reading fluency in the middle of Grade 1. The authors chose this time point because they argued that students who were below benchmark in the middle of Grade 1 were more likely to be at risk for reading difficulties than students who scored below benchmark in the beginning of the year, because those students tend to improve their literacy skills quickly after receiving instruction. The authors then rank ordered the students using their Stanford Achievement Test (SAT 10, Harcourt, 2003) Word Reading subtest scores and students with adjacent scores were then randomly assigned to either the treatment condition or to the comparison condition within each school.

Thirty-nine Spanish-speaking ELs at risk for reading difficulties were assigned to the treatment condition while 39 students were assigned to the control condition. Students in the comparison condition received the BAU school reading intervention for students at risk for

reading difficulty. This instruction varied by school and included a variety of commercially available supplemental programs (e.g., Houghton Mifflin core reading curriculum, Fast Track Phonics, DISTAR, and the Harcourt intervention program) and supplemental instruction using leveled reading books. Three certified bilingual teachers and 11 bilingual instructional assistants (IAs), who were already employed at the school, delivered the intervention in small groups for 30 minutes a day, 5 days a week, for a total of 12 weeks. In three schools, the intervention was provided after school, while at the other four schools, the intervention was implemented during small group instruction.

The intervention consisted of 12 units composed of five lessons each. Each lesson was made up of two sections. The first section focused on developing students' decoding skills including phonemic awareness, phonics, word reading, vocabulary, and sentence reading. Additionally, it included instruction in contrastive analysis between English and Spanish letters and sounds. The second section of the lessons focused on developing students' academic language, content vocabulary, and comprehension strategies during stories read aloud. Students listened to two stories a week and engaged in before, during, and after reading activities to bolster comprehension and vocabulary knowledge. Vocabulary instruction focused on targeted vocabulary words relevant to the story, general academic vocabulary, and decodable words. Each scripted lesson utilized a brisk pace and evidence-based instructional methods including explicit modeling of how to use new learning strategies and skills, scaffolding instruction, providing students with multiple opportunities to respond in groups and individually, and providing students with immediate corrective feedback.

Baker et al. (2016) found that ELs in both conditions made significant growth on all reading outcomes including measures of fluency, comprehension, decoding, and vocabulary.

However, similar to Filippini et al. (2012), there were no significant difference between students who received the intervention, who spent significantly more time on phonemic awareness, vocabulary, and comprehension instruction, and students in the comparison condition, who spent significantly more time on phonics, word work, and sentence reading. Baker et al. (2016) noted that a plausible explanation for the lack of significant intervention effects was that teachers in both the intervention and control condition had participated in professional development on explicit instruction, and teachers in the control condition also included an explicit and systematic approach to teaching foundational reading, like teachers in the intervention condition. This suggests that it is possible that there may not have been a large difference in the quality of instruction between the intervention and control conditions. Additionally, the authors suspect that another plausible explanation for the lack of significant effects in the intervention group could be due to a relatively small dosage of intervention (12 weeks) and that their outcome measures may not have been sensitive enough to detect differences in growth of higher order skills such as vocabulary and comprehension.

Overall, the results from this study have important practical implications. First because ELs in both conditions made significant growth on all reading outcomes, this suggests that commercially available interventions for at-risk non-ELs may also be effective, with some adaptations, at improving literacy and language outcomes for at-risk ELs. They also suggest that ELs at-risk for reading difficulties should receive explicit and systematic supplemental instruction tailored to meet their instructional needs even before they have been identified as being at-risk for reading difficulties.

In the last included study examining the effects of reading interventions with instruction in language comprehension for ELs with or at risk for reading difficulties in grades K-1, Neuman

and Kaefer (2018) examined the effects of a supplemental shared book reading program for students in pre-K to Grade 1. The intervention was designed to enhance students' academic vocabulary, conceptual knowledge, and content knowledge through systematic vocabulary instruction using science-related texts. Students in this study were approximately one standard deviation below the norm on the Peabody Picture Vocabulary Test (PPVT-III, Dunn & Dunn, 1997) and therefore determined to be at risk for reading difficulties. Twenty-four classrooms, with a total of 148 students, were assigned to the treatment condition while 21 classrooms, with a total of 117, students were assigned to the control condition. Thirteen percent of the treatment sample and fourteen percent of the control sample were classified as current ELs. Students in the comparison condition received BAU morning meeting instruction.

Students in the treatment condition received an adaptation of the World of Words intervention program during morning meeting. This program utilizes shared reading to promote academic vocabulary and science content knowledge. The curriculum consists of ten topic-related text sets per grade level. Each text set is composed of five read-aloud books of varying genres that begin with predictable books and move on to narrative nonfiction books and expository texts. The intervention included cumulative instruction that utilize students' background knowledge to anchor new content. The lessons included explicit vocabulary instruction along with before reading, during reading, and after reading activities. Before reading, teachers introduced words that represent content-rich vocabulary within taxonomic categories along with a visual representation of each word. During reading, teachers read and re-read the text and stopped at key points to clarify the meaning of words or challenging concepts. After reading, the teacher asked questions that required students to think critically about the concepts learned through the text, apply their new knowledge, and engage in discussions around

the vocabulary and concepts. The World of Words intervention encouraged teachers to systematically scaffold students' learning of new words, encourage students to engage in discussions, and help students build connections between their existing knowledge and new concepts being taught.

The results from this study indicate greater gains for younger students. Students in kindergarten scored significantly higher on science-related vocabulary knowledge and science concepts than students in the comparison. Importantly, ELs who received World of Word showed significantly more growth than non-ELs in the treatment condition on science-related vocabulary knowledge. Additionally, ELs were just as successful as non-ELs at developing knowledge of science-related concepts, with the greatest gains observed in the younger grades. Similar to Vaughn and colleagues (2006a, 2006b), Neuman and Kaefer also suggest that the role of discussion may have been especially important in promoting vocabulary for students. These results suggest that integrating word knowledge and instruction within the context of content-related read-aloud books may be more effective at improving both vocabulary and content knowledge for both non-ELs, and even more so ELs, demonstrating the interconnectedness between word and content knowledge.

Summary of Interventions with Students in K – Grade 1

All five of the included studies that examined the effects of supplemental reading interventions that included instruction in language comprehension included explicit and systematic instruction and demonstrated that these interventions improved language and literacy outcomes for ELs at-risk for reading difficulties in grades K-1, though these effects were not always statistically significant when compared to BAU instruction. Only one study examined effects for students in kindergarten while five studies examined effects for students in Grade 1.

Vocabulary and listening comprehension were the most frequently instructed language comprehension subcomponent skills in the included interventions. All five studies included instruction in these subcomponent skills. Oral language and reading comprehension were included in three studies and morphology, syntax, and semantics were only included in one study each. Unsurprisingly, phonological awareness and decoding were included in most (four) of the studies.

Interventions with Students in Grades 2 – 5

Four studies examined the effects of reading interventions that included instruction in language comprehension on language and literacy outcomes for ELs with or at-risk for reading difficulties in grades 2-5 (Denton et al., 2014; Jacob et al., 2016; Toste et al., 2009; Wanzek & Roberts, 2012). Toste et al. (2019) examined the efficacy of a supplemental multisyllabic word reading intervention on measures of decoding, spelling, and text comprehension for students in grades 4 and 5 who were struggling readers. The authors also examined whether EL status influenced the impact of the intervention. Toste et al. (2012) utilized a two-step screening process to determine students' at-risk status. First, the school district used universal screening data to nominate the lowest-performing students who were not already receiving intensive reading intervention. These students were then administered the Test of Word Reading Efficiency- Second Edition (TOWRE-2, Torgesen et al., 2012). Students who scored less than or equal to the 25th percentile were classified as at-risk and were eligible to participate in the study.

A total of 114 students were randomly assigned to one of three conditions: (a) multisyllabic word reading, (b) multisyllabic word reading plus motivational beliefs training, and (c) comparison condition. This review will focus on the multisyllabic word reading and comparison groups only. Students in the comparison group received BAU instruction. This

instruction typically included small group guided reading lessons, computer-based programming, and independent sustained silent reading.

Eight tutors hired and trained by the research team implemented the multisyllabic word reading intervention to small groups of 3 to 4 students for 40 minutes a day, 4 days a week, for a total of 10 weeks. Each scripted lesson in the multisyllabic word reading contained various instructional components that focused decoding, morphology, encoding, fluency, and text reading. Of particular relevance to this review, the morphology instruction focused on explicitly teaching students high-frequency affixes. Each lesson included instruction in three prefixes that were preselected from a list of the most commonly used prefixes in grades 3 to 9.

The multisyllabic word reading intervention demonstrated statistically significant effects on decoding, as measured by the Woodcock Johnson (WJIII; McGrew et al., 2007)) Word Attack and Word Identification subtests, as well as on one measure of reading comprehension, as measured by the Gates-MacGinitie Reading Test-4 (GMRT- 4; MacGinitie et al., 2000), and spelling as measured by the WJIII Spelling subtest for the overall sample of students. Of interest to this review, Toste et al. (2019) found a significant interaction between EL status and treatment on spelling outcomes. Overall, the effect of the multisyllabic word reading intervention were similar for ELs and non-ELs, except for on spelling outcomes where non-ELs in the multisyllabic word reading intervention made greater gains than ELs.

Wanzek and Roberts (2012) examined the relative effects of varying instructional emphases on reading outcomes for students in Grade 4 with or at risk for reading difficulties. The authors implemented a two-step process to identifying students with impaired reading who were at high risk for being identified as having a reading disability. First, classroom teachers were asked to list students who had either (a) been identified with dyslexia or a learning disability in

reading, (b) is receiving supplemental reading instruction, tutoring, or Title 1 services, or (c) is reading below grade level. Then, the authors screened the identified students using the comprehension subtest of the GMRT- 4. Students who scored at or below the 25th percentile were eligible to participate in the study. A total of 101 students, who were predominantly Spanish-dominant ELs, were identified to participate in the study. ELs made up 52.4% of the word recognition emphasis intervention group, 62.5% of the comprehension emphasis intervention group, and 74% of the comparison condition. Students in the comparison condition received BAU reading instruction. Additionally, 47.8% of the students in the comparison condition received supplemental intervention in addition to BAU.

Six teachers hired by the research team implemented the interventions to small groups of students for 30 minutes a day, 5 days a week, for 28 weeks. Based on a battery of assessment given at the beginning of the school year, students were assigned to one of four conditions through stratified random assignment: (a) word recognition emphasis, (b) comprehension recognition emphasis, (c) responsive emphasis, or (d) comparison. This review will only focus on the word recognition emphasis, comprehension recognition emphasis, and comparison groups. The two intervention conditions differed in the instructional emphasis placed on word recognition or comprehension components. Both intervention groups received the same amount of time in text reading, time spent on vocabulary instruction, and received instruction in all five key elements of reading.

In the word recognition emphasis intervention, students received the Wilson Reading System, which is a systematic program that focuses on phonological awareness and decoding in isolation, and then provides students with opportunities to apply that knowledge to text. Students also had opportunities to practice word reading, spelling, sentence reading and writing, and text

reading daily. In the comprehension emphasis intervention, students received Collaborative Strategic Reading (CSR) and spent most instructional time learning comprehension skills and strategies (e.g., Preview, Click and Chunk; Get the Gist; Wrap Up) by using expository texts.

Wanzek and Roberts (2012) did not find any significant differences between the word recognition emphasis, comprehension recognition emphasis, and comparison groups on measures of word reading, vocabulary, reading comprehension or listening comprehension. Importantly, the authors found that ELs demonstrated accelerated learning and performed better than non-ELs on decoding outcomes in all study conditions. Results from this study suggest that regardless on instructional foci, ELs demonstrated accelerated learning in decoding compared to non-ELs. One possible explanation for why ELs outperformed non-ELs on these measures regardless of the type of intervention they were receiving is that the main impact came not from the instructional focus but rather from receiving small group instruction with ample opportunities to practice reading text.

Interestingly, the authors found that the students in the two intervention conditions had statistically similar reading outcomes to students in the control condition who received the school's typical literacy instruction. Wanzek and Roberts (2012) note that one possible explanation for the lack of significant differences on reading outcomes between groups is that all students in the study were receiving additional literacy supports such as test preparation skills, which included practice in passage comprehension, and in some CSR strategies. Additionally, the study was conducted at schools who have historically implemented science-based reading instruction and have been recognized by the state as being effective at teaching literacy. Therefore, the authors suggest that students in the comparison condition were receiving strong reading instruction, which may explain why there were no significant differences on reading

outcomes between the groups. Wanzek and Roberts (2012) also point out that it is possible that older students with or at-risk for reading difficulties need more intensive interventions than what was provided in this study to see accelerated growth. Specifically, the authors note the short duration of the lessons (30 minutes) and of the study (28 weeks) and suggest that a higher overall dosage may have resulted in different outcomes.

Denton et al. (2004) examined the effects of two supplemental interventions, Read Well and Read Naturally, on decoding and reading comprehension outcomes for Spanish-dominant ELs enrolled in a transitional bilingual education program in grades 2-5. Both interventions combined instructional strategies previously validated for non-ELs, such as explicit phonics instruction and repeated reading of connected text, along with research-based strategies for ELs, such as preteaching vocabulary and engaging in conversations about text. To be eligible to participate in the study, students had to be recommended by their teacher as having reading difficulties. Additionally, students had to be Spanish-speaking ELs who had adequate oral English proficiency to benefit from intervention and at least basic Spanish reading proficiency.

Ninety-nine students participated in this experimental study with 19 students assigned to the Read Well treatment group, 14 students assigned to the Read Well comparison group, 32 students assigned to the Read Naturally treatment group, and 28 students assigned to the Read Naturally comparison group. Because students were enrolled in a bilingual education program, the language of their core reading instruction varied with some students receiving literacy instruction predominantly in English, some students receiving literacy instruction predominantly in Spanish, and some students receiving literacy instruction in both languages. Students in the comparison condition received BAU instruction. This instruction varied by class. Of the small subsample of nine teachers interviewed, most teachers reported that they did not provide neither

Spanish nor English phonics instruction in their reading programs. However, all teachers indicated that their classroom instruction did typically include both vocabulary instruction as well as comprehension instruction consisting primarily of oral questioning prior to reading a book.

Both the Read Well and Read Naturally interventions were implemented in a pull-out setting during school by 23 undergraduate university students for 40-minutes a day, 3 times a week, for 10 weeks. Students in the Read Well intervention received explicit and systematic decoding instruction combined with ample opportunities to practice reading decodable texts and contextualized vocabulary and comprehension instruction. The intervention utilized two types of stories in its instruction: stories that students read on their own and stories that students read with the instructor. The stories and story sections that students read on their own were decodable while the story parts that the instructors read contained more complex and sophisticated language. Each lesson provided students with considerable repetition and practice reading text with immediate teacher feedback and correction. Students in the modified Read Naturally treatment group worked on developing their oral reading fluency through repeated reading of text and received vocabulary and comprehension instruction. Additionally, students set goals for their oral reading fluency rate and monitored their own progress.

Students who were in the Read Well treatment group and received systematic and explicit phonics instruction, made significant growth on word identification but not in word attack or passage comprehension compared to students who received BAU classroom instruction with no phonics instruction. Given the brief duration of the intervention, these results may have significant implications as they suggest that even small amounts of systematic and explicit phonics instruction can strengthen ELs' decoding abilities. The lack improvement in reading

comprehension could be explained by students' insufficient automaticity, fluency, and vocabulary knowledge. Read Well's vocabulary instruction was not systematic and was embedded into decoding instruction and text reading, suggesting that informal vocabulary instruction may not be sufficient for ELs.

Students in the Read Naturally treatment condition who received repeated reading instruction did not demonstrate improved decoding or reading comprehension skills compared to students in the BAU comparison condition. This suggests that ELs at-risk for reading difficulties need explicit instruction in vocabulary, phonics, and phonological awareness to make gains in decoding and comprehending text. Additionally, because the intervention was only 10 weeks long and, similar to authors in the previously described studies, Denton et al. (2004) note that a possible explanation for this lack of significant findings for the Read Naturally group is that the intervention was not intensive enough. The authors suggest that increasing intervention intensity may impact student reading outcomes.

In the last included study examining the effects of reading interventions with instruction in language comprehension for ELs with or at risk for reading difficulties in grades 2-5, Jacob et al. (2016) evaluated the effects of Reading Partners, a program that provides supplemental one-on-one tutoring, on sight-word efficiency, reading fluency, and reading comprehension to students at-risk for reading difficulties in grades 2 to 5. To be deemed eligible to participate, students had to be recommended by a teacher or school staff member and their reading test scores were reviewed. A total of 1,1666 students, 55% of whom were classified as ELs, were randomly assigned within schools and grade levels to either the Reading Partners Program or to the comparison condition which received their school's BAU supplementary reading instruction.

Community volunteers implemented the Reading Partners Program either during school in a pull-out setting or after school. Students received one-on-one instruction for 45 minutes a day, twice a week, for a total of 28 weeks. The Reading Partners Program provided students with a structured an individualized curriculum by assessing students' reading fluency, accuracy, and comprehension, and then placing them at their appropriate instructional level. Each lesson began with the tutor reading a story aloud to students while modeling fluent reading and checking for comprehension by asking students questions regarding the content of the story and vocabulary words encountered. Next, the tutors worked with students on a worksheet that either introduced a new skill or reinforced a previously taught skill. Lastly, students applied their new skill knowledge when reading text aloud while their tutor provided them with instructional support.

Jacob et al. (2016) found that the supplemental Reading Partners program demonstrated positive and statistically significant impact on students' reading comprehension, sight word efficiency, and oral reading fluency. Further, these effects did not differ significantly and demonstrated positive impact on reading outcomes for ELs as well. On average, students' reading scores increase by a month-and-a-half to two months compared to students in the comparison condition. Results from this study suggest that the Reading Partners program has the potential to impact upper elementary students' reading outcomes. These results have important implications because the study also found that the program is far more cost effective for schools to implement than traditional commercially available programs. Additionally, the intervention was implemented by volunteers who were not trained teachers and still produced positive and statistically significant effects on reading comprehension scores for students in grades 4 and 5. These results suggest that it may be possible to implement low-cost intervention programs that do not deplete school resources and still improve student reading outcomes.

Summary of Interventions with Students in Grades 2 – 5

All but one of the included studies (Jacob et al., 2016) that examined the effects of supplemental reading interventions that included instruction in language comprehension for ELs at risk for reading difficulties in grades 2-5 included explicit and systematic instruction. All of the interventions examined demonstrated positive effects on language and literacy outcomes; however, similar to the effects from interventions for students in grades K-1, these effects were not always statistically significant when compared to BAU instruction. Reading comprehension and vocabulary were the most frequently instructed language comprehension subcomponent skills in the included interventions. None of the included studies provided instruction in syntax and semantics, two studies included instruction in morphology, and only one study included instruction in listening comprehension. Interestingly, all of the included studies included instruction in decoding for upper elementary students.

Discussion

The purpose of this review is to summarize the current state of knowledge regarding the use and characteristics of reading interventions that contain instruction in language comprehension for elementary-aged ELs with or at risk for reading difficulties. This population is of particular interest because ELs represent one of the fastest growing student populations in the United States and represent a unique and heterogeneous group of students that are simultaneously developing English language skills while learning in English. Further, national data suggests that the proportion of ELs scoring at or above the NAEP Proficient and NAEP Basic levels in Grade 4 is 10% and 33% respectively (USDOE, 2022). Students who struggle with reading in third grade are likely to continue to struggle as they progress through their schooling and are more likely to drop out of school (Francis et al., 1996, Brasseur-Hock et al.,

2011). Further, reading comprehension is an essential skill needed to access not only the language arts curriculum, but also the general education curriculum across content areas. Importantly, students rarely outgrow reading difficulties and need evidence-based instructional strategies to prevent and remediate reading difficulties (Vaughn et al., 2006a). Therefore, it is critical to identify effective interventions that not only remediate ELs' literacy skills, but also accelerate learning progress so that students can catch up with their peers and access grade-level content.

Despite a growing body of empirical research examining effective early reading interventions, this review located nine studies, five that focused on early elementary students and four that focused on upper elementary students, conducted between January 2000 and November 2023 that met the inclusion criteria. This number of studies is distinctly smaller than the number of high-quality intervention studies examining reading instruction for native English speakers with or at risk for dyslexia (e.g., see Hall et al., 2022) and examining language comprehension interventions for native English speakers (e.g., Silverman et al., 2020). Given this limited sample of studies and variability in the participant, intervention, and design characteristics, there are limited patterns and trends that can be drawn in this review.

Participant Characteristics

First, across the nine studies, there were inconsistent criteria applied to determining students' EL status. In some studies, schools designated students as English learners while in other studies, ELs were identified through parent reports. Further complicating the ability to synthesize results across studies, slightly less than half of the studies did not include information on how EL status was determined. Additionally, most studies did not include any information regarding students' language proficiency and did not distinguish whether the participants were

students who were currently classified as ELs or students who were former ELs and had lost their official EL designation upon achieving grade-level English language proficiency standards. Because language proficiency impacts reading outcomes, it is very important that future research include this information to determine whether interventions have a differential impact on ELs based on varying levels of language proficiency. This will help inform the field on which interventions are most effective for which populations of students.

Similarly, there were inconsistent criteria applied to determining students' at risk status. Six studies included either a norm- or criterion-referenced outcome measure as part of determining student reading risk; four studies included either a teacher or district referral; two studies included a researcher created word-reading measure; and two studies said they used a reading assessment but did not specify which. Without first establishing clear parameters around acceptable evidence of risk, it is challenging to interpret findings across studies because it is possible that students who were found to be at risk using the criteria from one study, would not be considered at risk using the criteria from another study and vice versa. This makes it difficult to interpret for whom the intervention effects are applicable.

Intervention Characteristics

Research has consistently demonstrated that there are various instructional practices that significantly improve literacy outcomes for students with or at risk for reading difficulties. These practices include explicit and systematic instruction in the five key reading concepts (i.e., phonemic awareness, vocabulary, comprehension, phonics, and fluency), small group and one-on-one instruction, and teacher modeling and feedback (Gersten & Baker, 2000; Swanson, 2008; Vaughn et al., 2000). Additionally, research has demonstrated that there are various instructional practices that significantly improve literacy outcomes for ELs including providing intense

explicit vocabulary instruction across several days using a variety of instructional activities, integrating oral and written English language instruction into teaching, and providing small-group instruction. The results of this review provide preliminary evidence to suggest that these approaches are being implemented with ELs with or at risk for reading difficulties and demonstrate improved student outcomes. All the interventions examined in this review except for one (Jacob et al., 2016) were described as explicit and systematic. Most studies described that the implementers modeled how to use new learning strategies and skills, scaffolded instruction, and provided multiple opportunities for students to respond. All but one study (Neuman & Kaefer, 2018) implemented the intervention in small-groups or one-on-one.

All nine of the studies included examined the effects of multicomponent interventions that ranged from providing instruction in three to nine literacy subcomponent skills. Importantly, the studies in this review were unable to determine the specific elements of the interventions that significantly impacted reading outcomes for ELs with or at risk for reading difficulties. Multiple studies suggested that the role of oral discussions may be especially important in promoting language comprehension for students; however, further research should be conducted to disentangle the effects of the instructional components.

These studies greatly varied in instructional dosage and ranged from 8 to 28 weeks of instruction. There was no clear recommendation for intervention dosage, including lesson length and frequency, across studies for ELs with or at risk for reading difficulties. It may be possible that some students require high intensity interventions, such as the intervention examined in the two Vaughn studies (2006a, 2006b), while others may not require such intensive interventions and could benefit from interventions of reduced scope or for shorter periods of time. Future research should examine how intervention intensity impacts language and literacy outcomes for

ELs at risk for reading difficulties, and whether there is a specific dosage that teachers should aspire to adhere to in order to generate the greatest positive impact on student outcomes.

Limitations

The findings from this systematic review should be considered in light of several limitations worth noting. First, the small number of studies that met the inclusion criteria limited our ability to examine whether participant characteristics (e.g., age, grade level, first language, EL language status), intervention features (e.g., dosage, setting, implementer, instructional foci), and methodological characteristics (e.g., study design) may have impacted student outcomes. Additionally, as previously noted, there were inconsistent criteria applied to determining students' EL and at-risk status. Therefore, the results from this review should be interpreted with caution as they may not be generalizable to the broader EL population.

The inclusion criteria for this review required interventions to include instruction in a subcomponent skill that contributes to language comprehension. However, the review did not define a required amount of language comprehension instruction a study must have to be included. Therefore, the time spent on instructing subcomponent skills that contribute to language comprehension varied greatly from study to study. While some studies included language comprehension skills as the main instructional foci (e.g., Neuman & Kaefer, 2018), other studies only had it as a very small component of the larger intervention (e.g., Toste et al., 2008). Though this review intended to examine the average amount of language comprehension instruction included in each intervention, very few studies reported time spent on these component skills. Future research should include this information to be able to examine how the dosage of instruction on skills that contribute to language comprehension instruction impacts student outcomes.

Lastly, in an effort to only include high-quality experimental or quasi experimental research, the search was limited to studies published only in peer-reviewed journals. This likely contributed the very small number of included studies. Including grey literature in the review could have aided in increasing the reviews' comprehensiveness, reducing publication bias by possibly disseminating studies with null findings, and providing a balanced representation of the available research (Paez et al., 2017).

Conclusion

The results from this systematic review provide promising evidence for the use of reading interventions that include instruction in subcomponent skills that contribute to language comprehension for elementary-aged ELs with or at risk for reading difficulties. However, these results should be interpreted with caution due to the small number of included studies and the large variability in participant, intervention, and study design characteristics. Importantly, this review highlights that there is a dearth of research examining interventions with instruction in language comprehension for ELs with or at-risk for reading difficulties. Additional rigorous research should be conducted to evaluate reading interventions serving diverse populations of ELs with or at risk for reading difficulties and should include detailed information regarding students' language proficiency to further examine the possible differential impact of these interventions.

References

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

Al Otaiba, S., McMaster, K., Wanzek, J., & Zaru, M. W. (2023). What we know and need to know about literacy interventions for elementary students with reading difficulties and disabilities, including dyslexia. *Reading Research Quarterly*, *58*(2), 313-332.

August D., Shanahan T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, D. L., Burns, D., Kame'enui, E. J., Smolkowski, K., & Baker, S. K. (2016). Does supplemental instruction support the transition from Spanish to English reading instruction for first-grade English learners at risk of reading difficulties?. *Learning Disability Quarterly*, *39*(4), 226-239.

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in Spanish and English. *Reading and Writing*, *25*, 251-281.

Baker, D. L., Richards-Tutor, C., Sparks, A., & Canges, R. (2018). Review of single subject research examining the effectiveness of interventions for at-risk English learners. *Learning Disabilities Research & Practice*, *33*(2), 64-74.

Boon, R. T., & Barbetta, P. M. (2017). Reading Interventions for Elementary English Language Learners with Learning Disabilities: A Review. *Insights into Learning Disabilities*, *14*(1), 27-52.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144-179.

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual differences*, *21*(4), 438-452.

Cheung, A. C., & Slavin, R. E. (2012). Effective reading programs for Spanish-dominant English language learners (ELLs) in the elementary grades: A synthesis of research. *Review of Educational Research*, *82*(4), 351-395.

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology*, *111*(6), 982-1000.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P. S., Bayraktar, A., Crowe, E. C., & Schatschneider, C. (2011). Testing the impact of child characteristics× instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, *46*(3), 189-221.

Covidence. (2021). Covidence systematic review software. Veritas Health Innovation.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934.

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *The Elementary School Journal*, *104*(4), 289-305.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test--Third Edition (PPVT-III) [Database record]. PsycTESTS.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Every Student Succeeds Act, Pub. L. No. 114-95, Stat. 1177 (2015).

Filippini, A. L., Gerber, M. M., & Leafstedt, J. M. (2012). A vocabulary-added reading intervention for English learners at-risk of reading difficulties. *International Journal of Special Education*, *27*(3), 14-26.

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712-751.

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. *The Elementary School Journal*, *118*(3), 477-504.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational psychology*, *88*(1), 3.

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional children*, *66*(4), 454-470

Gillon, G. T. (2018). Phonological awareness development in children with spoken language impairment. In *G. T. Gillon, Phonological Awareness: From Research to Practice* (2nd ed., pp. 95-117). The Guilford Press.

Goldenberg, C. (2020). Reading wars, reading science, and English learners. *Reading Research Quarterly*, *55*(51), S131-S144.

Good, R. H., & Kaminski, R. A. (2002). DIBELS oral reading fluency passages for first through third grades. (Technical report no. 10). Eugene, OR: University of Oregon

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257-285.

Gough, P. B., & Tunmer, P. B. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Graham, S., Silva, M., & Restrepo, M. A. (2022). Reading intervention research with emergent bilingual students: A meta-analysis. *Reading and Writing*, 1-32.

Grimm, R. P., Solari, E. J., McIntyre, N. S., & Denton, C. A. (2018). Early reading skill profiles in typically developing and at-risk first grade readers to inform targeted early reading instruction. *Journal of School Psychology*, *69*, 111-126.

Harcourt, B. (2003). *Stanford Achievement Test – Tenth Edition: Technical data report*. San Antonio, TX: Author.

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

Jacob, R., Armstrong, C., Bowden, A. B., & Pan, Y. (2016). Leveraging volunteers: An experimental evaluation of a tutoring program for struggling readers. *Journal of Research on Educational Effectiveness*, *9*(sup1), 67-92.

Joshi, R. M., Ji, X. R., Breznitz, Z., Amiel, M., & Yulia, A. (2015). Validation of the simple view of reading in Hebrew—A Semitic language. *Scientific Studies of Reading*, *19*(3), 243-252.

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391-398.

Kieffer, M. J., & Vukovic, R. K. (2012). Components and context: Exploring sources of reading difficulties for language minority learners and native English speakers in urban schools. *Journal of Learning Disabilities*, *45*(5), 433-452.

Larson, A. L., Cycyk, L. M., Carta, J. J., Hammer, C. S., Baralt, M., Uchikoshi, Y., An, Z. G., & Wood, C. (2020). A systematic review of language-focused interventions for young children from culturally and linguistically diverse backgrounds. *Early Childhood Research Quarterly*, *50*, 157-178.

Li, J. T., Tong, F., Irby, B. J., Lara-Alecio, R., & Rivera, H. (2021). The effects of four instructional strategies on English learners' English reading comprehension: A meta-analysis. *Language Teaching Research*.

Liberatti, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Medicine, 6(7), 1–28.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology*, *95*(3), 482-494.

Lonigan, C. J., Shanahan, T., & Cunningham, A. (2008). Impact of shared-reading interventions on young children's early literacy skills. In Developing early literacy: Report of the National Early Literacy Panel (pp. 153–171). Washington, DC: National Institute for Literacy.

Ludwig, C., Guo, K., & Georgiou, G. K. (2019). Are reading interventions for English language learners effective? A meta-analysis. *Journal of Learning Disabilities*, *52*(3), 220-231.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). Gates-MacGinitie Reading Tests, Forms S and T (4th ed.). Itasca, IL: Riverside.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300-335.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, *40*(2), 148-182.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). Technical Manual: Woodcock-Johnson III Normative Update. Rolling Meadows, IL: Riverside.

National Center for Education Statistics. (2022). English Learners in Public Schools. Condition of Education. U.S. Department of Education, Institute of Education Sciences. Retrieved May 27, 2023 from https://nces.ed.gov/programs/coe/indicator/cgf

Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology*, *52*, 15-24.

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233-240.

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children*, *82*(2), 144-169.

Roberts, G. J., Hall, C., Cho, E., Coté, B., Lee, J., Qi, B., & Van Ooyik, J. (2022). The state of current reading intervention research for English learners in grades K–2: A best-evidence synthesis. *Educational Psychology Review*, 1-27.

Rogde, K., Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, *15*(4).

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, *100*(1), 175.

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

Schwartz, Sarah. (2022, July 20). Which States Have Passed 'Science of Reading' Laws? What's in Them? Education Week. Retrieved May 22, 2023 from https://www.edweek.org/teaching-learning/which-states-have-passed-science-of-reading-laws-whats-in-them/2022/07

Silverman, R. D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K–5 students' language and literacy outcomes. *Reading Research Quarterly*, *55*, S207-S233.

Snyder, E., Witmer, S. E., & Schmitt, H. (2017). English language learners and reading instruction: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 136-145.

Solari, E. J., Kehoe, K. F., Cho, E., Hall, C., Vargas, I., Dahl-Leonard, K., Richmond, C. L., Henry, A. R., Cook, L., Hayes, L., & Conner, C. (2022). Effectiveness of interventions for English learners with word reading difficulties: A research synthesis. *Learning Disabilities Research & Practice, 37*(3), 158-174.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72-110.

Swanson, E. A. (2008). Observing reading instruction for students with learning disabilities: A synthesis. *Learning Disability Quarterly*, *31*(3), 115-133.

J. K. Torgesen, R. K. Wagner, and C. A. Rashotte (2012). Test of Word Reading Efficiency–Second Edition (TOWRE-2). Austin, TX: Pro-Ed.

Toste, J. R., Capin, P., Williams, K. J., Cho, E., & Vaughn, S. (2019). Replication of an experimental study investigating the efficacy of a multisyllabic word reading intervention with and without motivational beliefs training for struggling readers. *Journal of Learning Disabilities*, *52*(1), 45-58.

U.S. Department of Education. (2022). National Assessment of Educational Progress (NAEP), 2020 and 2022 Long-Term Trend (LTT) Reading and Mathematics Assessments. National Center for Education Statistics, Institute of Education Sciences.

U.S. Department of Education. (2017). Our nation's English learners: What are their characteristics? Washington, DC: National Center for Education Statistics, Institute of Education Sciences.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015). Early childhood education intervention report: Shared book reading. Retrieved from http://whatworks.ed.gov

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., Pollard-Durodola, S. D., Fletcher, J. M., & Francis, D. J. (2006a). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal*, *43*(3), 449-487

Vaughn, S., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention research: Findings from research syntheses. *Exceptional children*, *67*(1), 99-114

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E. & Francis, D. (2006b). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal*, *107*(2), 153-180.

Wanzek, J., & Roberts, G. (2012). Reading interventions with varying instructional emphases for fourth graders with reading difficulties. *Learning Disability Quarterly*, *35*(2), 90-101.

Table 1

Study Features

Authors	Grade Level	Sample Size	EL Primary Language	EL Determination	Reading Risk Determination
Baker et al. (2016)	1	78*	Spanish	Parent report	Pseudoword reading (below benchmark at midyear) ORF (below benchmark at midyear)
Denton et al. (2014)	2 – 5	93*	Spanish	NR	Teacher referral
Fillippini et al. (2012)	1	71*	Spanish	School determination	DIBELS NWF (<35)
Jacob et al. (2016)	2 – 5	1,166 (641 ELs)	NR	NR	Test scores (unspecified) Teacher referral
Neuman & Kaefer (2018)	PreK–1	265 (35 ELs)	NR	NR	PPVT (>1 standard deviation below the norm)
Toste et al. (2009)	4 – 5	109 (29 ELs)	NR	NR	District referral TOWRE-2 (≤ 25th percentile)
Vaughn et al. (2006a)	1	91*	Spanish	School determination	WLPB LWID (<25th percentile) Researcher created word-reading list (raw score < 2)
Vaugh et al. (2006b)	1	48*	Spanish	School determination	WLPB LWID (< 25th percentile) Researcher created word-reading list (raw score < 2)
Wanzek & Roberts (2012)	4	87 (43 ELs)	Spanish	School determination	Teacher referral GMRT (standard score < 25th percentile)

Note. WLPB = Woodcock Language Battery; GMRT = Gates-MacGinitie Reading Tests; DIBELS = Dynamic Indicators of Basic Early Literacy Skills; LWID = letter word identification; TOWRE-2 = Test of Word Reading Efficiency- Second Edition; PPVT = Peabody Picture Vocabulary Test; LWID = Letter Word Identification; ORF = oral reading fluency; NR = not reported

*Indicates the entire sample was comprised of ELs

Table 2

Grade Level Student Distribution

Grade Level	Approximate Number of Total Participants
K	88
1st	376
2nd	304
3rd	368
4th	467

Note. Neuman & Kaefer (2018) did not report student sample size at each grade level. The total number of participants for the study (N=265) was divided evenly for the included grade levels (K-1).

Table 3

Intervention Characteristics

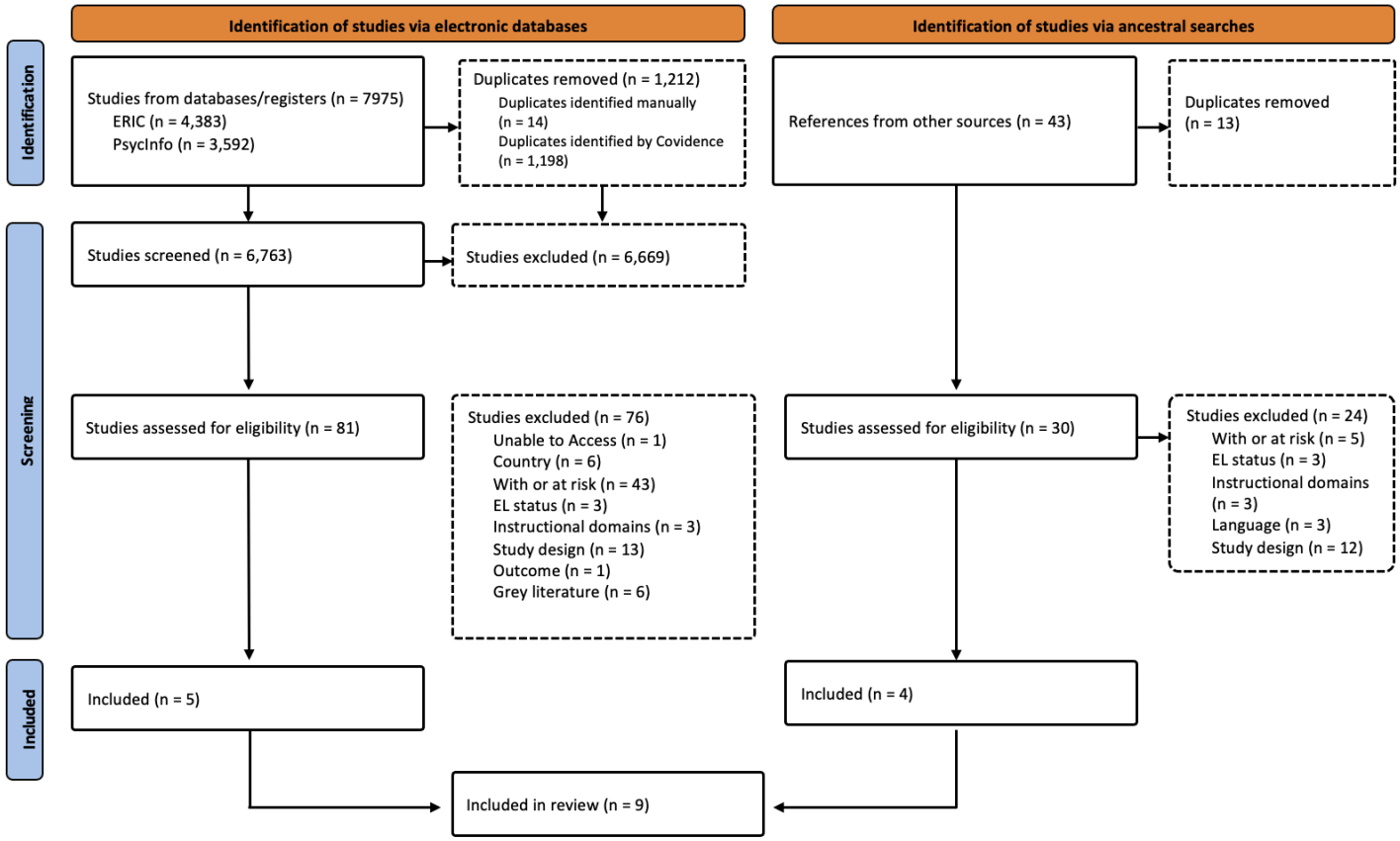
Authors	Group Size	Intervention Duration	Interventionist	Intervention Program	Literacy Domains Instructed
Baker et al. (2016)	Small groups	12 weeks (50 min lessons x 5 days a week)	Certified teachers; IAs (bilingual)	Transitions Lessons (researcher created)	PA, D, E, TR, V, RC, LC
Denton et al. (2014)	1-4	10 weeks (40 min lessons x 3 days a week	University undergraduates	Read Well Read Naturally	Read Well: PA, D, TR, V, RC, OL Read Naturally: D, TR, V, RC, OL, RR
Fillippini et al. (2012)	Small groups	8 weeks (15 min lessons x 4 days a week)	University undergraduate and graduate students	Core Intervention Model + researcher modifications	Vocab+ morphology: V, M, LC, PA, D, RR Vocab+ semantic relations: V, S, LC, PA, D, RR
Jacob et al. (2016)	1	28 weeks (45 min lessons x 2 days a week)	Community volunteers	Reading Partners	D, TR, V, RC, LC
Neuman & Kaefer. (2018)	Whole class	20 weeks (NR)	Certified teachers	World of Words	V, LC, OL
Toste et al. (2009)	3-4	10 weeks (40 min lessons x 4 days a week)	Tutors hired by the research team	Multisyllabic Word Reading (researcher created)	D, M, E, TR

Vaughn et al. (2006a)	3-5	28 weeks (50 min lessons x 5 days a week)	Certified teachers hired by research team (bilingual)	Proactive Reading	PA, TR, RR, D, E, OL, RC, LC, V
Vaugh et al. (2006b)	3-5	28 weeks (40 min lessons x 5 days a week)	Certified teachers hired by research team (bilingual)	Proactive Reading	PA, TR, RR, D, E, OL, RC, LC, V
Wanzek & Roberts (2012)	2-4	28 weeks (30 min lessons x 5 days a week)	Certified teachers hired by the research team	Wilson Reading System Collaborative Strategic Reading	Wilson Reading system: V, RC, PA, D, E, TR Collaborative Strategic Reading: V, M, RC, D, TR

Note. PA = phonological awareness; D = decoding; RR = rote reading of whole words; E = encoding; LK = letter knowledge; TR = connected-text reading fluency; V = vocabulary; S = syntax; M = morphology; LC = listening comprehension; RC = reading comprehension; OL = oral language; IA= instructional assistant

Figure 1

PRISMA Diagram



Manuscript 3

Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners

Isabel Vargas and Emily J. Solari

Author Note

Isabel Vargas, M.T., Department of Curriculum, Instruction, and Special Education, University of Virginia School of Education and Human Development.

Emily J. Solari, Ph.D., Professor, Reading Education Program, Department of Curriculum, Instruction, and Special Education, Virginia Literacy Partnerships / Director, University of Virginia School of Education and Human Development.

Abstract

This study evaluates the acceptability, appropriateness, and feasibility of a seven-week listening comprehension and vocabulary intervention for elementary-aged ELs and non-ELs. The intervention, Building Vocabulary and Early Reading Strategies (BVERS; Solari & Ciancio, 2014), is considered supplemental to the school's typical English language arts (ELA) curriculum and was not used as a substitute or replacement for core instruction. A total of 10 teachers, 8 general K-2 education teachers, 1 English as a Second Language (ESOL) teacher, and 1 reading interventionist, participated in the study and implemented the BVERS for 30 minutes a day, four days a week, for 7 weeks to a total of 193 students, including ELs and non-ELs, at 2 schools. Following the intervention, teachers answered questions from a 16-item measure of acceptability, appropriateness, and feasibility and participated in a 45-minute semi-structured focus group interview to discuss their experience implementing the intervention. Fidelity data was also collected through observations and coded using a fidelity checklist to measure intervention adherence. A convergent mixed methods design merging both quantitative and qualitative data from teachers and students was used to comprehensively examine the intervention's implementation process. Findings from this study suggest that BVERS is acceptable, appropriate, and feasible when implemented by teachers in elementary school practice settings. These findings indicate promise for the successful implementation of BVERS in diverse classrooms that may ultimately impact students' language and literacy outcomes.

Keywords: reading interventions, English learners, language comprehension, feasibility, implementation research

Examining the Feasibility of a Listening Comprehension and Vocabulary Intervention for Elementary-Aged English Learners and Non-English Learners

The foundation for children's academic success is their ability to read proficiently and comprehend what they read; reading comprehension is essential for students to engage across all academic subjects. Reading comprehension is the ability to derive meaning from text, and it is a skill that develops over time. Successful reading comprehension requires proficiency in two broad skill categories: code-based skills that contribute to word recognition (e.g., PA, alphabet and phonics knowledge, decoding) and language-based skills that contribute to linguistic comprehension (e.g., vocabulary knowledge, knowledge of syntax/language structure; Gough & Tunmer, 1986; Hoover & Gough, 1990). Students need to simultaneously decode isolated words quickly and accurately while also deriving meaning from the words they read to successfully comprehend text.

National data suggests that a large proportion of students in the United States (U.S.) are unable to adequately read and comprehend grade level text. The most recent iteration of the National Assessment of Educational Progress (NAEP) demonstrated that 33% of fourth grade students performed at or above the NAEP Proficient level on the reading assessment in 2022, and 37% of fourth grade students performed below the NAEP Basic level in the same year (U.S. Department of Education [USDOE],, 2023). These scores reveal a decline in overall student performance when compared to the 2019 scores. Importantly, on average, reading scores were lower across many student subgroups when compared to the national sample. For example, in 2022, the proportion of students scoring at or above the NAEP Proficient and NAEP Basic levels is lower, 10% and 33% respectively, for students currently classified as English learners (ELs), a quickly growing and highly diverse group of students that make up 10.3% of enrollment in U.S.

schools (USDOE, 2017). This is consistent with extant research that suggests that ELs perform worse on reading outcomes as compared to their non-English learner (non-EL) peers.

Reading difficulties can result in serious academic and vocational consequences for students as they get older. Longitudinal data has demonstrated that students with reading difficulties in the primary grades are likely to have persistent reading difficulties throughout their school years (Adolf, et al., 2010; Francis et al., 1996; McNamara et al., 2011; Torgesen & Burgess, 1998). Additionally, these students are also more likely to drop out of school and experience unemployment (Aro et al., 2019; Daniel et. Al., 2006; Korhonen, et al., 2014). Fortunately, research suggests that effective early reading intervention can help reduce the severity, and potentially the incidence, of reading difficulties among young students (Al Otaiba et al., 2009; Wanzek et al., 2018). Over the last few decades, scientific advances have informed our understanding of reading development and acquisition. Research has demonstrated that literacy instruction should include an explicit and systematic instructional approach in five key reading components: (a) phonological awareness, (b) phonics, (c) fluency, (d) vocabulary, and (e) comprehension (National Early Literacy Panel [NELP], 2008).

Despite a large and growing body of empirical research conducted around establishing how proficient reading develops and identifying effective instructional strategies to teach literacy skills, a profound gap exists between these empirical findings and the implementation of evidence-based instructional literacy practices in school settings (Solari et al., 2020b). Though there are several reasons that help explain why the evidence base is not being implemented in schools, one key factor is that reading research oftentimes occurs in tightly controlled settings and does not account for unique contextual factors that may impact an intervention's implementation in education. Intervention research oftentimes focuses on generating evidence of

an intervention's efficacy, primarily by demonstrating improvement on proximal reading outcomes such as student performance on a decoding measure, and often lacks a focus on examining the implementation process of the intervention and how it affects outcomes in real-classroom settings. For example, there is large diversity in school contexts, such as school leaders' and teachers' awareness and perceptions of science-based research, availability of resources, and diversity of the student populations schools serve. However, limited empirical research has focused on how these contexts impact the effective implementation of instructional literacy strategies (Solari et al., 2020b). Teachers often seek strategies that are not only effective, but also feasible to implement in their own classroom and contextualized for their students' needs (Solari et al., 2020b). The aim of the proposed study is to report findings of the acceptability, appropriateness, and feasibility of a research-based early reading intervention that targets listening comprehension and vocabulary through shared story reading implemented by teachers to improve reading and language outcomes for elementary students including both ELs and non-ELs.

Reading Comprehension Development

The simple view of reading (SVR) provides an empirically validated framework for understanding the component skills necessary for successful reading comprehension (Gough & Tunmer, 1986). The SVR states that both code-based skills and language-based skills contribute separately to students' reading development, and proficiency in both is necessary for successful reading comprehension. The SVR describes code-based skills, which are also referred to as code-focused and code-related skills, as efficient word recognition or "the ability to rapidly derive a representation from printed input that allows access to the appropriate entry in the mental lexicon, and thus, the retrieval of semantic information at the word level" (Hoover & Gough,

1990, p. 130). In other words, it is the ability to read words quickly and accurately. The SVR describes language-based skills, which are also referred to as meaning-based skills, language comprehension, listening comprehension, and linguistic comprehension, as "the process by which, given lexical (i.e., word) information, sentences and discourses are interpreted" (Gough & Tunmer, 1986, p. 7). It is the ability to derive meaning from both written and oral language. A student with the ability to decode isolated words quickly and accurately will not comprehend text if they are unable to derive meaning from the words they read (Gough & Tunmer, 1986). Similarly, a student who has a strong understanding of language and language structures will not be able to successfully comprehend a text if they are unable to decode the words in the text (Gough & Tunmer, 1986). Therefore, it is essential that students develop both code-based and language-based skills early in the primary grades to support reading comprehension.

The SVR has been empirically validated for both non-ELs (e.g., Gough & Tunmer, 1986; Lonigan et al., 2018) and ELs (e.g., Grimm et al., 2018; Hoover & Gough, 1990; Joshi et al., 2014). In their seminal study, Hoover and Gough (1990) provided initial support for the SVR by tracking the reading development of a sample of English and Spanish bilingual students through the early elementary grades. Their findings illustrate that code-based and language-based skills explained significant and unique variance in students' reading comprehension. Importantly however, is that although both sets of skills are necessary for successful reading comprehension, the relative contribution of these constructs change across time for both non-ELs and ELs.

Relative Contribution of Language Comprehension to Reading Comprehension

The relative contribution of language comprehension to reading comprehension increases over time as texts become more complex and students become more proficient decoders and focus more on gaining meaning from text (Adolf et al., 2010; Catt et al., 2005; García & Cain,

2014; Language and Reading Research Consortium, 2015). Though there may be some variability in young students' language comprehension skills, researchers have suggested that this variability is normally distributed around a substantial mean that is higher than what is required to comprehend typical texts encountered in the younger grades if the text were read aloud (Gough et al., 1996). Based on the SVR, individual differences in word reading skills are the primary contributor to reading comprehension during these early years of formal reading instruction for beginning readers. Inversely, in the later grades, as students become more automatic and efficient decoders, texts also become more complex and the vocabulary, syntax, and discourse demands of the texts increase. As older students become more proficient decoders, language comprehension skills contribute more than word reading skills to reading comprehension (Gough et al., 1996).

Empirical research has repeatedly supported these findings (Adolf et al., 2010; Catts et al., 2005; Language and Reading Research Consortium, 2015; Vellutino et al., 2007). In one seminal study, Gough et al. (1996) meta-analyzed 10 studies to examine the relationships between word reading, language comprehension, and reading comprehension over time for non-ELs. The authors reported 17 correlations between word reading and reading comprehension that decreased with the increasing age of participants from Grade 1 to college students.

García and Cain (2014) meta-analyzed 110 peer reviewed studies to examine the relationship between decoding skills and reading comprehension during reading development and to identify reader characteristics that contribute to this relationship. The authors found that the strength of word reading skills and reading comprehension relationship decreased with age. Specifically, they found that around 10 years of age, there was a reduction in the strength of the correlation between decoding and reading comprehension. As predicted by the SVR, as children

became more proficient decoders, the variance in reading comprehension explained by word reading decreased and the role of language comprehension increases. García and Cain (2014) also found that listening comprehension had a moderating effect on the relationship between decoding and reading comprehension. Higher listening comprehension scores were found to be associated with weaker correlations between decoding and reading comprehension, consistent with the SVR which posits that reading comprehension is the product of word recognition and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990).

Catts et al. (as cited in Language and Reading Research Consortium, 2015, p.152) used a series of regression analyses to examine the influence of word reading and listening comprehension on reading comprehension using a longitudinal sample of students in grades 2, 4, and 8. The authors found that the unique contribution of word reading to reading comprehension decreased over time as students got older from 27% in second grade, 13% in fourth grade, and 2% in eighth grade. Catts and colleagues found an opposite pattern for the impact of listening comprehension on reading comprehension. Listening comprehension contributed 9% in second grade, 21% in fourth grade, and 36% in eighth grade.

Most recently, the Language and Reading Research Consortium (2015) examined the influence of word recognition and listening comprehension on reading comprehension across time with participants in grades 1-3. Consistent with Catts et al. (2010), the authors found that the influence of decoding skills decreased with increasing age and the influence of listening comprehension increased. Extending Catts and colleagues' findings, the Language and Reading Research Consortium found that the increase on unique variance explained by listening comprehension was apparent by Grade 2, and vocabulary skills affected reading comprehension indirectly through both decoding and listening comprehension.

Though linguistic comprehension has been found to explain more variance in reading comprehension for all students as they progress beyond the primary grades, reading comprehension difficulties in ELs are more strongly predicted by language comprehension difficulties than word reading difficulties, and the relationship between language comprehension and reading comprehension difficulties is stronger for ELs than non ELs (Babayiğit, 2014; Cho et al., 2019; Proctor et al., 2005). Cho and her colleagues (2019) used the SVR to compare the sources of reading comprehension failure for ELs and non-EL students with reading difficulties. Specifically, they examined the roles of linguistic comprehension and word reading in explaining poor reading comprehension for each group of students. Using a sample of students in fourth grade, the authors found that although fourth-grade word reading was a significant contributor to reading comprehension difficulties regardless of language status, ELs outperformed non-ELs on measures of word-reading. Additionally, the authors found that non-ELs had a stronger performance on linguistic comprehension measures than ELs and that linguistic comprehension, skills such as listening comprehension and vocabulary, were on average more responsible for reading comprehension difficulties for ELs than they were for non-ELs.

This research suggests that language comprehension instruction should be robust in the elementary grades to equip students with the language skills required for successful reading comprehension of increasingly complex texts. This is particularly important for linguistically diverse learners with lower levels of English proficiency as they progress through school and language comprehension increasingly contributes towards reading comprehension. While research has demonstrated that there is heterogeneity in ELs' language-based skills (e.g., Ford et al., 2013; Gonzalez et al., 2016; Lesaux & Kieffer, 2010; Lonigan et al., 2018; Vargas et al., 2023), all ELs, including former ELs, are developing their English language skills while

simultaneously developing foundational literacy skills (Goldenberg, 2020). Therefore, explicit language and vocabulary instruction is essential and should be emphasized for all students, especially ELs, as language-based skills have been found to be strong predictors of later reading achievement (Lindsey et al., 2013).

Language Comprehension Instruction

Language comprehension relies on various subcomponent skills: breadth and depth of vocabulary knowledge; understanding of language structures including syntax and semantics, verbal reasoning; understanding of morphology; and literacy knowledge (Scarborough, 2001). While language comprehension can be developed in part through exposure to language in the environment, research has demonstrated that explicit instruction can support, and is often required, to develop requisite language skills (Connor et al., 2011; Silverman et al., 2022).

Several meta-analyses and literature syntheses have recently examined the effects of language comprehension interventions for elementary-aged students (Elleman et al., 2009; Fitton et al., 2018; Goodwin & Ahn, 2013; Larson et al., 2018; Marulis & Neuman, 2010; Rodge et al., 2019; Silverman et al., 2020; Stahl & Fairbanks, 1986). These studied found that on average, language comprehension interventions have positive effects on custom measures aligned with the intervention and smaller, yet still positive effects, on measures of general language knowledge and reading comprehension.

Silverman et al. (2020) meta-analyzed studies examining the effects of language comprehension interventions across measures of language comprehension (syntax, vocabulary/semantics, or morphology), listening comprehension, and reading comprehension measures. Silverman and colleagues found positive and significant effects on custom measures of vocabulary, listening comprehension, and reading comprehension, but did not find that these

results generalized to standardized measures of these outcomes. The authors found positive vocabulary and reading comprehension effects for ELs, and that the interventions were equally or more effective for ELs compared with non-ELs.

Most recently, Herrera et al. (2021) examined the past 20 years of research on the effectiveness of early literacy instruction with the goal of providing an update since the NELP report. The review explored instruction in six language and literacy domains: language, phonological awareness, print knowledge, decoding, early writing, and general literacy. The authors found that in the 132 high-quality experimental or quasi-experimental studies evaluated, language (i.e., the ability to comprehend or use spoken language, which can include vocabulary, listening comprehension, syntax, or narrative understanding and production) was the most frequently taught instructional domain. Of the 86 interventions that taught language and also included language outcomes, 67 interventions included shared book reading, 78 interventions focused on vocabulary instruction, and 71 included instruction focused on listening comprehension in addition to vocabulary instruction. The authors found that interventions that included instruction in the language domain resulted in statistically significant effects on language outcomes. These effects were larger when researchers used custom developed language outcome measures rather than general measures of language comprehension. Additionally, they found that interventions that included both vocabulary and listening comprehension instruction significantly improved language performance.

Vocabulary Interventions

Language comprehension encompasses vocabulary knowledge. In their seminal meta-analysis, Stahl and Fairbanks (1986) reviewed vocabulary intervention studies for students in Grade 2 through college. The authors found a statistically significant mean effect size of .97 ($p<$

.01) for vocabulary instruction on reading comprehension outcomes containing taught words and a smaller but significant effect of .30 for general measures of reading comprehension. Further, they found that providing definitional and contextual information of a word (i.e., students are provided with a definition and exposed to the word in context) was the most effective vocabulary teaching method.

Elleman et al. (2019) meta-analyzed 37 vocabulary interventions in Pre-K to Grade 12. The authors found that vocabulary instruction was associated with an increase in students' ability to comprehend text with an overall moderate effect size on custom measures of vocabulary. However, vocabulary instruction was found to be less effective on measures of general vocabulary knowledge. Further, the authors found that interventions with higher levels of discussion were associated with greater effects on researcher-created measures of vocabulary. Additionally, Elleman and colleagues found that vocabulary interventions had greater effects for students with reading difficulties than for students without reading difficulties.

Similarly, Marulis and Neuman (2010) meta-analyzed the effects of vocabulary interventions on the receptive and expressive language of Pre-K and kindergarten aged children. The 64 studies produced a moderate to large overall effect size of .88 ($p < .001$*)* on word learning in pre-K and kindergarten. In contrast with Elleman et al. (2019) who examined effects for students in PreK to Grade 12, the authors reported a larger overall effect size and found that there was no significant difference between gains on vocabulary measures for students in Pre-K and kindergarten at-risk and not at risk for learning difficulties.

Shared Book Reading Interventions

One language-based intervention approach commonly used with young children is shared book reading, an instructional approach in which an adult reads a book aloud to a child or a

group of children and use planned interactive activities to actively engage the children in text through language (What Works Clearinghouse [WWCH], 2015) . Shared book reading interventions have been associated with improved oral language and comprehension outcomes for elementary students (e.g., Coyne et al., 2004; NELP, 2008; Dowdall et al., 2009; Fitton et al., 2018). The National Early Literacy Panel (2008) found that book-sharing interventions produced statistically significant and moderate-sized effects on children's print knowledge and oral language skills.

Coyne et al. (2014) developed a rich repeated shared storybook intervention meant to increase students' linguistic and vocabulary knowledge and increase their story comprehension. This intervention incorporated research-based principles including rich dialogic discussions (i.e. activating prior knowledge, eliciting responses about story elements, and making connections), multiple readings of storybooks, explicit instruction on target words important to the understanding of the story, and rich language instruction such as extended discourse before and after reading. The authors found that kindergarten students at risk for reading difficulties who received the shared story intervention scored significantly higher than students who received business as usual literacy instruction on experimenter-developed measures of explicitly taught vocabulary. Students with lower receptive vocabulary demonstrated greater growth on target vocabulary words in relation to students who did not receive the storybook intervention. Coyne et al. provide evidence that teaching word meanings explicitly within the context of a storybook can increase students' knowledge of target vocabulary.

Shared book reading interventions have also been associated with improved outcomes for ELs. Fitton et al. (2018) examined the impact of shared book reading on language and literacy outcomes among ELs 12-years old or younger. After meta-analyzing 54 studies conducted in the

U.S., the authors found that the overall combined effect size for the impact of shared book reading on ELs' language and literacy outcomes was positive and moderate, 0.28 ($p < .001$). These results were robust to design characteristics and were only moderated by children's special education status: children with developmental disorders exhibited smaller amounts of growth from the shared reading interventions.

Most relevant to the current study, a series of studies have examined the effects of *Building Vocabulary and Early Reading Strategies* (BVERS) (Solari & Ciancio, 2014), a supplemental listening comprehension and vocabulary intervention that utilized shared book reading and explicit vocabulary instruction (Henry & Solari, 2020; Solari et al., 2020a; Henry et al., 2022). Solari et al. (2020a) examined the effects of BVERS, which includes explicit instruction, modeling, and guided practice with listening comprehension strategies (e.g., direct recall, making connections to background knowledge, narrative retell, and making inferences) for elementary-aged students with autism spectrum disorder (ASD). The intervention was implemented in small groups by trained graduate and undergraduate students. The authors found that students who received BVERS had slight improvements on proximal measures of vocabulary and listening comprehension, and significant increases on proximal measures of narrative retell.

Henry and Solari (2020) used an adapted version of BVERS (Solari & Ciancio, 2014) to examine the effects of integrating oral language and listening comprehension into shared book reading for elementary students with ASD. The supplemental intervention was implemented in small groups by special education classroom teachers, and consistent with Solari et al. (2020a), included instruction in vocabulary and written expression. The authors found that the students who received the intervention demonstrated significant improvements on measures of expressive

vocabulary, narrative ability, and listening comprehension. Henry et al. (2022) extended these finding by examining the feasibility and initial efficacy of BVERS implemented virtually for elementary-aged students with ASD. The authors found that students demonstrated growth in listening comprehension, but unlike Henry and Solari (2020), students did not show growth in narrative retell or vocabulary. Based on student attendance, a parent satisfaction survey, and treatment integrity data, Henry et al. (2022) found that BVERS was feasible to implement virtually.

Feasibility Research

The last few decades have produced a large body of intervention research that has informed our understanding of reading development and instruction. This research has focused on generating evidence of the effectiveness of interventions by demonstrating improved literacy and language outcomes for students under controlled experimental conditions. Though ideally interventions supported by rigorous research will be implemented into school practice and will generate similar improved learning outcomes, translating these intervention effects in routine classroom-based settings has oftentimes posed a challenge (Gadke et al., 2021). One explanation for this is that while intervention research has predominantly focused on treatment outcomes, it has rarely focused on the implementation process to examine the intervention procedures and how the intervention works in practice settings (Gadke et al., 2021).

Implementation outcomes, separate from treatment outcomes, are defined as, “the effects of deliberate and purposive actions to implement new treatments, practices, and services” (Proctor et al., 2011, p. 65). Implementation outcomes allow researchers to examine the facilitators and barriers to implementing new interventions in diverse settings and allows researchers to disentangle the effects of the intervention as compared to the effects of

implementation. For example, if an intervention is found to be ineffective, it could be a result of the intervention being ineffective in that particular setting, or it can be the result of an effective intervention being implemented incorrectly (Proctor et al., 2011).

Implementation outcomes, such as acceptability, appropriateness, and feasibility, evaluate the fit between an intervention and some criterion (e.g., personal; technical; social; practical; Weiner et al., 2017). An intervention's acceptability is defined as the "perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory" (Proctor et al., 2011, p. 67). Acceptability measures how well the intervention meets the personal need of various interested parties such as educators who implement the intervention, students who receive the intervention, and the broader school community (Weiner et al., 2017). It examines participants' perception of the fairness, reasonableness, and intrusiveness of the intervention (Gadke et al., 2021). An intervention's appropriateness is the "perceived fit, relevance, or compatibility of the innovation or evidence-based practice for a given practice setting, provider, or consumer; and/or perceived fit of the innovation to address a particular issue or problem" (Proctor et al., 2011, p. 68). Appropriateness measures how well the intervention fits socially and technically and examines how well the intervention addresses a specific concern (Weiner et al., 2017). Feasibility is defined as the "extent to which a new treatment, or an innovation can be successfully used or carried out within a given agency or setting" (Proctor et al., 2011, p. 68). Feasibility examines whether the intervention is practical and whether teachers are able to deliver the intervention relatively easily and consistently given their existing resources and circumstances (Weiner et al., 2017).

Gadke et al. (2021) present a framework for examining implementation outcomes by analyzing various dimensions of the intervention including social validity, practicality,

integration, adaptability, implementation, and effectiveness (see Figure 1). Social validity refers to the acceptance and approval of an intervention, and should be collected from multiple interested parties such as teachers implementing the intervention as well as students receiving the intervention. Practicality examines whether the intervention can be implemented in a certain setting given contextual and environmental factors unique to that practice setting. Such factors include participant commitment and the availability of resources such as time and materials (Gadke et al., 2021). Integration looks at the alignment between the intervention and features of the practice setting such as organizational structures, physical environment, and existing curricula (Gadke et al., 2021). Integration examines the fit between the intervention, school culture, and the school and teachers' teaching philosophy.

Adaptability explores whether an intervention can be adjusted to fit the needs of different settings such as with different student populations and in different formats such as whole class or small group (Gadke et al., 2021). Adaptability examines whether the intervention will produce comparable outcomes irrespective of its delivery format or the student population. Implementation, also commonly referred to as fidelity, examines the degree to which the intervention was implemented as it was intended by the program developers. By examining implementation, researchers focus on how closely the implementation adhered to the program protocol, the dosage of the intervention, and the quality of the intervention delivered (Weiner et al., 2017). Researchers are able to identify the varying levels of treatment integrity as well as the environment and conditions under which the implementation was most successful (Gadke et al., 2021). Lastly, effectiveness, though not the main focus of feasibility research, considers whether the intervention shows some preliminary evidence of positive outcomes for the intended population (Gadke et al., 2021). Examining the effectiveness of the intervention also allows

researchers to explore whether the outcome measures used are sensitive to change and whether the data collection protocol is feasible.

Current Study

Building upon extant research on language comprehension instruction, the proposed study will evaluate the acceptability, appropriateness, and feasibility of BVERS for ELs and non-ELs. Though previous research has established BVERS's efficacy and feasibility for students with ASD when implemented by special education teachers in small group settings (Henry & Solari, 2020; Solari et al., 2020) and when implemented by trained undergraduate research assistants in one-on-one virtual settings (Henry et al., 2022), this study seeks to extend the literature by examining its implementation in a whole class setting by general education teachers and in small groups by reading and ESOL interventionists. In doing so, this study seeks to help address the persistent research to practice gap by examining how, why, and under what conditions BVERS can be implemented effectively in routine classroom-based settings (Solari et al., 2020b). Utilizing Gadke et al.'s (2021) feasibility research framework, the current study will use implementation outcomes (Proctor et al., 2011) to monitor and evaluate teachers' implementation efforts and will address the following research questions:

1. Is BVERS, when delivered in whole class settings by general education classroom teachers and in small groups by reading and ESOL interventionists, acceptable for use with students in grades K-2?
2. Is BVERS, when delivered in whole class settings by general education classroom teachers and in small groups by reading and ESOL interventionists, appropriate for use with students in grades K-2?

3. Is BVERS, when delivered in whole class settings by general education classroom teachers and in small groups by reading and ESOL interventionists, feasible for use with students in grades K-2?
4. Are teachers able to implement BVERS to fidelity?

Method

Participants and Setting

This research was conducted in compliance with the University Institutional Review Board. Written informed teacher and parental consent were obtained prior to data collection. Participants included teachers and students across two schools, and within each school, across three grades, namely kindergarten through second grade.

Schools. The study was conducted at two public elementary schools (K-5) in central Virginia (see Table 1). School A qualified for Title I services and was located in a small suburban district. The school served a total of 446 students. Seventy-four percent of students qualified for free and reduced-price lunch and ten percent of students were ELs. The school had recently adopted the University of Florida Literacy Institute (UFLI) Foundations as part of their literacy curriculum, and the participating teacher from School A had been implementing the program daily for four months prior to the beginning of the current study.

School B was located in a small rural district. The school served a total of 506 students. Sixty percent of students qualified for free or reduced priced lunch and fourteen percent of students were ELs. The school had recently adopted the Orton-Gillingham approach for their literacy instruction, and all participating general education teachers from School B had been utilizing this approach for four months prior to the beginning of the current study.

Teachers. Teachers who taught English language arts to students in kindergarten through Grade 2 were eligible to participate. A total of 10 teachers, 8 general K-2 education teachers, 1 English as a Second Language (ESOL) teacher, and 1 reading specialist, participated in the study and implemented the BVERS curriculum (see Table 2). One of the teachers taught at School A, while the remaining nine teachers taught at School B. The eight general education teachers implemented BVERS in a whole-group setting while the reading and ESOL interventionists implemented BVERS in a small-group setting. Two of the general education teachers taught kindergarten, three taught first grade, and three taught second grade. Six of the teachers had a master's degree and four teachers had a bachelor's degree. All teacher participants possessed either a Collegiate Professional License or a Postgraduate Professional license and had an average of 10.1 years of teaching experience. Seven teachers had a PreK-6 endorsement, two teachers had a PreK-3 endorsement, three teachers had an additional special education (SPED) endorsement, and one teacher had an ESOL endorsement. One teacher was a long-term substitute but had a preK-6 teaching license and a special education endorsement.

Students. One hundred and ninety-three students participated in the intervention. Twenty student attended School A and one hundred seventy-three students attended School B. Forty-nine students were currently classified as an EL and spoke a variety of native languages. In Virginia public schools, multilingual students are identified for language screening based on registration documents or a home language survey parents or guardians fill out during the school enrollment process. During school registration, parents or guardians must respond to the following three questions: (a) What is the primary language used in the home, regardless of the language spoken by the student? (b) What is the language most often spoken by the student? and (c) What is the language that the student first acquired? If the parent or guardian responds with a language other

than English for one or more questions, the student will enter the language screening process. If the student does not have a current WIDA ACCESS for ELLS (2023) score (administered within the past year), the student are administered a WIDA screening assessment (i.e., WIDA Screener, WIDA Model, Kindergarten MODEL, WIDA Screener for Kindergarten) to determine EL status and eligibility for language services.

Measures

Acceptability, Appropriateness, and Feasibility

Following the intervention, teachers answered questions from a 16-item measure of acceptability, appropriateness, and feasibility (adapted from Weiner et al., 2017; see Appendix A). The survey asked teachers to rate their level of agreement for each statement using a 5-point Likert-scale that ranged from 1 (“strongly disagree”) to 5 (“strongly agree"). Four items on the survey focused on the acceptability of BVERS, four items focused on the appropriateness of the intervention for non-ELs, four items focused on the appropriateness of the intervention for ELs, and four items focused on the feasibility of implementation. Additionally, teachers participated in a 45-minute semi-structured focus group interview. The interview protocol was adapted from Komesidou and Hogan (2023) and included 14 questions regarding the acceptability, appropriateness, and feasibility of BVERS (see Appendix B). The interviews were conducted virtually in grade-level groups ranging from one to three teachers that all taught at the same school. Each interview was audio recorded and transcribed.

Acceptability. The current study aimed to measure how well BVERS fit each teacher’s needs, preferences, and expectations. To assess BVERS’s acceptability, the current study examined one dimension of the intervention, social validity (Gadke et al., 2021). To assess the social validity of BVERS, students who had received parental consent to participate in the data

collection process completed a five-item social validity survey (see Appendix C). The survey was administered and read aloud by each student's teacher, and it asked students to rate their level of agreement with various statements using a 3-point Likert-scale measure where a sad face represented "disagree", a neutral face represented "neither agree nor disagree", and a happy face represented "agree". Additionally, this study collected social validity data from teachers through survey items and the semi-structured interview where teachers were asked to describe their experience implementing BVERS and the components of the intervention they liked and disliked.

Appropriateness. To examine BVERS's appropriateness, this study explored how well BVERS matched the school's and teacher's professional values and norms, and how much teachers perceived the intervention to meet their students' needs. To assess BVERS's appropriateness, the current study examined three dimensions of the intervention: (1) integration, (2) adaptability, and (3) effectiveness (Gadke et al., 2021). To explore BVERS's integration within existing systems, during the focus group interview, teachers were asked to describe their school and district climate related to accepting and implementing new interventions. Additionally, teachers were asked how well BVERS fit with their teaching philosophy, values, and norms, and how the intervention did or did not fit within the existing infrastructure of their school and classroom. To measure the adaptability of BVERS, teachers were asked to describe the types of alterations, if any, the intervention would need to be effective in their classroom setting. Lastly, to examine the effectiveness of BVERS, teachers were asked to describe anecdotally whether they perceived the intervention to meet the needs of the students, both ELs and non-ELs, in their class. Additionally, they were asked to describe what changes, if any, they

had observed in their students' listening comprehension, vocabulary knowledge and use, and comprehension skills since the beginning of implementation.

Feasibility. The current study examined the feasibility of teachers implementing BVERS in both whole class and small group settings within their existing resources and circumstances (e.g., time and effort). To assess BVERS's feasibility, the current study examined one dimension of the intervention, practicality (Gadke et al., 2021). To assess practicality, teachers were asked to describe the barriers and facilitators they faced during implementation. Additionally, they were asked whether they were able to implement BVERS with their available resources, time, training, and materials.

Implementation Fidelity

To examine the extent to which teachers enacted BVERS as intended and as designed by the intervention developers, the current study examined one dimension of the intervention, implementation, also referred to as fidelity (Gadke et al., 2021). To measure fidelity, the lead author conducted three classroom observations per teacher during a BVERS lesson over the course of seven weeks. Each observation was recorded and coded for fidelity of implementation by the lead author. To ensure that each intervention session was implemented consistently, a fidelity checklist was used during each observation (see Appendix D). The fidelity checklist included separate sections for the three key instructional elements of the intervention: (1) read aloud, (2) vocabulary, and (3) writing. Additionally, each lesson was coded for the overall quality of student engagement, teacher feedback and scaffolding, and organization of materials. Fidelity was measured by examining how closely each teacher's lesson implementation matched with the script and instructional activities described in the intervention lesson plans. Teachers were scored using a binary scale where a score of one represented that the instructional component was

present and implemented as intended and a score of zero represented that the component was missing or implemented differently than as intended. The overall fidelity score was a percentage representing the number of intervention components implemented as designed out of the total number of possible components. Fifty percent of the recorded lessons were randomly selected and independently double-coded for fidelity by a second observer, a doctoral student in a reading education program. This observer established interobserver agreement (>90%) with the lead author using the fidelity checklist prior to conducting the independent observations.

Study Procedures

Instructor Training and Coaching

The lead author trained teachers during a one-day, four-hour professional development session. The session consisted of on an overview of the study design with an emphasis on the importance of examining the feasibility of implementation, the theoretical underpinnings of BVERS, the importance of linguistic comprehension for reading comprehension, and characteristics of evidence-based literacy instruction. The training also included a detailed outline of the intervention's routines along with video exemplars of experienced instructors implementing the intervention. The teachers engaged in facilitated discussions regarding the logistics of implementation and were given various opportunities to explore the intervention materials and practice implementing a lesson while receiving feedback from the lead author. In addition to the initial training, teachers received in-person and virtual coaching. Each teacher was observed implementing a BVERS lesson three times and received immediate feedback from the lead author to support high-fidelity implementation.

Intervention Implementation

The BVERS intervention was considered supplemental to the school's typical English language arts (ELA) curriculum and was not used as a substitute or replacement for core instruction. All students continued to receive the school's ELA curriculum throughout the course of the study. General education teachers implemented BVERS during their regularly scheduled ELA time or during their intervention block in a whole class setting for 30 minutes a day, four days per week, over seven weeks. The reading and ESOL interventionists implemented BVERS using a "pull-out" model of instruction for small groups of students for 30 minutes a day, four days per week, over seven weeks.

Instructional Components

BVERS is a scripted listening comprehension and vocabulary intervention developed by Solari and Ciancio (2014). The intervention targets listening comprehension strategies and vocabulary through explicit instruction and the use of preselected children's storybooks to provide opportunities for instruction in one of the following four comprehension skills: (1) direct recall of information from texts; (2) making connections to background knowledge; (3) narrative retell; and (4) making inferences. The sequencing of these comprehension skills is systematic and progresses from fundamental to more complex skills. These skills are also cumulative, and mastery of each previous skill is required before progressing to the next.

Each comprehension skill unit consists of two to four weeks of instruction. Each lesson follows the same format with four components: (1) introduction or review of the target comprehension strategy; (2) read aloud of the whole text or part of the text (15 minutes); (3) vocabulary instruction (5 minutes); and (4) a writing activity to support comprehension (10 minutes). Each comprehension skill unit begins with an introductory anchor lesson with explicit instruction in the new comprehension skill. Then, five concurrent lessons focus on the same

storybook to provide students with various opportunities to engage with guided practice with each comprehension skill using repeated readings of familiar text.

Read Aloud

Each read aloud begins with setting the purpose for reading by introducing a guiding question that focuses on a main idea from the book. Next, the teacher reads the story aloud and stops at predetermined points throughout the book to ask students scripted questions and to model think-alouds of discussion and reasoning strategies related to the target comprehension skill. The questions include both closed- and open-ended question types and offer students multiple opportunities to respond with different answers related to the target comprehension skill. During the first and last day of the five-day sequence focused on each book, students listen to the entirety of the text read aloud. On the second, third, and fourth days, students listen to portions of the text and engage in additional guided discussions and practice with the target comprehension skill. Each read aloud concludes by revisiting the guiding question. If students are unable to independently answer the question, the teacher uses a prompting hierarchy progressing from moderate to more intense scaffolding to support oral responses when needed.

Additionally, visual supports are used during the read aloud portion of the intervention. This includes pictures and hand signals that accompany each comprehension skill. These visual supports help participants to demonstrate their understanding and express their responses in nonverbal ways.

Story Vocabulary

BVERS includes explicit instruction in story vocabulary words. During the first four lessons for each story, teachers provide explicit instruction in two to three novel vocabulary words. These words have been preselected and are words that are either essential for successful

comprehension of the story or high utility Tier 2 words (Beck et al., 2008). During the initial encounter with the vocabulary word during the read aloud, the teacher defines the word with a short and child-friendly definition. After the read aloud concludes, the teacher redefines the word and students engage in an activity to support retention. These activities are interactive and include acting out the word, finding examples and non-examples using pictures, and generating synonyms and antonyms. They provide students with multiple exposures with each word and encourage them to practice using the words in new contexts. During the fifth and final lesson for each book, the teacher reviews the eight to twelve previously taught words.

Written Expression

The final component of each lesson is a writing or drawing activity created to facilitate further discussion of the story and reinforce the target a comprehension skill addressed during the week. These activities included students generating ideas, summarizing important story details, drawing pictures about the story, and writing short responses about the text. This writing component is used to support text comprehension development and reinforce the target comprehension skill rather than to teach novel writing skills.

Analytic Approach

To explore the acceptability, appropriateness, and feasibility of BVERS, the current study used a convergent mixed methods design merging both quantitative and qualitative data from teachers and students to comprehensively discuss the intervention's implementation process. The study utilized parallel questions with the same sample of teachers during the qualitative focus group interviews and quantitative survey questionnaire to examine the same dimensions of implementation to allow for the comparison, and possibly merging, of the various data sources (see Table 3). Quantitative teacher and student survey data were analyzed by describing the

mean scores and the variability of the scores for each implementation outcome (i.e., acceptability, appropriateness, and feasibility). Fidelity of implementation was examined using the percentage of intervention components implemented as designed by the developers. Lastly, focus group interviews were recorded and transcribed. We then employed a deductive analytical approach to examine implementation outcomes by applying Gadke et al.'s (2021) dimension of feasibility research framework to examine teacher participants' responses based on six dimensions of implementation (i.e., social validity, integration, adaptability, effectiveness, practicality, and implementation).

Results

A total of 10 teachers (100%) completed the acceptability, appropriateness, and feasibility survey (see Table 4) and 8 teacher (80%) participated in semi-structured focus group interviews. Connie, a second-grade long-term substitute, finished her contractual term the week after implementation concluded and did not respond to requests for an interview. Ginny, a first-grade teacher, withdrew from the study at the conclusion of implementation due to changes in her work responsibilities and time constraints. Though she did not participate in the interview, she did complete the survey and consented to having her data to remain part of the study.

Acceptability

Social Validity

Teachers. On average, teachers rated the four items on acceptability highly (M= 4.33, SD= .79). Seventy percent of the teachers either agreed or strongly agreed with the statement, "This intervention meets my approval," and only one teacher, Melissa, a kindergarten teacher at School A, disagreed with the statement. Melissa commented, "We really enjoyed the first day of the read aloud of the whole book, but I just wish all the questions were tied into the first read

aloud…because they [the students] liked the books! But for us, BVERS was just too long. It was too much in depth… I think some things were dragged out too much…Kindergarten is very much an exposure grade, and we expose [them] to a lot of things. So, we don't spend like a week looking at a [writing] web… my kids were getting very bored. (Interview, 2/29/2024)" Conversely, Victoria, a kindergarten teacher at School B, gave BVERS the highest possible acceptability score. She stated, "To these kids, it probably felt too repetitive because they're simply not used to reading a book more than once. But I think it's beneficial to them to read books and listen to books several times because you're going to remember more things each time you do…so, I think it's beneficial even if it's not their favorite thing. (Interview, 2/29/2024)" Unlike Melissa, she also approved of the structure of the writing component. Victoria said, "I like that the writing follows a bit of a pattern. They [the students] know, 'Okay we're going to make a list. And then we know we're going to use the list to do our writing.' And so, I think just having the procedures in place helps them with their confidence. (Interview, 2/29/2024)"

All teachers either agreed or strongly agreed that the BVERS was appealing, and 90% of teachers said that they liked the intervention. First-grade teachers expressed that BVERS challenged their traditional way of conducting a read aloud, something that they welcomed (see Table 5). Francine, a first-grade teacher, said, "I like how you discuss the books [in the lesson], and I like the questions that were given. They might not have been questions that I would have come up with on my own. So, it was good and just a different way for them [the students] to think about the books sometimes. (Interview, 2/29/2024)" Maureen, another first-grade teacher added, "Personally, I don't really ask questions during the book. I just, I've never really done like active read alouds. Like, this is only my second year of teaching. So BVERS was fun, like to think of a read aloud in that way. (Interview, 2/29/2024)" The reading and ESOL interventionists

rated BVERS highly on acceptability. Faith, a reading interventionist noted, “I enjoyed it! If someone handed it again to me with everything in it [the books], I’d be like, ‘Ahh, yes!’ (Interview, 3/1/2024)”. Clara, an ESOL interventionist added, “You can tell it was teacher created because it’s teacher friendly and shows that they knew what goes on in the classroom. I loved it! (Interview, 3/1/2024)”

Students. Eighty-two students received consent to participate in the data collection process. On average, the students rated the five social validity items highly (see Table 6). Seventy-one percent of students agreed that they liked the books teachers read aloud to them in the BVERS lessons, while only 5% of students disagreed and the remaining 24% of students neither agreed nor disagreed. Additionally, 46% of students agreed that BVERS helped them learn new words while 13% of students disagreed. Most students, 66% and 68% respectively, also agreed that BVERS helped them learn how to answer questions using information they heard in the story and how to make connections between the story and their background knowledge, two comprehension strategies explicitly taught in the intervention. Seventy-seven percent of students felt that they were able to understand the books they were reading during the intervention, and 53% of students wanted to continue using BVERS in their classroom at the end of the seven-week study.

Appropriateness

On average, teachers rated the four items on appropriateness, though slightly lower than acceptability, highly (M= 4.09, SD= 1.15). Teachers found the intervention to be slightly more appropriate for their non-EL students (M= 4.1, SD= 1.22) than for their EL students (M= 4.08, SD= 1.1). Ninety percent of teachers found BVERS to be suitable and applicable to their non-EL students while 70% thought it to be suitable and applicable for their EL students.

Integration

At School A, Victoria described that for the past few years, her district and school have been open to trying out new literacy curriculum. **"**I would say that we are eager to adopt curriculum. Like, we're currently looking at new curriculum for next year. The only issue is we tend to over jump [switch curriculums] sometimes. (Interview, 2/29/2024)" Victoria then went on to explain how over the past two years, the school had implemented three new instructional approaches and interventions that were only focused on word-reading skills (e.g., phonological awareness, alphabet and phonics knowledge, decoding). While she had received a weeklong training for one approach, she received no support for the others. "It was just a manual in our hand, and they said, 'Here you go. Figure it out.' (Interview, 2/29/2024)" She described how even though the school is using an intervention program, they've been told to implement it as core instruction rather than as an intervention. She notes, "The kids aren't getting it. So, we're like eager to jump [curriculums]. But we're not jumping in the right direction. (Interview, 2/29/2024)"

During the current study, Victoria's literacy blocks were an hour and forty minutes long and consisted of a combination of whole group instruction on phonological awareness and small group instruction focused on phonological awareness, decoding, and writing. She stated that read alouds were part of her writing instruction, but that they were not structured and interactive like BVERS. When prompted on whether she believed BVERS would fit into the school's values and existing infrastructure, she expressed concern that the intervention went too in depth for kindergarteners and the lessons took too long. However, she did note that BVERS seemed like a "great program" and that she felt like the manual was "easy enough to follow." She followed up by saying, "I think it'd be welcomed [for older students] if the literacy blocks were set up

differently… But we don't get much of a say in our schedule. (Interview, 2/29/2024)" Victoria noted that because the current curriculum they were using focused specifically on phonics and phonological awareness, she believed BVERS would be a good complimentary intervention.

At School B, teachers also described their school and district as eager to try out new interventions. Sam, a second-grade teacher, said, "I mean, for as long as I've been here, it kind of feels like we've been intervention hopping a little bit. It just seems like we try to like get a feel for different ones and then maybe they don't work out. (Interview, 2/27/2024)" This sentiment was shared across each group of teachers that was interviewed. Like School A, School B had also recently adopted a new instructional approach to literacy and were in their first year of implementation. When asked about whether she thought BVERS would fit into the existing literacy practices at her school, Clara, an ESOL interventionist noted, "I think it really complements the language comprehension and vocabulary piece…I think for a lot of people, there was no organized system for doing that. I think everyone was doing their own thing for that. I was doing it before when I was in the [general education] classroom and yeah, you have books that you would read that were just kind of like fluff, you know. (Interview, 3/1/2024)"

Adaptability

Teachers were asked whether they believed that BVERS could be modified to fit the needs of their students. Though Melissa appreciated the script and in-text prompts as reminders of what types of questions she should be asking, she noted that BVERS felt a bit rigid at times and she found herself wanting to ask her own questions and provide additional scaffolding for her students. She said, **"**We are a very low-income school. And I mean, most my kids did not attend preschool, and so, like, they don't have that background knowledge that some of the kids in like a more privileged school may have…and sometimes the questions weren't worded for a

kid who's never had preschool, like they don't know what background knowledge or any of that stuff is unless they've heard it from me, which is again, is in passing. (Interview, 2/29/2024)" She mentioned that this could be due specifically because of the makeup of her class and their age, "Maybe it's a kindergarten problem, but it's really hard to stick to a script because a lot of my kids get off task, and they like can't always get back on. So, it's harder to stay on script in kindergarten, just because I kind of have to like restate the question, and I want to word it in a way as to not to give them the answer. (Interview, 2/29/2024)" Victoria, who also taught kindergarten, demonstrated an alternative perspective. "It's [BVERS] scripted enough that it's easy to follow, but I do feel like I'm still able to use my degree. (Interview, 2/29/2024)" She continued by saying, "I know that some people don't like things that are scripted because they feel like they're just a robot. And then they, you know, think what does that say about the purpose of teachers and education and the philosophy of all of that? But I do feel like I still have some autonomy [to make instructional decisions]."

Teachers were asked what they felt needed to be altered in the intervention to make it more effective for their students. Most teachers responded that they would not make any alterations to the read aloud and vocabulary instructional components. However, multiple teachers voiced that the writing component of the intervention was consistently very difficult for their students due to time constraints and because they were unable to offer them individualized support during whole-class instruction. This would oftentimes cause teachers to either shorten the writing instruction or skip it altogether. Mackenzie noted, "For the writing piece I feel like, they definitely improved, which was cool to see, but it was hard to be able to get around to everybody and make sure that they were all doing it the right way. The third grade [teachers] always says that they [the students] need to be able to write more, and I feel like a lot of my kids

are now because of doing this. They can write like a full paragraph which is really cool, but some of them I feel like I didn't get to like make it as effective as it could have been because I couldn't help them all. (Interview, 2/27/2024)" Clara offered a possible solution, "I think to keep it in a 30 min chunk of time, I would have to extend it [the lesson] over longer than 5 days. It would have to be like, eight, so you could have like three days of just really focused on the writing and then really focus on the editing and sharing piece because I think that's really important. (Interview, 3/1/2024)"

Effectiveness

Though the current study does not report students' performance on proximal and distal language and literacy measures, teachers provided anecdotal evidence during their interview to support BVER's effectiveness for both ELs and non-ELs. When discussing non-ELs, multiple teachers noted that students began making connections between the stories and their lived experiences. Melissa noted, "They [the students] were doing a better job with their listening skills and with their book interaction skills. They're growing with their ability to listen to a story and realize, 'Oh I can learn from this story,' or, 'Oh I can connect to this story in a meaningful way, or this story can help me understand stuff about myself or about the world.' (Interview, 2/29/2024)" Faith, also noted, "I did see a difference. And as they've gone through [the lessons], they understand the process more and they get it…They started saying things like 'Oh look, that's like Owen and his blanket!' … They were making connections and figuring it out! (Interview, 3/1/2024)"

When discussing BVER's effect for ELs, multiple teachers specifically pointed out that they believed the vocabulary logs and instruction were making a large impact. Clara stated, "They [ELs] started making connections and remembering the vocabulary words, so that was

really neat to see. They definitely remembered them and could apply them too… the small group allowed ELs the space to talk through it more. (Interview, 3/1/2024)" Mackenzie, a second-grade teacher said, "I thought it was really positive for them [ELs]. I loved that the vocabulary cards had the pictures on them and that the activities went along with it. I feel like that helped them be able to access the material. They really seem to enjoy the read alouds, even if some of them maybe didn't understand exactly what was happening. (Interview, 2/27/2024)" The first-grade team at School B felt similarly. Maureen stated, "I think it was nice for my EL population to have those picture cards with the sentence stems. I think it helps them a lot to give them a sentence stem. And then they create whatever back half of the sentence they need to. So, I think we definitely saw some language improvement there. (Interview, 2/29/2024)" However, even with the visual supports, Melissa did not believe BVERS was effective for ELs in her class. "Honestly, for my students who have basically no English, it kind of went over their head. And they would mimic whatever you're saying… I would say the pictures definitely help. We do a lot of pictures in kindergarten. So, they really like the pictures. But I don't know if they would be able to apply it later. (Interview, 2/29/2024)"

Multiple teachers mentioned, that even though the seven weeks of implementation were over, they felt that BVERS left a lasting impact not only on their students, but also on their instructional approach to teaching literacy. In the second grade interview, Sam shared, "The students are sad it's over. But today, I was like, 'Let's make a list of books that we can read instead of BVERS!' and like, I want to start doing some more interactive read alouds that are focusing on similar skills but they [the students] get to come up with some books that they want to read. So, yeah. We're planning building upon it, but not doing obviously the actual intervention. (Interview, 2/27/2024)" Victoria mentioned, "So, we are reading new books, and

like, I caught myself stopping a few times more while I was reading it, and like having [students] think, pair, share more times. (Interview, 2/29/2024)" She went on to explain how in the past, she wouldn't ask students question throughout the read aloud, but rather at the end. Now, she makes sure to check for understanding during reading and encourages students to interact and discuss the text more than before.

Feasibility

On average, teachers rated the four items on feasibility highly (M= 4.18, SD= 0.78). Ninety percent of teachers found the intervention implementable and possible, while 70% of teachers found the intervention easy to use.

Practicality

Across all grade levels, teachers mentioned that they really valued that the materials for the intervention were all provided for them and that helped make BVERS very practical and easy to implement. They mentioned that teachers were often handed curricula and expected to create the physical materials such printing pictures, copying student materials, and assembling kits. However, with BVERS, teachers received folders that included all the materials, including the student materials, they needed for the week. Each folder contained each week's book with the before and after reading cards located on the inside covers, vocabulary logs, picture cards, and the week's lesson plans. However, there was disagreement amongst the teachers regarding their preferences for how the materials were organized. While some teachers enjoyed that each week's folder contained all the materials, other teachers would have strongly preferred if the lesson plans would have included the script for the entirety of the lesson (including the read aloud prompts, vocabulary lesson, and writing lesson) all in a binder instead. Melissa mentioned, ""For kindergarten, it was hard. There was a lot, but maybe it's just me. There was just a lot of pieces

to pull from. And the kids get really distracted in between activities ... It [BVERS] had like multiple pieces to it and like multiple things to look at. And there's a lot to process for me… it was just a lot of transitioning that we are not used to. (Interview, 2/29/2024)" However, she continued by saying, "But like once I got into routine, I was okay."

Getting comfortable with the lesson format was initially challenging for most teachers, however, their confidence and comfort with using the intervention increased throughout the seven weeks. Victoria said, "I think the first little while, it took me a bit to really figure out, even though I practiced it. It took me a while to figure out what order to do things in and where to look to see what comes next and definitely part of that was simply the hectic nature of my group…but once I got the hang of that, then okay, I already know we're doing vocabulary next, then we're doing writing. (Interview, 2/29/2024)" Clara mentioned, "At the beginning, we were all in panic like, 'Oh my gosh, are we doing this right?' There are just so many components and I think we got a little messed up at first…But then once we figured it out, I think we were good. The color coding is wonderful. (Interview, 3/1/2024)" Faith then added, "I even wrote it all out for like the first two weeks! But, yeah, once you figured it out, it was easy. (Interview, 3/1/2024)" Teachers also mentioned that one of the reasons they felt it was difficult to implement BVERS initially was because, due to logistical and weather delays, the teacher training had occurred two months prior to beginning implementation. Teachers weren't sure if it would have made a difference, but they felt that had the training been closer to implementation, the first couple of weeks of BVERS may have gone smoother.

The most frequently mentioned barrier to implementation was time. Teachers often felt that though a single BVERS lesson was meant to take 30 minutes, this was not realistic. Various teachers mentioned that finding 30 minutes for implementation was especially challenging when

teachers had a lot of students who were being pulled out of their class for additional intervention services. Maureen mentioned, “I had to like basically restructure my whole literacy block to fit BVERS in, and that's only because I have a tremendous EL population in my classroom as well as tier services. And so, I have kids being pulled out very frequently during my literacy block during all day. So, it was actually hard to find, like 30 min of uninterrupted time throughout the day and that was the hardest part for me. It was just finding uninterrupted time with all my students in the classroom. (Interview, 2/27/2024)”

Of all the components to implement, writing was the most challenging in part because it was the most time consuming and teachers believed that students were not used to engaging in the writing process. Melissa said, “Honestly, they hated the writing, and our [typical] writing is very like different. It's not laid out in maps or anything like that. We usually use the piece of paper, I'm sure you know, with the picture and the lines. It's very simplistic and only for one day. And like once they're done, they're done…We didn't really do a lot of the writing [in BVERS] because we’d run out of time. And they like weren't opposed to it. (Interview, 2/29/2024)” Maureen had a similar experience and stated, “I think that in the writing process, you have essentially like the brainstorm, and then the web, and then the writing. I think that's a really good strategy. But I feel like writing as a whole is just a struggle. And a lot of my students are not able to do that without specific like help. So, it's super hard to do that whole group, and so sometimes we just didn't even do it. (Interview, 2/29/2024)” Sam tried to adapt the writing component to make it more feasible for her class to complete. She said, “Just finding the time to do the writing [was hard]. I would often just put it into one of my rotations because I always try to have a writing rotation in my literacy block. But by that time, it didn't feel like linear enough, if that makes sense. They would kind of forget what the book was about or forget what they were

supposed to be doing between the time I did the reading and the time that they were supposed to write about it. (Interview, 2/27/2024)"

Fidelity of implementation

Implementation. Teacher's adherence to the activities, as measured by the fidelity checklist, was moderate. Coding agreeability between the lead author and second observer was 93.27%. On average, teachers were able to adhere to 78.2% of the intended intervention components (SD= 15.20; see Table 11). Adherence increased from the beginning of implementation to midway through the intervention, and then decreased towards the end of implementation. On average, teachers adhered to 73.22% (*SD* = 13.22*,* range 55 – 91) of the instructional protocol at the beginning of implementation, to 84.60% (*SD* = 13.92*,* range 64 – 100) at the middle of implementation, and to 77.10% at the end of the seven weeks of intervention (*SD* = 17.77*,* range 42 – 100).

Fidelity was also examined for each separate key element in a BVERS lesson: (1) read aloud, (2) vocabulary, and (3) writing. Teachers, on average, adhered to 77.5% (*SD*= 41.83%) of the instructional components during the read aloud portion of the lesson. Teachers displayed the strongest fidelity to the before reading activities (*M* = 92%, *SD* = 27%) which included reviewing the comprehension skill along with its corresponding hand signal, introducing or reviewing the story, and presenting the lesson's guiding question. They showed moderate adherence (*M*= 80.95%, *SD*= 40%) to the during reading activities which included reading in-text cues and questions, scaffolding students' answers, and encouraging students to use the hand signal associated with each comprehension skill. Teachers showed poor adherence to, and often skipped, the after reading activities (*M*= 54.76%, *SD*= 50.07%) which included revisiting and

answering the lesson's guiding question, scaffolding students' answers, and reinforcing the comprehension skill.

Teachers demonstrated poor adherence to the vocabulary component of the lessons (*M*=63.82%, *SD*= 48.17%). During the vocabulary component of the lesson, teachers were asked to define each target word, provide a teacher model of the word, and encourage students to use the word in a sentence through an activity. However, teachers oftentimes skipped words and either skipped or shortened the vocabulary activities for each word.

The writing component was only implemented during 46.67% of the observed lessons. Teachers frequently ended their lessons by telling the students that they would complete their student journal activity the following day due to limited instructional time. However, when implemented, teachers displayed strong intervention adherence during the written component of the BVERS lessons (*M*= 86.11%, *SD*=35.07%), which included providing students with clear instructions, modeling the writing activity, and prompting students to contribute at least one idea in their writing and the discussion about writing.

Lastly, each observation was coded for the overall quality of student engagement, teacher feedback and scaffolding, and the organization of materials. Teachers demonstrated strong adherence to these protocols (*M*= 93.33%, *SD*= 25.05%).

Discussion

The purpose of this study was to examine the acceptability, appropriateness, and feasibility of BVERS, a listening comprehension and vocabulary intervention, when implemented by general classroom teachers in a whole class setting and when implemented by reading and ESOL interventionists in small groups. Additionally, this study examined whether teachers were able to implement BVERS to fidelity as its developers intended. Despite the large and growing body of

empirical research surrounding how proficient reading develops and investigating effective instructional strategies to teach literacy skills, there is still a large and persistent gap between empirical findings and the implementation of evidence-based instructional literacy practices in schools. An important contributing factor to this challenge is that there are very few studies that have focused on examining the implementation process of different interventions and how unique contextual factors can impact the intervention's effect in schools. This study aims to help address this research to practice gap by examining the barriers and facilitators of the implementation of a research-based intervention to better understand how to translate the effects found in rigorous research into effects in routine classroom-based settings.

Acceptability

Acceptability explores the degree to which an intervention meets the needs of teachers, students, and the school. Acceptability can vary greatly between various interested parties because it is a function of each person's unique needs, preferences, attitudes, and willingness to try and adapt to new things (Weiner et al., 2017). Further, acceptability can differ greatly based on student population, school setting, and school culture (Gadke et al., 2021). While an intervention may be deemed acceptable for one group of students, it may also be deemed unacceptable for another. An intervention's acceptability plays a crucial role in determining whether the intervention is likely to be implemented in practice settings because teachers are more likely to implement and adhere to interventions that meet their approval (Gadke et al., 2021).

Role of Teaching Philosophy

As expected, there was some variability in the perceived acceptability of the intervention between teachers. The teacher who rated BVERS lowest on acceptability was a kindergarten teacher who worked at School A. She noted that she did not approve of the intervention specifically

for kindergarten students and voiced her concern regarding the length, repetitiveness, and complexity of the lessons for her students. This teacher was the same participant who described kindergarten as an "exposure grade" (Interview, 2/29/2024). She described how she believed that the purpose of kindergarten was to introduce students to concepts and to develop their general understanding, rather than to their deep understanding of any one concept by spending extended periods of time repeatedly working on the same skill.

In stark contrast, the kindergarten teacher at School B perceived the intervention to be highly acceptable for her kindergarten students and felt as though the structured repetitive intervention was what students needed to continue developing their listening comprehension skills. She acknowledged that young students may not be used to, or even like, the repetitive nature of some instructional strategies, such as engaging in the writing process for various days, but that she believed it to be effective at helping build their literacy skills and helping them gain a deeper understanding of how to be stronger readers and writers. This contrast between two kindergarten teachers helps highlight the potential role that personal preferences and teaching philosophies play on teachers' perceived acceptability of an intervention.

Role of Familiarity

A common theme teachers noted when discussing the acceptability of BVERS was that BVERS took an instructional approach they were already using (i.e., read alouds), and adapted it to be more explicit and systematic. Further, teachers cited that BVERS was especially appealing because it utilized high quality authentic children's literature, most of which teachers were already familiar with and enjoyed. This suggests that familiarity with the general materials and intervention protocols may aid in its perceived acceptability. Interventions that do not introduce

entirely new concepts and activities, but rather enhance existing structures, may be more likely to be accepted by teachers into their classrooms.

Similar to the teachers, students reported very high levels of approval for the books used in BVERS and reported feeling like they were able to understand the books being read. Anecdotally, many students shared that they had heard many of the stories before, again suggesting that familiarity with the materials and content may impact student measures of acceptability.

Despite there being some slight variation in the degree of acceptability, overall, the results of this study support the acceptability of the intervention by both teachers as well as students. This finding is consistent with Henry and Solari (2020) who reported that special education teachers rated BVERS as being acceptable when implemented in small groups with students with ASD. The current study expanded these findings as BVERS was implemented in both small group and whole class settings, and the participants in the current study included general classroom teachers, reading and ESOL interventionists, and diverse populations of students including ELs. Further, there was notable interest among both teachers and students in continuing to use BVERS beyond the seven-week implementation period of the current study, suggesting the potential for long-term integration into classroom practices within diverse settings.

Appropriateness

The evaluation of an intervention's appropriateness focuses on whether teachers perceive the intervention to be a good fit for their students (Weiner et al., 2017). It investigates various dimensions of implementation including whether the intervention is seen as consistent with the norms, values, and instructional philosophy of the school and whether the intervention can be easily integrated into existing school structures. Further, it examines if the intervention is seen as efficacious at addressing the problem for the population for which it was intended and whether it

has the flexibility to be adapted to better suit the needs of diverse settings and populations of students.

Overall, the results of this study support the appropriateness of the intervention for both ELs and non-ELs. While teachers generally found BVERS to be appropriate and were able to provide anecdotal preliminary evidence of effectiveness, there were some perceived challenges related to the integration, adaptability, and varying effectiveness for different student populations. Addressing these challenges could enhance BVERS's overall efficacy and sustainability within diverse classroom settings.

Integration With Existing Curricula

When examining the dimension of intervention integration, teachers at both schools expressed that their districts were willing, and at times overly eager, to explore new literacy curricula and interventions. While they appreciated having access to new materials and ideas, teachers expressed frustration in the constant implementation of the new curricula because of the lack of training, support, and consistency they received. They expressed that within the past few years, they've noticed a pattern of trying out new interventions, but not following through with them long enough to see any long-term success. At both schools, the current curricula in place focus on developing students' code-based skills that contribute to word reading (e.g., PA, alphabet and phonics knowledge, decoding); however, they do not focus on developing students' language-based skills that contribute to linguistic comprehension (e.g., vocabulary knowledge, knowledge of syntax/language structure). Because of this, teachers expressed that BVERS, which does focus on developing students' linguistic comprehension, would be a welcomed complimentary addition to their existing literacy structure and would help address a current gap in their instruction.

Integration into Existing Classroom Schedules

One of the main concerns surrounding the integration of BVERS into the current instruction was finding uninterrupted time where all their students were present to implement the curriculum. This was especially problematic for teachers who described their classrooms as having a "revolving door" due to the high number of students being pulled out for intervention (e.g., ESOL, reading, special education) services. Additionally, though meant to be a 30-minute intervention, BVERS took teachers considerably more time to implement in their classrooms.

As previously noted, the writing component was particularly challenging to integrate into teachers' existing schedules as teachers across all grade levels stated that it would oftentimes take upwards of 20 to 30 minutes just for the writing instruction. Teachers expressed that they would oftentimes have to adjust their instruction rather than complete the writing component as intended. These adjustments varied from only completing some of the writing activities but not all, putting the writing activity into their literacy rotations rather than at the end of the 30-minute BVERS lesson, or skipping the writing all together.

Most often, general education teachers admitted to skipping the writing altogether because they felt like they were unable to quickly adapt the lesson to both fit their timeframe and provide students with the additional support they needed. However, both interventionists described that rather than skipping the writing, they would let one lesson span across multiple days until the students had completed all the components. They even suggested modifying the intervention to two weeks per book, something that many of the general education teachers disagreed with, so that they could dedicate enough time to writing. This suggests that teachers who implemented the intervention in small groups found it to be more adaptable than teachers who implemented it whole class. The interventionists who implemented BVERS in small groups were able to provide students

with the additional support they needed to successfully engage with the writing process and had the flexibility to adjust their schedules to accommodate the progress of their groups.

Though the general education teachers had originally committed to teaching BVERS during their typical ELA block, towards the middle of the seven-week implementation period, most teachers had moved BVERS to their 30-minute intervention block, which was when students got pulled for additional services. When probed on what prompted the change, teachers explained they did not want BVERS to take up their ELA block because it was the only time they received instructional support from teaching assistants, and they needed to prioritize teaching students foundational word recognition skills during this time. This suggests that even though teachers said that they recognized that there was a gap in their current instruction and that they needed to teach more language-based skills, there was some incompatibility between teachers' assumptions of the importance of linguistic comprehension (as compared to the importance of word recognition skills) and the theoretical underpinnings of the intervention. This misalignment impacted teachers' ability and willingness to integrate BVERS into their existing systems to allow for all students to receive the intervention. To help address this integration challenge and bolster successful implementation, researchers need to evaluate the degree of fit between teachers' beliefs of reading development and instruction with the intervention components, and then seek ways to promote alignment.

Potential for Meeting Diverse Needs

When examining the adaptability and effectiveness of BVERS for different populations of students, teachers generally agreed the intervention, though scripted, still allowed for some teacher autonomy and provided them with the opportunity to respond to their student's needs. Most teachers perceived the read aloud and vocabulary components to be fitting and applicable for both their non-EL students as well as their EL students. When describing the effect on non-ELs, teachers

exclusively focused on how they observed growth in their students' ability to make connection between the texts and themselves. Students not only made these connections during the BVERS lessons but would often transfer this skill to other subject areas and made connections between new content to their lived experiences.

When describing the effects for ELs, teachers almost exclusively discussed students' progress with vocabulary and their ability to remember the target words and, at times, their ability to apply them in novel situations. Though they weren't always sure that their ELs were fully comprehending each story in its entirety, teachers believed that the embedded supports in BVERS, such as the vocabulary logs and picture cards, were instrumental at helping ELs access the language in the stories.

It is important to further examine this preliminary support for BVERS's potential to impact ELs' language, and possibly literacy, outcomes. Though there has been some progress in recent years in improving ELs' language and literacy outcomes, the proportion of ELs with reading difficulties in U.S. schools is still significantly greater than the proportion of non-ELs with reading difficulties (Roberts et al., 2022). These disparities in early literacy performance are due to various interrelated factors in addition to the significant contribution of English language comprehension to reading comprehension. ELs face an opportunity gap and are more likely to be enrolled in low performing, under resourced schools (Cosentino de Cohen et al., 2005), and as a result, have decreased access to high quality teachers, texts, and educational opportunities. The absence of these resources and educational opportunities may further exacerbate reading disparities for ELs by limiting their ability to engage with and participate in experiences that contribute to language and literacy acquisition (Solari et al., 2022). Therefore, it is essential to identify effective

interventions that are feasible for teachers to implement in diverse settings to help improve ELs' language and literacy outcomes.

Overall, the results from the current study suggest that while teachers generally found BVERS to be appropriate and efficacious, there are some adjustments that could be made to increase how teacher's view its appropriateness. The writing component was a point of contention for many teachers in part because teachers' teaching philosophy did not always match what BVERS was asking them to do (i.e., engage with the writing process with young students) and because of time constraints. Additionally, general education teachers believed that BVERS, because it was implemented in a whole class setting, did not allow for the instructional supports students needed to be successful in writing. Adjustments such as additional flexibility and instructional time during the writing component and additional teacher training on science-based reading research could help enhance teachers' perception of BVERS's appropriateness and help promote sustainability within diverse classroom settings.

Feasibility

Feasibility examined the extent to which BVERS could be successfully implemented within a given school. It explored whether the intervention could be implemented relatively easily and consistently within the existing resources and circumstances of the school such as time and resource availability (Gadke et al., 2021). Overall, the results of this study support the feasibility of the intervention.

Teachers across all grade levels noted that having all the materials, including the student materials, prepared for them ahead of time helped make implementing BVERS much more feasible. This was not something that they commonly experienced with other interventions, so they were exceedingly grateful. Additionally, the teachers frequently mentioned that having the before

reading, after reading, and in-text stickers color coded and already in the books was very helpful and helped them get accustomed to the routines in the intervention quickly. A common sentiment amongst teachers was that there were a lot of individual parts to the intervention (e.g., the before reading cards, the after reading cards, the vocabulary logs, the picture cards, and the teacher's guide) and it was initially challenging to keep track of materials and the sequence of activities. However, most teachers mentioned that despite these initial challenges, by the second week of implementation, they felt much more confident in the instructional routines and the sequence of activities. One suggestion made to help with learning the intervention was to provide teachers with a visual checklist they can keep nearby to make sure they have completed each component as oftentimes, a teacher would accidentally skip over a piece, such as the after reading card, and not notice until much further along in the lesson.

As previously mentioned, a key challenge to BVERS's feasibility was the time constraints put on teachers' literacy and intervention blocks. While some classes were able to complete a lesson in the 30-minute session, others took significantly longer, stretched a lesson over multiple days, or simply didn't complete each lesson. However, teachers did mention that this challenge was not necessarily unique to BVERS and that unfortunately, they did not have control over their own schedule as it was school mandated. This suggests that to increase the feasibility of an intervention, the intervention should not only should provide teachers with materials that require minimal preparation, but they should also be designed with shorter lessons possibly spread out over a longer period rather than longer lessons for less days.

Fidelity

A key goal of this study was to determine if teachers were able to implement BVERS to fidelity as intended by the developers of the intervention (Proctor et al., 2021). Examining fidelity

of implementation is a key step for translating effects from clinical research environments to practical setting as unique implementation challenges may arise during teacher implantation in classroom-based settings which may hinder the integrity of the intervention and its effects. Overall, fidelity was moderate and hovered just under 80%, which is slightly lower than fidelity in a previous study examining the feasibility of BVERS when implemented in a one-on-one virtual setting (Henry et al., 2022). Similar to Henry et al. (2022), fidelity scores were highest for the read aloud portion of the intervention sessions and lower for the vocabulary and written activity components.

Something that was not reflected in the teacher fidelity results were the behavioral challenges teachers faced in their classroom and how that may have impacted teachers' ability to implement BVERS as intended. Though behavioral data was not collected or analyzed, anecdotal evidence from the lesson observations and recordings suggest that teachers with more challenging behaviors present in their class were consistently scored as having lower fidelity of implementation while teachers with less behavioral challenges in their class were scored as having higher fidelity of implementation. During observations, teachers oftentimes had to redirect students and, at times, completely stop the intervention to direct their efforts to behavior management. This shift in attention may have impacted teachers' ability to concentrate on the instructional protocols in the intervention and possibly miss key components leading to lower fidelity scores.

Interestingly, higher fidelity scores were associated with less experienced teachers, and lower fidelity scores were associated with more experienced teachers. For example, Francine had the second most years of teaching experience, but the lowest fidelity of implementation. Even after multiple rounds of coaching and offers to model a lesson, she continued to utilize BVERS as a loose set of guidelines for conducting an interactive read aloud with some short vocabulary

instruction if time allowed. These results suggest that less experienced teachers may be more willing to utilizing scripted interventions, such as BVERS, as a scaffold for their own teaching while more experienced teachers may not feel like they need, or even want, to follow a script since they may feel more confident in their abilities due to their extensive experience.

Limitations and Future Directions

This study should be considered in light of several limitations. One limitation in the current study is that the BVERS intervention training for teachers occurred approximately two months prior to implementation. Due to the nature of conducting research in routine school and classroom-based settings, there were a lot of delays to implementation outside of the research team's control. These delays were due in part because of a long postponement in one districts' approval processes, school cancellations due to weather, and school holiday breaks. Though the teachers had access to the curriculum during this time, this could have impacted teachers' adherence to BVERS's instructional protocols as it is possible teachers had not reviewed the lessons since their training. Future research should carefully consider the timeline of implementation closely and ensure that the teacher training and the beginning of implementation are done within a close timeframe. Additionally, refresher trainings should be made available to teachers immediately prior to and during the implementation of the intervention to aid with adherence to the intervention protocol.

A second limitation to this study is the small, limited sample of teachers that participated in the implementation of intervention and the small number of ELs who received consent to participate in the data collection process. One key dimension of feasibility research is to examine the generalizability of the intervention. However, the small samples impacted this study's ability to generalize findings, specifically when examining the student (i.e., EL and non-EL) subgroups

because the student group sample sizes varied greatly with only 15 ELs in the sample, and when examining the teacher subgroup samples (i.e., kindergarten teachers, first grade teachers, second grade teachers, interventionists) because the teacher sample size ranged from two to three teachers per group. Further, our teacher sample was fairly homogenous with all teachers identifying as White women and possessing a traditional Collegiate Professional License or a Postgraduate Professional license. A larger and more diverse teacher and student sample would allow for greater generalization of findings to broader populations of educators and students.

Additionally, there was only one teacher who worked at School A while the other nine teachers worked at School B. This made it challenging to examine how the dimensions of implementation differed based on different school and district climates, norms, and instructional philosophies. Specifically, with only one teacher offering her perspective for School A, we could not compare multiple experiences and create a cohesive understanding of the school's sociopolitical context. Future research should focus on examining the feasibility of the intervention in various settings with multiple teachers at each setting. This would allow researchers to triangulate evidence to create more wholistic representations of individual schools and their unique contexts. They could then compare various types of schools and assess which context supported effective implementation and examine the specific features of the schools and districts that were associated with successful implementation.

Lastly, though this study did collect data regarding BVERS effectiveness through the teacher survey and from anecdotes from teachers during the semi-structured focus group interviews, the current study did not report on any proximal or distal language and literacy student outcomes. Though it is not the primary focus of feasibility research, examining an intervention's effectiveness can allow researchers to explore if there is preliminary evidence of

potential for improving outcomes and whether the outcome measure being used are appropriate and sensitive to change. Future feasibility research should include student outcome measures to help guide researchers during further intervention development or during the refinement of the intervention prior to beginning a pilot study or a randomized controlled trial.

Conclusion

This study sought to evaluate the acceptability, appropriateness, and feasibility of BVERS, a research based listening comprehension and vocabulary intervention, for use in K-2 classrooms when implemented by general education teachers and interventionists. Developing students' language comprehension is essential as the relative contribution of language comprehension to reading comprehension increases over time as texts become more complex and students become more proficient decoders (Adolf et al., 2010; Catt et al., 2005; García & Cain, 2014). To help develop students' reading skills, teachers need access to interventions that are not only science-based but are also feasible to implement in their own classroom and are contextualized for their students' needs. Importantly, the findings from this study suggest that BVERS is acceptable, appropriate, and feasible for ELs and non-ELs in K-2 when implemented by teachers in practice settings. These findings indicate promise for the successful implementation of BVERS in diverse classrooms that may ultimately impact students' language and literacy outcomes.

References

Al Otaiba, S., Connor, C. M., Foorman, B., Schatschneider, C., Greulich, L., & Sidler, J. F. (2009). Identifying and intervening with beginning readers who are at-risk for dyslexia: advances in individualized classroom instruction. *Perspectives on Language and Literacy*, 35, 13–19.

Aro, T., Eklund, K., Eloranta, A. K., Närhi, V., Korhonen, E., & Ahonen, T. (2019). Associations between childhood learning disabilities and adult-age mental health problems, lack of education, and unemployment. *Journal of Learning disabilities*, *52*(1), 71-83.

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

Babayiğit, S. (2014). The role of oral language skills in reading and listening comprehension of text: a comparison of monolingual (L1) and bilingual (L2) speakers of English language. *Journal of Research in Reading, 37(1)*, S22–S47.

Beck, I. L., McKeown, M. G., & Kucan, L. (2008). *Creating robust vocabulary: Frequently asked questions and extended examples*, Vol. 10. Guilford Press.

Catts, H.W. , Hogan , T.P. , & Adlof , S.M. (2005). Developmental changes in reading and reading disabilities. In H.W. Catts & A.G. Kamhi (Eds.), The connections between language and reading dis abilities (pp. 25 – 40). Mahwah, NJ: Erlbaum

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111(6),* 982-1000.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P. S., Bayraktar, A., Crowe, E. C., & Schatschneider, C. (2011). Testing the impact of child characteristics× instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, *46*(3), 189-221.

Cosentino de Cohen, C., Deterding, N., & Chu Clewell, B. (2005). Who's left behind? Immigrant children in high and low LEP schools. Urban Institute, Program for Evaluation and Equity Research.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E.M., Reboussin, B. A., &Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities, 39(6)*, 507–514.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Every Student Succeeds Act, Pub. L. No. 114-95, Stat. 1177 (2015).

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712-751.

Ford, K. L., Cabell, S. Q., Konold, T. R., Invernizzi, M., & Gartland, L. B. (2013). Diversity among Spanish-speaking English language learners: profiles of early literacy skills in kindergarten. *Reading and Writing*, *26*(6), 889-912. https://doi.org/10.1007/s11145-012-9397-0.

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

Gadke, D. L., Kratochwill, T. R., & Gettinger, M. (2021). Incorporating feasibility protocols in intervention research. *Journal of School Psychology*, *84*, 1-18.

Gonzalez, J., Pollard-Durodola, S., Saenz, L., Soares, D., Davis, H., Resendez, N., & Zhu, L. (2016). Spanish and English early literacy profiles of preschool Latino English language learner children. *Early Education and Development*, *27*(4). 513-531.

Good, R. H., & Kaminski, R. A. (2002). DIBELS oral reading fluency passages for first through third grades. (Technical report no. 10). Eugene, OR: University of Oregon

Goodrich, J. M., Thayer, L., & Leiva, S. (2021). Evaluating achievement gaps between monolingual and multilingual students. *Educational Researcher, 50*(7), 429-441.

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257-285.

Henry, A. R., & Solari, E. J. (2020). Targeting oral language and listening comprehension development for students with autism spectrum disorder: A school-based pilot study. *Journal of Autism and Developmental Disorders, 50(10)*, 3763–3776.

Henry, A. R., Conner, C., Zajic, M. C., & Solari, E. J. (2022). Feasibility and initial efficacy of an adapted telepractice listening comprehension intervention for school-aged children with autism. *Journal of Autism and Developmental Disorders, 52*(5), 1862-1872.

Herrera, S., Phillips, B. M., Newton, Y. C., Dombek, J. L., & Hernandez, J. A. (2021). Effectiveness of Early Literacy Instruction: Summary of 20 Years of Research. REL 2021-084. *Regional Educational Laboratory Southeast*.

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391-398.

Korhonen, J., Linnanmäki, K., & Aunio, P. (2014). Learning difficulties, academic well-being and educational dropout: A person-centered approach. *Learning and individual differences*, *31*, 1-10.

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50,* 151–169.

Larson, A. L., Cycyk, L. M., Carta, J. J., Hammer, C. S., Baralt, M., Uchikoshi, Y., An, Z. G., & Wood, C. (2020). A systematic review of language-focused interventions for young children from culturally and linguistically diverse backgrounds. *Early Childhood Research Quarterly*, *50*, 157-178.

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

Lonigan, C. J., Goodrich, J. M., & Farver, J. M. (2018). Identifying differences in early literacy skills across subgroups of language-minority children: A latent profile analysis. *Developmental psychology*, *54*(4), 1-48.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300-335.

McNamara, J. K., Scissons, M., & Gutknecth, N. (2011). A longitudinal study of kindergarten children at risk for reading disabilities: the poor really are getting poorer. *Journal of Learning Disabilities, 44(5)*, 421–430.

Martin, N. (2013a). *Expressive one-word picture vocabulary test: Spanish-bilingual edition* (4th ed.). Novato: Academic Therapy Publications.

Martin, N. (2013b). *Receptive one-word picture vocabulary test: Spanish-bilingual edition* (4th ed.). Novato: Academic Therapy Publications.

National Center for Education Statistics (2019). Status and trend in the education of racial and ethnic groups. Retrieved from https://nces.ed.gov/programs/raceindicators/indicator_RAD.asp

National Center for Education Statistics. (2023). English Learners in Public Schools. Condition of Education. U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/programs/coe/indicator/cgf

National Early Literacy Panel (U.S.) & National Center for Family Literacy (U.S.). (2008). Executive summary: Developing early literacy: report of the National Early Literacy Panel. National Institute for Literacy.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: toward a model of comprehension. *Journal of Educational Psychology*, *97(2)*, 246–256.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and policy in mental health and mental health services research*, *38*, 65-76.

Roberts, G. J., Hall, C., Cho, E., Coté, B., Lee, J., Qi, B., & Van Ooyik, J. (2022). The state of current reading intervention research for English learners in grades K–2: A best-evidence synthesis. *Educational Psychology Review, 1-27*.

Rogde, K., Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, *15*(4).

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

Schrank, F. A., & Wendling, B. J. (2018). The Woodcock–Johnson IV: Tests of cognitive abilities, tests of oral language, tests of achievement. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 383–451). The Guilford Press.

Silverman, R. D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K–5 students' language and literacy outcomes. *Reading Research Quarterly, 55*, S207-S233.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72-110.

Snyder, E., Witmer, S. E., & Schmitt, H. (2017). English language learners and reading instruction: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 136-145.

Solari, E.J. & Ciancio, D. (2014). *Building Vocabulary and Early Reading Strategies (Eager BVERS).*

Solari, E. J., Henry, A. R., McIntyre, N. S., Grimm, R. P., & Zajic, M. (2020a). Testing the effects of a pilot listening comprehension and vocabulary intervention for individuals with autism. *Research in Autism Spectrum Disorders, 71*, 101501.

Solari, E. J., Terry, N. P., Gaab, N., Hogan, T. P., Nelson, N. J., Pentimonti, J. M., Petscher, Y., & Sayko, S. (2020b). Translational science: A road map for the science of reading. *Reading Research Quarterly*, *55*, S347-S360.

Solari, E. J., Kehoe, K. F., Cho, E., Hall, C., Vargas, I., Dahl-Leonard, K., Richmond, C. L., Henry, A. R., Cook, L., Hayes, L., & Conner, C. (2022). Effectiveness of interventions for English learners with word reading difficulties: A research synthesis. *Learning Disabilities Research & Practice, 37*(3), 158-174

The Nation's Report Card. (2023). *NAEP Report Card: Reading National Achievement-Level Results.* Retrieved from https://www.nationsreportcard.gov/reading/nation/achievement/?grade=4

Torgesen, J. K., & Burgess, S. R.(1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal–correlational and instructional studies.In J.Metsala& L.Ehri(Eds.), Word recognition in beginning reading(pp. 161–188). Hillsdale, NJ: Erlbaum.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of Word Reading Efficiency—Second Edition* (TOWRE-2). AUSTIN, TX: Pro-Ed.

U.S. Department of Education. (2003). *National Assessment of Educational Progress (NAEP) 2003 reading report card*. National Center for Education Statistics, Institute of Education Sciences.

U.S. Department of Education. (2015). *State Nonfiscal Survey of Public Elementary/Secondary Education*, *2014–15*. National Center for Education Statistics, Institute of Education Sciences. Retrieved from https://nces.ed.gov/ccd/stnfis.asp

U.S. Department of Education. (2017). *Our nation's English learners: What are their*

characteristics? National Center for Education Statistics, Institute of Education Sciences.

Vargas, I., Daucourt, M. C., Hall, C., Hart, S. A., & Solari, E. J. (2023). Examining the heterogeneous early literacy profiles of first-grade students who are English learners. *Reading and Writing*, 1-23.

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

Virginia Department of Education. (2022). *English Learner Education*. Retrieved from https://www.doe.virginia.gov/teaching-learning-assessment/specialized-instruction/english-learner-education

Wanzek, J., Stevens, E. A.,Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. Journal of Learning Disabilities, 51(6), 612–624.

Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation science*, *12*, 1-12.

WIDA. (2023). ACCESS for ELLs Grades K-12. Board of Regents of the University of Wisconsin System.

Wiederholt, J. L., Bryant, B. R. (2012). *Gray Oral Reading Tests—Fifth Edition* (GORT-5). Austin, TX: Pro-Ed.

Wiig, E. H., Semel, E., & Secord, W. A. (2013a). Clinical evaluation of language fundamentals–fifth edition (CELF-5). Journal of Psychoeducational Assessment, 33(5), 495–500.

`

Table 1

School Student Demographics

School	Enrollment	Percent of free or reduced lunch	Percent EL	Percent White	Percent Black	Percent Hispanic	Percent Asian	Percent Multiple Races
A	446	73.5	10.3	50.0	17.0	16.6	2.9	13.2
B	506	60.3	14.0	58.5	9.7	18.6	2.8	10.5

`

Table 2

Teacher Demographic Data

Teacher	Gender	Age	Race and Ethnicity	Current Grade Level	Highest Level of Formal Education	Teaching Endorsement	Years of Teaching Experience
School A							
Melissa	Female	31	White, non-Hispanic	Kindergarten	Master's Degree	PreK-6, SPED	8
School B							
Victoria	Female	40	White, Hispanic	Kindergarten	Master's Degree	PreK-6	8
Francine	Female	43	White, non-Hispanic	First Grade	Bachelor's Degree	PreK-6	18
Maureen	Female	34	White, non-Hispanic	First Grade	Bachelor's Degree	PreK-6	2
Ginny	Female	33	White, non-Hispanic	First Grade	Master's Degree	PreK-3	1
Sam	Female	26	White, non-Hispanic	Second Grade	Master's Degree	PreK-6	4
Mackenzie	Female	29	White, non-Hispanic	Second Grade	Master's Degree	PreK-6	5
Connie	Female	56	White, non-Hispanic	Second Grade (Long-Term Substitute)	Bachelor's Degree	PreK-6, SPED	11
Faith	Female	59	White, non-Hispanic	Reading Interventionist	Bachelor's Degree	PreK-3, SPED	34
Clara	Female	34	White, non-Hispanic	ESOL Specialist	Master's Degree	ESOL	10

`

Table 3

Mixed Methods Convergent Design

Implementation Outcome (Proctor et al., 2011)	Corresponding Dimension of the Intervention (Gadke et al., 2021)	Measures
Acceptability	Social Validity	Post-intervention teacher survey
		Post-intervention teacher semi structured interview
		Post-intervention student survey
Appropriateness	Integration	Post-intervention teacher survey
		Post-intervention teacher semi structured interview
	Adaptability	Post-intervention teacher survey
		Post-intervention teacher semi structured interview
		Lesson observations
	Effectiveness	Post-intervention teacher semi structured interview
		Lesson observations
Feasibility	Practicality	Post-intervention teacher survey
		Post-intervention teacher semi structured interview
	Implementation	Fidelity of implementation observation checklist

`

Table 4

Teacher Acceptability, Appropriateness, and Feasibility Survey Results

Question	% Strongly Disagree	% Disagree	% Neither Agree nor Disagree	% Agree	% Strongly Agree
This intervention meets my approval.	—	10	20	20	50
This intervention is appealing.	—	—	—	50	50
I like this intervention.	—	—	10	40	50
I welcome the use of this intervention.	—	—	20	30	50
This intervention seems fitting for my non-EL students.	10	—	20	20	50
This intervention seems suitable for my non-EL students.	10	—	—	40	50
This intervention seems applicable to my non-EL students.	10	—	—	40	50
This intervention seems like a good match for my non-EL students.	10	10	10	20	50
This intervention seems fitting for my EL students.	—	10	20	20	50
This intervention seems suitable for my EL students.	—	10	20	20	50
This intervention seems applicable to my EL students.	—	20	10	20	50
This intervention seems like a good match for my EL students.	—	10	20	20	50
This intervention seems implementable.	—	—	10	60	30
This intervention seems possible.	—	—	10	50	40
This intervention seems doable.	—	—	—	60	40
This intervention seems easy to use.	—	20	10	40	30

Note. Teachers' responses were measured using a 5-poing Likert scale: 1= Strongly disagree, 2= Disagree, 3= Neither agree nor disagree, 4= Agree, 5= Strongly agree.

`

Table 5

Teacher Acceptability Survey Results

Survey Item	Kindergarten n=2	First Grade n=3	Second Grade n=3	Interventionists n=2	Teacher Interview Quotes
This intervention meets my approval.	3.50 (2.12)	3.67 (0.58)	4.33 (1.15)	5.00 (0.00)	"I really liked it. I really liked the cards at the beginning of the book and the end of the book. I like that it is color coded. I like that the discussion prompts are in the book, on the page, and color coded. It was very handy. Loved that!" -Melissa (K)
This intervention is appealing.	4.50 (0.70)	4.00 (0.00)	4.67 (0.58)	5.00 (0.00)	"I like how you discuss the books [in the lesson], and I like the questions that were given. They might not have been questions that I would have come up with on my own. So, it was just a different way for them [students] to think about the books sometimes." - Francine (1st)
I like this intervention.	4.00 (1.41)	4.00 (0.00)	4.67 (0.58)	5.00 (0.00)	"I enjoyed it! If someone handed it again to me with everything in it [the books], I'd be like, 'Ahh, yes!'" – Faith (Intervention)
I welcome the use of this intervention.	4.00 (1.41)	4.00 (0.00)	4.33 (1.15)	5.00 (0.00)	"Personally, I don't really ask questions like during the book. I just, I've never really done like active read aloud. Like, this is only my second year of teaching. So BVERS was fun, like to think of a read aloud in that way." – Maureen (1st)
Total	4.00 (1.20)	3.92 (0.29)	4.50 (0.80)	5.00 (0.00)	

Note. Teachers' responses were measured using a 5-poing Likert scale: 1= Strongly disagree, 2= Disagree, 3= Neither agree nor disagree, 4= Agree, 5= Strongly agree.

`

Table 6

Student Social Validity Survey Results

	% Disagree	% Neither Agree nor Disagree	% Agree
I liked the books in BVERS.	5	24	71
BVERS helped me learn new words.	13	41	46
I learned how to answer questions using information I heard right in the book.	15	19	66
I learned how to make connections to my background knowledge.	14	18	68
I understood the books we were reading.	8	15	77
I want to keep using BVERS	20	27	53

Note. N=82

`

Table 7

Means and Standard Deviations on Student Social Validity Survey by Student Subgroups

		n	I liked the books in BVERS.	BVERS helped me learn new words.	I learned how to answer questions using information I heard right in the book.	I learned how to make connections to my background knowledge.	I understood the books we read.	I want to keep using BVERS.
Grade								
	Kindergarten	31	2.55 (0.68)	2.52 (0.68)	2.61 (0.76)	2.45 (0.72)	2.71 (0.64)	2.39 (0.84)
	1st	14	2.71 (0.47)	2.36 (0.74)	2.86 (0.53)	2.64 (0.74)	2.71 (0.61)	2.50 (0.76)
	2nd	33	2.73 (0.52)	2.15 (0.67)	2.24 (0.75)	2.58 (0.75)	2.67 (0.60)	2.18 (0.77)
EL Status								
	EL	15	2.80 (0.58)	2.60 (0.70)	2.80 (0.75)	2.733 (0.73)	2.67 (0.61)	2.53 (0.80)
	Non-EL	63	2.62 (0.61)	2.23 (0.70)	2.43 (0.77)	2.49 (0.74)	2.70 (0.61)	2.27 (0.78)
Setting								
	Whole Group	61	2.57 (0.62)	2.20 (0.70)	2.36 (0.80)	2.41 (0.78)	2.62 (0.66)	2.15 (0.81)
	Small Group	17	2.94 (0.24)	2.82 (0.39)	3.00 (0.00)	3.00 (0.00)	2.94 (0.24)	2.94 (0.24)
Total		78	2.65 (0.58)	2.33 (0.70)	2.50 (0.75)	2.54 (0.75)	2.69 (0.61)	2.32 (0.80)

Note. Students' responses were measured using a 3-point Likert scale: 1= Disagree, 2= Neither agree nor disagree, 3= Agree.

`

Table 8

Teacher Appropriateness for non-ELs Survey and Interview Results

Survey Item	Kindergarten n=2	First Grade n=3	Second Grade n=3	Interventionists n=2	Teacher Interview Quotes
This intervention seems fitting for my non-EL students.	4.00 (1.41)	3.00 (1.73)	4.33 (1.15)	5.00 (0.00)	"Reading the stories again everyday pretty much… that got them to understand more." – Faith (Intervention)
This intervention seems suitable for my non-EL students.	4.50 (0.71)	3.00 (1.73)	4.67 (0.58)	5.00 (0.00)	"They were doing a better job with their listening skills and with their book interaction skills. They're growing with their ability to listen to a story and realize, 'Oh I can learn from this story,' or, 'Oh I can connect to this story in a meaningful way, or this story can help me understand stuff about myself or about the world.'" – Melissa (K)
This intervention seems applicable for my non-EL students.	4.50 (0.71)	3.33 (1.15)	4.67 (0.57)	5.00 (0.00)	"I think it was just like a natural time to talk about the vocabulary, I mean the cards that were in there and the pictures worked really well for them to build their vocabulary." - Sam (2nd)

`

This intervention seems like a good match for my non-EL students.	3.50 (2.12)	3.00 (1.73)	4.33 (1.15)	5.00 (0.00)	"I think they really liked the books. I think 5 days might have been just a little too long for first graders even though we were breaking the book up. You know, it got to the it got to some point where they were like, 'How many days are we going to read this?'" -Francine (1st)
Total	4.13 (1.13)	3.08 (1.38)	4.50 (0.72)	5.00 (0.00)	

Note. Teachers' responses were measured using a 5-poing Likert scale: 1= Strongly disagree, 2= Disagree, 3= Neither agree nor disagree, 4= Agree, 5= Strongly agree.

`

Table 9

Teacher Appropriateness for ELs Survey and Interview Results

Survey Item	Kindergarten n=2	First Grade n=3	Second Grade n=3	Interventionists n=2	Teacher Interview Quotes
This intervention seems fitting for my EL students.	3.50 (2.12)	3.67 (0.58)	4.33 (1.15)	5.00 (0.00)	"They started making connections and remembering the vocabulary words, so that was really neat to see. They definitely remembered it and could apply it too… the small group allowed ELs the space to talk through it more." – Clara (Intervention)
This intervention seems suitable for my EL students.	3.50 (2.12)	3.67 (0.58)	4.33 (1.15)	5.00 (0.00)	"I thought it was really positive for them. I loved that the vocabulary cards had the pictures on them and that the activities went along with it. I feel like that helped them be able to access the material. They really seem to enjoy the read alouds, even if some of them maybe didn't understand exactly what was happening." – Mackenzie (2nd)
This intervention seems applicable for my EL students.	3.50 (2.12)	3.33 (1.15)	4.33 (1.15)	5.00 (0.00)	"I did see a difference. And as they've gone through, they understand the process more and they get it…They started saying things like 'Oh look, that's like Owen and his blanket!' …

`

					They were making connections and figuring it out." – Faith (Intervention)
This intervention seems like a good match for my EL students.	3.50 (2.12)	3.67 (0.58)	4.33 (1.15)	5.00 (0.00)	"I think it was nice for my EL population to have those picture cards with the sentence stems. I think it helps them a lot to give them a sentence stem. And then they create whatever back half of the sentence they need to. So, I think we definitely saw some language improvement there." – Maureen (1st)
Total	3.50 (1.60)	3.58 (0.67)	4.33 (0.98)	5.00 (0.00)	

Note. Teachers' responses were measured using a 5-poing Likert scale: 1= Strongly disagree, 2= Disagree, 3= Neither agree nor disagree, 4= Agree, 5= Strongly agree.

`

Table 10

Teacher Feasibility Survey and Interview Results

Survey Item	Kindergarten n=2	First Grade n=3	Second Grade n=3	Interventionists n=2	Teacher Interview Quotes
This intervention seems implementable.	4.50 (0.71)	3.67 (0.58)	4.00 (0.00)	5.00 (0.00)	"It was so easy like after a little bit. You didn't even have to think about it anymore. It was like second nature. Everything was there. You knew everything. It was just nice to have everything prepped and ready to go."- Maureen (1st)
This intervention seems possible.	4.00 (1.41)	4.00 (0.00)	4.33 (0.58)	5.00 (0.00)	"I think it seemed complicated at first, but then it got easier to manage as time went on." – Sam (2nd)
This intervention seems doable.	4.50 (0.71)	4.00 (0.00)	4.33 (0.58)	5.00 (0.00)	"I like that we were given all the materials like, even like the logs and little cards. I [liked] that everything was given to us already made…we weren't having to look in so many different places to find out what we needed and needed to do."- Francine (1st)
This intervention seems easy to use.	3.00 (1.41)	3.67 (0.58)	3.67 (1.52)	5.00 (0.00)	"It's scripted enough that it's easy to follow, but I do feel like I'm still able to use my degree." – Victoria (K)
Total	4.00 (1.07)	3.83 (0.39)	4.08 (0.79)	5.00 (0.00)	

`

Table 11

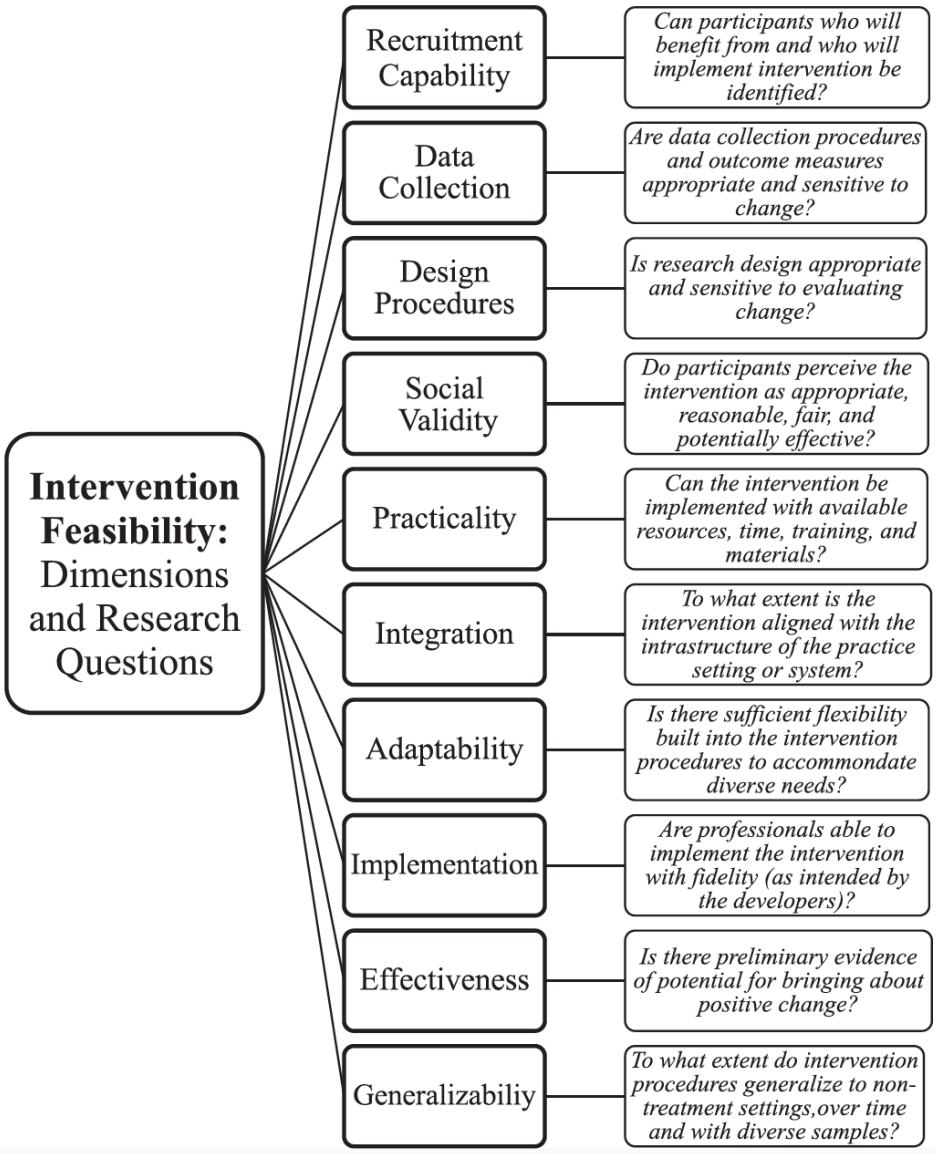
Teacher Fidelity of Implementation Results

Teacher	% Treatment Integrity			Overall Treatment Integrity (SD)
	Observation 1	Observation 2	Observation 3	
School A				
Melissa	91.00	92.00	81.00	88.00 (6.08)
School B				
Victoria	70.00	75.00	57.00	67.33 (9.29)
Francine	56.00	64.00	42.00	54.00 (11.14)
Maureen	78.00	94.00	100.00	90.67 (11.37)
Ginny	59.00	100.00	74.00	77.66 (20.74)
Sam	83.00	87.00	100.00	90.00 (8.88)
Mackenzie	83.00	100.00	76.00	86.33 (12.34)
Connie	55.00	65.00	74.00	64.67 (9.50)
Faith	75.00	74.00	79.00	76.00 (2.65)
Clara	79.00	95.00	88.00	87.33 (8.02)
Total				78.20 (15.20)

`

Figure 1

Feasibility Research Framework (Gadke et al., 2021)



`

Appendix A: Teacher Intervention Acceptability, Appropriateness, and Feasibility Survey (Weiner et al., 2017)

Item	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. This intervention meets my approval.	1	2	3	4	5
2. This intervention is appealing.	1	2	3	4	5
3. I like this intervention.	1	2	3	4	5
4. I welcome the use of this intervention.	1	2	3	4	5
5. This intervention seems fitting for my non-EL students.	1	2	3	4	5
6. This intervention seems suitable for my non-EL students.	1	2	3	4	5
7. This intervention seems applicable to my non-EL students.	1	2	3	4	5
8. This intervention seems like a good match for my non-EL students.	1	2	3	4	5
9. This intervention seems fitting for my EL students.	1	2	3	4	5
10. This intervention seems suitable for my EL students.	1	2	3	4	5
11. This intervention seems applicable to my EL students.	1	2	3	4	5
12. This intervention seems like a good match for my EL students.	1	2	3	4	5
13. This intervention seems implementable.	1	2	3	4	5

`

14. This intervention seems possible.	1	2	3	4	5
15. This intervention seems doable.	1	2	3	4	5
16. This intervention seems easy to use.	1	2	3	4	5

`

Appendix B: Teacher Semi-Structured Interview Protocol

Implementation Outcome (Proctor et al., 2011)	Corresponding Dimension of the Intervention (Gadke et al., 2021)	Question
Acceptability	Social Validity	Describe your overall experience with BVERS. What components did you like? What components did you dislike?
		Have you elicited information from students regarding their experiences with the intervention? If so, what were they?
Appropriateness	Integration	How would you describe your school and/or district climate related to accepting and implementing new interventions?
		How well does this intervention fit with your values and norms and the values and norms within your school?
		How well does the intervention fit with existing infrastructure, work processes, and practices in your setting?
		What, if any, kinds of infrastructure changes will be needed to accommodate the intervention?
		Describe your typical literacy block. How does the intervention compare (e.g., whole group vs. small group, instructional components, duration) to other existing programs in your setting?

`

	Adaptability	What kinds of changes or alterations, if any, do you think you will need to make to the intervention so it will work effectively in your setting? Are there components that should not be altered?
	Effectiveness	To what extent did the intervention meet the diverse needs of the students in your class including ELs and non-ELs? Have you seen improvement in student listening comprehension, target vocabulary knowledge, or comprehension skill usage during intervention?
Feasibility	Practicality	What barriers complicated implementation and what factors facilitated implementation? Do you feel that you were able to implement BVERS with your available resources, time, training, and materials?
Fidelity	Implementation	How complicated do you perceive the intervention to be? Please consider the following aspects of the innovation: duration, scope, intricacy and number of steps involved and whether the innovation reflects a clear departure from previous practices. How confident are you that you will be able to successfully implement the innovation without researcher support?

`

Appendix C: Student Social Validity Survey

1. **I liked the books in BVERS.**		
 Disagree	 Neutral	 Agree

2. **BVERS helped me learn new words.**		
 Disagree	 Neutral	 Agree

3. **I learned how to answer questions using information I heard right in the book.**		
 Disagree	 Neutral	 Agree

4. **I learned how to make connections to my background knowledge.**		
 Disagree	 Neutral	 Agree

5. **I understood the books we were reading.**		
 Disagree	 Neutral	 Agree

6. **I want to keep using BVERS.**		
 Disagree	 Neutral	 Agree

Appendix D: Fidelity of Implementation Checklist

BVERS Fidelity Checklist

Book _____ Date _____ OBS _____
Lesson _____

Anchor Lesson (If Applicable)

Y	N	Introduces skill
Y	N	Provides teacher model/think-aloud
Y	N	Guides student practice
Y	N	Concludes lesson and links to next activity

Before-Reading Card

Y	N	Review comprehension skill
Y	N	Review hand signal
Y	N	Introduce/review story
Y	N	Today's guiding question

Read Aloud

Y	N	Reads all stickers
Y	N	Scaffolds student answers
Y	N	Encourages use of hand signal

After-Reading Card

Y	N	Asks for/restates guiding question
Y	N	Scaffolds student answers
Y	N	Reinforces comprehension skill

Vocabulary

W1	W2	W3	
Y N	Y N	Y N	Defines word
Y N	Y N	Y N	Provides teacher model
Y N	Y N	Y N	Encourages use of word in sentence

Student Journal

Y N	Provides clear instruction for activity
Y N	Provides teacher model
Y N	Prompts child to contribute at least one idea

Overall

Y N	Instructor kept child engaged
Y N	Instructor reinforced appropriate contributions
Y N	Instructor scaffolded when needed
Y N	Instructor was organized and had materials prepared

NOTES

_____
_____
_____
_____
_____
_____

TOTAL ____ YES
____ NO