Title
Examining the Simple View of Reading among Subgroups of Spanish-Speaking English Language Learners

Author
Grimm, Ryan

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Examining the Simple View of Reading among Subgroups of Spanish-Speaking English Language Learners

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Education

by

Ryan Ponce Grimm

Committee in charge:

Professor Michael M. Gerber, Chair

Professor H. Lee Swanson, University of California Riverside

Professor Karen Nylund-Gibson

Professor Mian Wang

December 2015

The dissertation of Ryan Ponce Grimm is approved.

_____
H. Lee Swanson

_____
Karen Nylund-Gibson

_____
Mian Wang

_____
Michael M. Gerber

August 2015

Examining the Simple View of Reading among Subgroups of

Spanish-Speaking English Language Learners

ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my parents who were fully supportive of me pursuing a doctorate in Special Education, Disability, and Risk Studies. Both work in the social services and my desire to conduct research with people and their learning processes is a direct result of my upbringing. My parents' support manifested in multiple and countless ways throughout my time at UC Santa Barbara and the culmination of my work would not have been possible without them and their unwavering care. My mother would say that she looked out for my "health and welfare" while my father looked out for my "academics." For me, however, the two were neither distinct nor discernible. Health and welfare, perhaps operationalized as happiness, was closely tied to my academic pursuits. And academic pursuits could not have progressed if I was not healthy enough to seek and chase them. One was wholly entwined with the other, just as my work has been wholly entwined with my parents' support. I know I did not thank them enough along the way, so I hope this is at least a very small remedy for that.

My brothers were my punching bags when we were youths and, to this day, I am amazed at how much they put up with me. But whenever I encountered difficulties, I was always able to call on them for support and I hope they know likewise. To my brothers, I am happy to be your punching bag while you both progress through graduate school yourselves.

I certainly want to thank my advisor, Michael Gerber, who taught me to think conceptually about not only research questions, but also in applied ways and holistically about Special Education and how various stakeholders may be impacted by our work. He challenged and guided my thinking during all of our meetings, which I directly attribute to helping me develop the research questions for this study. Additionally, he helped me to

understand my research should not and will not simply be an academic endeavor and real lives must always be considered. Mike's dedication to serving those involved with Special Education is inspirational and I can only hope to demonstrate something comparable.

I provide a special thank you to Karen Nylund-Gibson whose mentorship in terms of statistical modeling is something I could not have imagined upon my arrival at UC Santa Barbara. She quickly recognized my passion for quantitative methodology and readily took me on as a mentee to foster it. Karen also provided me with opportunities to be a Teaching Assistant for quantitative methodology classes and opportunities to publish. Certainly, the methodological approach utilized in this study would not have been possible without her guidance and I look forward to many more years of working with Karen. I will always remember her response when I first asked to work with her because it provided me with confidence to pursue academia as a career: "Yes, of course."

I would like to thank the other members of my committee, H. Lee Swanson and Mian Wang. I had the pleasure of learning from and working with them prior to beginning this study and I certainly appreciate their willingness to serve on my committee. I knew their feedback would result in helping me think deeper and more critically about the questions in this study. Whether it was speaking with Lee and Mian in class, via email, or in person at conferences, they helped guide and improve my thinking about this research. Their feedback made this a stronger study than I could ever have accomplished on my own.

Of course, I would like to thank my friends whose support is indescribable. Thank you to Erika, Veronica, Myley, Sean, Ann, Mark, Jane, and, of course, Cecile. Thank you all for the conversations, in which we discussed our work and developed our research agendas. Thank you for the times in which we kept each other grounded and ensured we maintained

social lives to keep our balance. Your friendship ensured this dissertation was not a lonely endeavor. I also want to thank Cecile for reading and proofreading much of my work. Your feedback made my writing clearer and more articulate. I will never forget the times we all enjoyed together during graduate school and I am very much looking forward to our reunions.

VITA OF RYAN PONCE GRIMM
June 2015

EDUCATION

University of California, Santa Barbara — June 2015
Santa Barbara, CA
Ph.D., Special Education, Disability, and Risk Studies
Advisor: Dr. Michael Gerber

University of California, Santa Barbara — June 2015
Santa Barbara, CA
M.A., Research Methodology
Advisor: Dr. Karen Nylund-Gibson

San Jose State University — 2009
San Jose, CA
Level 1 Mild/Moderate Education Specialist Credential

Stanford University — 2003
Stanford, CA
M.A., Social Sciences in Education

University of California, Los Angeles — 2000
Los Angeles, CA
B.A., American Literature and Culture, Minor: Education

PUBLICATIONS

Quirk, M., Mayworm, A., Furlong, M., **Grimm, R.,** & Rebelez, J. (2015) Dimensionality and measurement invariance of a school readiness screener by gender and parent education levels. *International Journal of School & Educational Psychology*. Advance Online Publication. doi: 10.1080/21683603.2015.1053644

Grimm, R.P. (2015). A review of "Handbook of quantitative methods for educational research." *Structural Equation Modeling.* Advance Online Publication. doi: 10.1080/10705511.2014.955103

Guzman-Orth, D., **Grimm, R.P.**, Orosco, M., Lussier, C., Gerber, M.M., & Swanson, H.L. (2014). Psychometric properties of the Working Memory Rating Scale for Spanish-speaking English language learners. *Journal of Psychoeducational Measurement.* Advance online publication. doi: 10.1177/0734282914558710

Nylund-Gibson, K., **Grimm, R.**, Quirk, M., & Furlong, M. (2014). A latent transition mixture model using the three-step specification. *Structural Equation Modeling, 21*(3), 439-454.

Yun, J. T., Grimes, M. S., Dang, M., **Grimm, R.**, Wigginton, R., Espinosa, E., DiGuilio, L.,Hunt, E., Baldwin, E. E., & Harmon, L. (2014). *Online Instruction Pilot Project (OIPP) Final Evaluation Report*. Santa Barbara, CA: University of California Educational Evaluation Center.

Manuscripts Under Review

Quirk, M., **Grimm, R.,** Furlong, M., & Nylund-Gibson, K.L. *Latino/a children's school readiness profiles and associations with longitudinal reading growth trajectories.*

Manuscripts In Preparation

Grimm, R.P., Gerber, M.M., Leafstedt, J. & Swanson, H.L. *Literacy development in English Language Learners from kindergarten through eighth grade: A latent growth curve model.*

Nylund-Gibson, K., & **Grimm, R.P.** *Consequence of classification: A demonstration of different methods for including distal outcomes in mixture models of classification.* Paper accepted for *Frontiers in Psychology: Quantitative Psychology and Measurement*.

PRESENTATIONS

Grimm, R.P., Gerber, M.M., Swanson H.L. (2015, April). *Examining the Simple View of Reading among subgroups of Spanish-speaking English language learners.* Poster, Council for Exceptional Children Conference and Expo, San Diego.

Grimm, R. & Nylund-Gibson, K.L. (2014, September). *A demonstration of different methods to include distal outcomes in mixture models of classification.* Poster, Society for Research in Child Development, Special Topic Meeting: Developmental Methodology, San Diego.

Grimm, R., Nylund-Gibson, K.L., Quirk, M.P., Furlong, M.L. (2014, April). Latent transition analysis with different measurement models: Linking kindergarten readiness to early reading trajectories. In A.M. Gonzalez (Chair), *New twists on latent transition analysis: Unique adaptations and applications.* Symposium, American Educational Research Association, Philadelphia.

Grimm, R., Nylund-Gibson, K.L., Quirk, M.P., Furlong, M.L. (2013, August). Linking kindergarten readiness and elementary school academic performance: A latent mixture modeling approach. In A.M. Gonzalez (Chair), *Applications and advancements in latent transition analysis.* Symposium, American Psychological Association Annual Conference, Honolulu.

Grimm, R., Gerber, M.M., Swanson, H.L. (2013, July) *Examining profiles of reading development in elementary-aged Spanish speaking English Language Learners.* Paper, Society for the Scientific Study of Reading Annual Conference, Hong Kong.

Grimm, R., Victorino, C., Nylund-Gibson, K.L. (2013, April). Changes in Students' Math Attitudes During Middle School: A Latent Transition Analysis. In K. Nylund-Gibson (Chair), L*atent transition analysis: Recent advances and applications.* Symposium, American Educational Research Association, San Francisco.

Nylund-Gibson, K.L., Quirk, M.P., **Grimm, R.**, Furlong, M.L. (2013, February). *Kindergarten Student Entrance Profile predicts academic achievement through grade 5.* Poster, National Association of School Psychologists, Seattle.

Grimm, R.P., Gerber, M.M., Leafstedt, J. & Swanson, H.L. (2012, July) In V. Murphy (Chair), *Comparing models of growth in reading between at-risk English Language Learners and typically developing peers.* Paper, Society for the Scientific Study of Reading, Montreal.

UNIVERSITY TEACHING EXPERIENCE

Teaching Assistant for Graduate Level Statistics Courses
Gervitz Graduate School of Education, University of California, Santa Barbara

Roles include: Independently lead a 1-hour statistical software lecture for students (Mplus and SPSS), creating annotated lab handouts for students, creating and grading homework, holding 2-hour office hours, and serving as an assistant to the professor when needed.

- ***Structural Equation Modeling;*** Spring 2015 (Prof: Dr. Karen Nylund-Gibson)
- ***Psychometrics;*** Winter 2015 (Prof: Dr. George Marcoulides)
- ***Factor Analysis;*** Winter 2015 (Prof: Dr. Karen Nylund-Gibson)
- ***Advanced Multivariate* Statistics**; Fall 2014 (Prof: Dr. George Marcoulides)
- ***Structural Equation Modeling;*** Spring 2014 (Prof: Dr. George Marcoulides)
- ***Factor Analysis;*** Winter 2014 (Prof: Dr. George Marcoulides)
- ***Advanced Multivariate Statistics****;* Fall 2013(Prof: Dr. Andrew Ainsworth)
- ***Structural Equation Modeling****;* Spring 2013 (Prof: Dr. Christine Victorino)
- ***Factor Analysis;*** *Winter 2013 (*Prof: Dr. Karen Nylund-Gibson)
- ***Advanced Multivariate Statistics;*** Fall 2012 (Prof: Dr. Monica Dabos)
- ***Survey Design;*** *Fall 2012* (Prof: Dr. Karen Nylund-Gibson)

Teaching Assistant, The Modern University; Summer 2014
University of California, Santa Barbara
Freshman Summer Start Program
Professor: Dr. Ralph Gallucci

PROFESSIONAL EXPERIENCE

Statistical and Software Support for Faculty and Students 2014
University of California, Santa Barbara
Department of Education

Behavioral Therapist 2010 to Present (summers)
Camp Altitude, Los Gatos, CA

Special Day Class Teacher 2009 to 2010
Ralston Middle School, Belmont, CA

Lead Teacher 2003 to 2009
Esther B. Clark School at The Children's Health Council, Palo Alto, CA

RESEARCH EXPERIENCE

Kindergarten Student Entrance Profile January 2013 to Present
Researcher
University of California, Santa Barbara
Principal Investigators: Matthew Quirk, Ph.D.
Michael Furlong, Ph.D.

University of California Educational Evaluation Center (UCEC) April 2012 to August 2013
Graduate Student Researcher, UC Online Instruction Pilot Project
University of California, Santa Barbara
Principal Investigator: John Yun, Ph.D.

Growth in Literacy November 2010 to Present
Researcher
University of California, Riverside/
University of California, Santa Barbara
Principal Investigators: H. Lee Swanson, Ph.D.
Michael Orosco, Ph. D.
Michael Gerber, Ph.D.

Project La Patera November 2010 to Present
Researcher
University of California, Santa Barbara
Principal Investigator: Michael Gerber, Ph.D.

COMMUNITY SERVICE

Conference Chair 2014 to 2015
University of California Special Education, Disability and Risk Studies Annual Conference
Doctoral Student Advisory Committee

Treasurer and Local Arrangements Chair 2012 to 2014
University of California Special Education, Disability and Risk Studies Annual Conference
Doctoral Student Advisory Committee

Intervention Specialist, Assessor 2011 to 2013
McEnroe Reading and Language Arts Clinic

ACADEMIC SERVICE

Reviewer, *Learning and Instruction* (2015)
Reviewer, *Developmental Psychology* (2014)
Reviewer, *Educational and Psychological Measurement* (2014)

AWARDS

California Council for Exceptional Children August 2014
$1000 Scholarship

Children's Health Council, Palo Alto, CA. 2009
Employee Excellence Award

San Jose State University 2009
Advisory Board Member, Special Education Intern Credential Program

San Jose State University 2009
Spotlight Intern Award, Special Education Intern Credential Program

PROFESSIONAL AFFILIATIONS

American Educational Research Association
- Division C: Learning & Instruction
- Division D: Measurement & Research Methodology
- Cognition and Assessment (SIG)
- Disability Studies in Education (SIG)
- Educational Statisticians (SIG)
- Special Education Research (SIG)
- Structural Equation Modeling (SIG)

Council for Exceptional Children
    Division for Learning Disabilities
    Division for Research
Society for the Scientific Study of Reading

ABSTRACT

Examining the Simple View of Reading among Subgroups of Spanish-Speaking English Language Learners

by

Ryan Ponce Grimm

The Simple View of Reading (SVR; Gough & Tunmer, 1986; Hoover & Gough, 1990) has a longstanding history as a model of reading comprehension, but it has mostly been applied to native English speakers. The SVR posits reading comprehension is a function of the interaction between word-level reading skills and oral language skills. It has been useful in identifying subgroups of English monolinguals characterized by difficulties in word-level reading, oral language comprehension, or both (e.g. Catts, Adlof, & Weismer, 2006). However, applications investigating heterogeneous subgroups in samples of non-native English speakers are lacking. This study uses the SVR as a framework to explicitly model heterogeneity within a group of Spanish-speaking English language learners (ELLs). First, using latent profile analysis, this study empirically identified subgroups of ELLs based on reading and language skills in both Spanish and English. Three subgroups were identified, two based on relative language proficiency in Spanish and English. The first subgroup demonstrated the highest achievement across all measures, but was also characterized by relative strengths in Spanish compared to English. The second subgroup performed at the average level across most measures, but was also characterized by relative

strengths in English compared to Spanish. The third group performed the lowest and did not show demonstrate substantial relative strengths in either language. Second, a regression mixture model was conducted to examine whether the SVR functioned differently across subgroups. Results demonstrated the predictive relationships posited in the SVR were moderated by membership in the subgroups and that Spanish-speaking ELLs should not be treated as a homogenous population in terms of reading comprehension and its component skills. This study is one of the first to treat Spanish-speaking ELLs as a heterogeneous group and sheds light on conflicting results found in previous research. Implications and directions for future research are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Beginning in elementary school, proficient reading skills are fundamental to academic achievement across the curriculum as students must read texts in multiple subjects such as language arts and social studies (Storch & Whitehurst, 2002). Prior to third grade, students learn prereading skills such as phonemic awareness and the alphabetic principle that enable them to decode individual words. After third grade, students shift from learning these skills to applying them to narrative and expository texts in order to extract meaning and knowledge (Storch & Whitehurst, 2002). Therefore, learning and reading comprehension after third grade partially depend on how well students developed reading skills during early elementary (Kendeou, van den Broek, White, & Lynch, 2009; Snow, Burns, & Griffin, 1998). Students with stronger reading skills early on are more apt to experience academic success later in their academic careers. Furthermore, there is evidence that early reading skills have been linked to a variety of outcomes beyond academic achievement such as behavioral difficulties, high school dropout, and entry into the justice system (e.g. Connor, Alberto, Compton, & O'Connor, 2014; Herbers et al., 2012; Hernandez, 2011; Jones, Brown, & Aber, 2011; Miles & Stipek, 2006; Reynolds et al., 2002; Snow, Burns, & Griffin, 1998) such that stronger reading skills are associated with more positive outcomes. Given the importance of early reading skills, it is crucial for educators and other professionals who work with young students to be able to identify students struggling to develop reading skills.

Simple View of Reading

In order to identify struggling students, practitioners need theoretical models of reading comprehension that describe the development of reading acquisition processes. One well-known model of reading comprehension called the Simple View of Reading (SVR) was put forth by Gough and Tunmer (1986). In this model, reading comprehension[1] is depicted as a multiplicative interaction between linguistic comprehension and decoding skills. Linguistic comprehension (in this paper, the terms *linguistic comprehension* and *oral language* are used interchangeably) refers to a person's skills in understanding spoken language. This domain subsumes both semantic (vocabulary) and syntactic knowledge (recognition of grammar and sentence structures). Decoding skills refer to the skills necessary for a person to map speech sounds onto letters and letter combinations, and then combine them to accurately read individual words. However, correctly utilizing letter-sound correspondence rules is not a sufficient definition of decoding, according to Hoover and Gough. As irregular words do not follow traditional letter-sound correspondence rules, the researchers recognized decoding must include the ability to recognize words without explicitly using such rules. However, at the same time, word recognition necessitates the use of standard rules equating letters and letter combinations with sounds. This is an especially complex relationship in the English language as words with similar spellings may have distinct pronunciations (e.g. *bomb* and *comb*). Thus, as the SVR states, accurate word reading is a prerequisite for reading comprehension.

It is important to emphasize the multiplicative nature of the interaction between the two components of the SVR. Since it is multiplicative, a reader cannot demonstrate reading comprehension if she has zero skill in one of the two domains. For instance, if a child is

[1] An ontological discussion of the nature and complexity of reading comprehension is beyond the scope of this study, but, for an example, see Snow (2002).

unable to accurately decode any printed words, then it follows she cannot read and, therefore, cannot comprehend text. On the other hand, suppose a child is able to accurately decode individual words, but is unable to understand oral language when spoken to. Her inability to understand language precludes her from being able to assign meaning to what she has read. Thus, though she may be able to decode individual words – that is, correctly pronounce a printed word – she is unable to demonstrate reading comprehension. Additionally, both components are considered necessary for adequate reading comprehension, but neither component, by itself, is sufficient.

An important aspect of the SVR is that the associations between decoding and reading comprehension as well as oral language and reading comprehension are not static over time. In early elementary, decoding skills are stronger predictors of concurrent reading comprehension while oral language becomes a stronger predictor in middle elementary and later (e.g. Gough & Tunmer, 1990; Kendeou, van den Broek, White, Lynch, 2009; Kershaw & Schatschneider, 2010; Storch & Whitehurst, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007). As students attain mastery in decoding, there is less variance to differentiate readers and decoding becomes less predictive of reading comprehension. Thus, decoding may be viewed as a simpler skill relative to oral language given its temporal predictive capacity. Furthermore, decoding draws on lower level processes such as phonological awareness whereas oral language comprehension draws on higher level processes such as semantic (vocabulary) and syntactic (grammar) knowledge (Vellutino et al., 2007). Over time, as texts become more complex, reading comprehension places stronger demands on oral language comprehension processes. That is, older students may be able to decode unfamiliar

words, so their reading comprehension depends on whether they know the meanings of the words they read and whether they can understand the sentence structure of the text.

The SVR has received considerable attention from researchers and has been shown to be an elegant model broadly depicting causal mechanisms influencing reading comprehension (e.g. Kirby & Savage, 2008). Part of the model's utility lies in its ability to distinguish readers of varying ability levels. For instance, a struggling reader may be thought to be experiencing difficulty with either decoding or oral language or both (Hoover & Gough, 1986). Identifying the source of difficulty allows practitioners to focus their remediation efforts, hopefully leading to more effective instruction for the struggling readers (Aaron, Joshi, Gooden, & Bentum, 2008). For instance, Catts, Adlof, and Weismer (2006) studied middle school students' reading achievement and found subgroups of students could be characterized by deficits in decoding or linguistic comprehension or no deficits. This led the authors to advocate for the use of the SVR in terms of a classification system to identify struggling readers. Additional research has also documented the effectiveness of using SVR components to identify struggling readers (e.g. Catts, Fey, Zhang, & Tomblin, 1999; Nation, Cocksey, Taylor, & Bishop, 2010). Even within a subgroup of poor readers, Catts, Hogan, and Fey (2003) were able to further classify students into more specific subgroups based on relative skill strengths in oral language and word level skills. In other words, researchers have repeatedly found the SVR to be effective in terms of explaining heterogeneity within various samples of students.

Reading Achievement and Latino/a Students

The SVR studies cited thus far generally sought to identify and characterize readers who experienced difficulties due to mild or moderate learning disabilities. However, this is

not the only population to historically underperform relative to their peers. Latino/a students have also historically underperformed on measures of reading achievement compared to White students (e.g. Hemphill & Vanneman, 2011; Kao & Thompson, 2003; Lee, 2002; Reardon & Galindo, 2009; Rumberger & Anguiano, 2004; Snow, Burns, & Griffin, 1998). While a number of variables may contribute to this achievement gap, such as poverty (e.g. Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hernandez, 2011; Rumberger & Anguiano, 2004), examination of these variables is beyond the scope of this study. This gap in literacy achievement persists throughout elementary and secondary school (Hemphill & Vanneman, 2011; Kao & Thompson, 2003; Lee, 2002; Reardon & Galindo, 2009; Rumberger & Anguiano, 2004) and may influence lower rates of post-secondary degree attainment by Latino/a students. The Bureau of Labor Statistics (BLS; 2014) reports that only 26% of Hispanic or Latino/a workers attained an Associate degree or higher compared to 38% of African-Americans, 48% of Whites, and 67% of Asian-Americans. This is problematic because the BLS has projected the fastest growing occupations between 2012-2022 will require a post-secondary degree (BLS, 2013). Since Latino/a students are the fastest growing demographic in the United States (Gándara & Contreras, 2009), they represent a sizeable population with the potential to contribute to skilled occupations requiring a post-secondary education.

Given knowledge of this achievement gap, it is important to identify and provide intervention to Latino/a students who are struggling with literacy and reading early in their academic careers. Research has shown Latino/a students may even enter kindergarten at a disadvantage compared to their White counterparts (Lonigan et al., 2013; Rumberger &

Anguiano, 2004) and this is predictive of later achievement in early elementary (Quirk, Nylund-Gibson, & Furlong, 2013).

A subset of Latino/a students are learning English and speak Spanish as a first language referred to as Spanish-speaking English language learners (ELLs). According to the U.S. Department of Education's National Center for Education Statistics (NCES, 2014), the percentage of ELLs in U.S. schools grew from 8.7% in 2002-03 to 9.1% in 2011-12. The latter percentage represents approximately 4.4 million students. Furthermore, the greatest percentage of ELLs reside in urban neighborhoods (NCES, 2014), with greater numbers of low-income areas and less access to resources and high quality teachers (e.g. Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003) compared to suburban environments. Among K-12 students with limited English proficiency (LEP), more than three-quarters (77.2%) report Spanish as their home language (Batalova & McHugh, 2010). Thus, Spanish-speaking ELLs represent a sizeable group of potential workers who can contribute to the U.S. labor force and economy.

Research has demonstrated ELLs struggle with reading comprehension compared to native-English speakers (Fry, 2007; Kieffer, 2008, 2010; Snow, Burns, & Griffin, 1998). The most recent data from the National Assessment of Educational Progress indicate that only 20% of Latino/a students score proficient or better in grade 4 reading (NAEP, 2013). Disaggregating Latino/a students by English language proficiency, only five percent of Latino/a students classified as ELL scored proficient with none scoring as advanced. This is in stark contrast to 26% of Latino/a non-ELL who scored as proficient or better. Further, 71% of Latino/a ELL scored at below basic compared to 35% of Latino/a non-ELL. Finally, 23% of Latino/a ELL scored at basic while 38% of Hispanic non-ELL scored as such. Even

studies that have controlled for poverty demonstrate ELLs underperform native-English speakers. For example, Kieffer (2010) found ELLs remained one and a half times more likely to experience reading difficulties by third grade compared to their native-English speaking counterparts after accounting for poverty. Thus, ELL status appears to be a risk factor for students struggling with English-reading skills.

Research that seeks to understand reading acquisition processes within the ELL population is timely and necessary. While there is copious reading research regarding native-English speakers, less is understood about students who must navigate between two languages. Additionally, the vast majority of the extant literature concerning Spanish-speaking ELLs treats these students as a single, homogenous population. Language proficiency exists along a spectrum and future research should account for heterogeneity among Spanish-speaking ELLs. For example, what characteristics differentiate the five percent of Spanish-speaking ELLs who scored as proficient on the NAEP from the rest of Spanish-speaking ELLs? Though classified as ELL, are their English speaking and/or reading abilities more proficient than those of other Spanish-speaking ELLs?

The Simple View of Reading and English as a Second Language

Evidence for the tenability of the SVR as applied to students who speak English as a second language has been confirmed in empirical studies (e.g. Mancilla-Martinez & Lesaux, 2010; Pasquarella, Gottardo, & Grant, 2012; Proctor, August, Snow, & Barr, 2010). For example, Proctor et al. (2010) examined a sample of native-Spanish speaking fourth graders who were literate in both Spanish and English. Their findings resulted in a model of Spanish reading comprehension that paralleled the SVR. Namely, Spanish oral language and Spanish decoding skills were both related to Spanish reading comprehension, but oral language was

the stronger predictor in fourth grade, which is consistent with the dynamic nature of the SVR. These results were similar to those found by Hoover and Gough (1990) who also used a sample of bilingual students. This latter study explicitly included a longitudinal component demonstrating the changing relationships between decoding and oral language to reading comprehension over time. However, neither of these studies focused on Spanish-speaking ELLs exclusively. Rather, their samples consisted of students who were considered biliterate and received reading instruction in Spanish and English. Proctor et al. (2010) operationally defined biliterate students as only those who were able to draw upon "the entire host of Spanish and English literacy skills working concomitantly" (Proctor et al., 2010, p. 9). By definition, Spanish-speaking ELLs may not be able to utilize English literacy skills with the same precision as they can use Spanish skills, particularly with regards to oral language. Therefore, it is difficult to generalize these findings to Spanish-speaking ELLs in early to middle elementary.

The SVR has also been applied to Spanish-speaking LM learners in late elementary and middle school, but results were not consistent with typical findings from studies of the SVR in native-English speaking students (Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, & Snow, 2009) . Specifically, Mancilla-Martinez et al. (2009) found English word reading (i.e. decoding) skills were more strongly associated with English reading comprehension than linguistic comprehension skills. This was surprising as studies with native-English speaking students have found linguistic comprehension to be a stronger predictor in older students who have developed their decoding abilities (e.g. Nation & Snowling, 2004; Ouellette & Beers, 2010; Roth, Speece, & Cooper, 2002; Snow et al., 2008; Vellutino et al., 2007). The authors posited reading comprehension processes in older LM

students may be similar to those of younger native-English speakers. However, the sample did not include students who were classified as limited English proficient, so these findings may not extend to ELLs. Similar to the vast majority of studies concerning LM and ELL students, the authors did not consider the possible linguistic heterogeneity that may have potentially existed within their sample. For example, the strength of the association between word reading and reading comprehension may not have applied equally to students with varying levels of relative language proficiency. These questions remain unanswered, especially with Spanish-speaking ELLs in younger grades than middle school.

Statement of the Problem

If heterogeneity with respect to reading comprehension has repeatedly been found with native-English speakers, then it may be reasonable to expect heterogeneity in terms of reading comprehension among Spanish-speaking ELLs. Yet the vast majority of the research literature has not acknowledged such heterogeneity. That is, research generally treats ELLs as a single homogenous group. Thus, results from these studies may not generalize to students who differ in terms of relative language proficiency. A simple example illustrates this point. Catts et al. (2006) found three levels of English oral language proficiency in their sample. Moreover, decoding skills were found to differentiate the three groups. If such heterogeneity has been found in English, then similar results can reasonably be expected in other languages, particularly Spanish as the two languages share orthographic similarities. Moreover, heterogeneity regarding reading comprehension among Spanish-speaking ELLs may be more complex than that in a sample of students who speak only one language. The question arises regarding the functioning of decoding and oral language to reading comprehension in students who are exposed to both languages.

To this author's knowledge, only two studies have attempted to examine heterogeneous reading profiles among Spanish-speaking ELLs. Ford, Cabell, Konold, Invernizzi, and Gartland (2012) utilized cluster analysis and found four distinct profiles of kindergarten Spanish-speaking ELLs based on measures of early literacy skills (i.e. phonological awareness, alphabet knowledge, and orthography) assessed in the fall. However, given the age of their sample, these researchers did not examine reading comprehension, nor did they examine heterogeneity in terms of the SVR. Guzman-Orth (2013) used latent transition analysis to identify four distinct latent classes based on Spanish and English bilingual oral language proficiency. The latent classes in her study were ordered such that there were high, medium, and low performing bilingual latent classes as well as a latent class characterized by English dominance. Further, findings from her study revealed these classes were stable across two years and predictive of reading comprehension performance in a third year. However, the primary focus of her study was heterogeneity in oral language skills and did not test a formal model of reading comprehension such as the SVR.

This study seeks to fill this gap in the literature by using latent profile analysis to identify empirically-derived latent classes of second grade Spanish-speaking ELLs based on measures of decoding and oral language assessed in both Spanish and English. Second, this study proposes to examine whether the SVR – using the same measures assessed in third grade – functions similarly across the emergent latent classes with a regression mixture model. That is, will the predictive capacities of decoding and oral language to reading comprehension be similar in magnitude across the classes? Finally, this study will be

conducted with both Spanish and English reading comprehension yielding a model that will allow for cross-language comparisons.

Chapter 2

Literature Review

This study explores the functioning of the SVR across multiple subgroups of elementary-aged Spanish-speaking ELLs using structural equation modeling methodology. As such, the literature presented here focuses on empirical studies that report on the contributions of word-level reading and oral language to reading comprehension using structural equation modeling. The literature collected and reviewed was peer-reviewed and obtained by searching the ProQuest Social Sciences electronic database. This database was chosen because it provides a comprehensive search of 23 social sciences databases, including ERIC and PsycINFO, two commonly utilized databases in educational research. This was viewed as providing a more efficient and comprehensive search compared to searching ERIC or PsycINFO individually. Different combinations of the search terms "English language learner", "Simple View of Reading", "reading comprehension", and "reading growth" were used to search for relevant research articles. The References sections of the chosen articles were also examined for potentially relevant research literature.

Research was included if it was published in English, utilized a structural equation modeling framework, and included the three primary components of the SVR. The literature search was restricted to articles utilizing structural equation models to align with the methodology used in this study. Research articles that presented augmented SVR models were included if they examined the relationships between the three primary SVR components as part of the analysis. Research conducted using samples of students who spoke languages other than Spanish and English was not included here as the focus of this study concerns Spanish-speaking ELLs. Though Hoover and Gough (1990) did not utilize

structural equation modeling, this seminal article was included to review the initial empirical evidence for the SVR. The literature search resulted in a total of 27 articles.

While Gough and Tunmer (1986) provided the conceptual framework for the SVR, Hoover and Gough (1990) provided the empirical basis for the model's viability. The researchers followed a sample of 254 Spanish and English speaking bilingual students from Kindergarten to fourth grade. They administered a battery of annual assessments that included measures of decoding, listening comprehension, and reading comprehension. The authors found the product of decoding and listening comprehension explained a substantial proportion of the variance in reading comprehension each year beginning in first grade (Kindergarten data were not analyzed). Moreover, the authors tested a linear combination of the two components, but found the multiplicative interaction consistently explained a slightly greater proportion of variance. To provide further evidence for the SVR, the researchers further tested their model according to level of reading comprehension skill. Consistent with the theorized model, they found readers with decreased reading comprehension skills had relatively disparate levels of decoding and listening comprehension skills, but this profile was not found in readers with higher levels of reading comprehension. Adequate readers were found to have similar levels of both decoding and listening comprehension skills. That is, among struggling readers, some were found to have relatively high decoding skills compared to their listening comprehension skills while others exhibited the opposite pattern. This aligned with the definitions of reading disability provided above. The SVR, then, appeared to provide an elegant model of reading comprehension, capable of describing typical reading development while providing an explanation for atypical reading development.

Subsequent studies have also found evidence for the SVR using correlational and multiple regression techniques. One primary drawback of these approaches, however, is they do not account for potential measurement error in the variables of interest. Structural equation models, on the other hand, are directly able to model measurement error, which can lead to increased accuracy when reporting results. Another advantage lies in the conceptualization of the SVR components as latent variables. The measurement of latent variables typically requires multiple observed variables, thereby strengthening the representation of the latent construct when the observed variables are highly related to the latent construct. For these reasons, and because the present study utilizes a structural equation modeling framework, the research literature that follows includes only studies that investigated the SVR using structural equation modeling techniques.

SVR in English Using SEM

Recent support for the original conceptualization of the SVR has been found in samples of English speakers. Kendeou, Savage, and van den Broek (2009) and Tunmer & Chapman (2012) both tested the SVR in samples of young children (ages ranged from four to eight years old) using factor analytic techniques. Not surprisingly, both decoding and oral language latent factors were found to uniquely explain large portions of variance in reading comprehension. Similarly, Kendeou, van den Broek, White, and Lynch (2009) created latent factors of decoding and oral language skills in kindergarten and found these to be stable through second grade. Furthermore, second grade decoding and oral language factors predicted a concurrently measured observed reading comprehension variable. Taken together, these results suggest using a latent variable framework to represent the SVR is not only viable, but is likely preferred over using single measures of decoding, oral language,

and reading comprehension. While these studies were conducted with samples of English monolinguals, these results are viewed as creating a foundation for the theoretical and methodological approaches utilized in this study.

The original SVR characterized decoding and oral language as independent components, each contributing uniquely to reading comprehension. However, more recently this notion has been challenged. Research using structural equation modeling has examined whether a predictive relationship exists between decoding and oral language, as opposed to simply a correlational relationship. Both Kendeou et al. (2009) and Tunmer and Chapman (2012) examined this, but results were not consistent. The former study found oral language could predict concurrent decoding in preschool, but not in kindergarten or second grade, while the latter study found oral language could predict concurrent decoding in third grade. From an intuitive standpoint, it may make sense that oral language is able to predict decoding. For example, if a student is presented with an unfamiliar word, she may be more likely to correctly decode the word if she currently has the word in her vocabulary. As she begins to decode the word by individual graphemes, the letter-sound correspondences may trigger her memory to recall words with similar beginnings. This would allow her to correctly decode the word without reading it in its entirety. The key here is she must already possess the word as part of her oral vocabulary. However, it is somewhat puzzling why this relationship may exist in preschool, disappear during kindergarten through second grade, and reemerge in third grade. But it should also be emphasized these findings may be sample-specific as only two studies came to these conclusions. Another possibility may arise from the different measures used to operationalize the SVR components in the studies. Though a relationship between oral language and decoding may exist, this relatively minor adjustment

to the SVR does not change the larger premise concerning the predictive capacities of oral language and decoding to reading comprehension.

A small number of studies have sought to alter the original SVR as applied to English monolinguals using SEM techniques. In such cases, researchers included additional components such as cognitive processes (Vellutino, Tunmer, Jaccard, & Chen, 2007) and fluency (Adlof, Catts, & Little, 2006; Kershaw & Schatschneider, 2010; Silverman, Speece, Harring, & Ritchey, 2013) or question the multiplicative interaction between decoding and oral language (Kershaw & Schatschneider, 2010). Although these studies provide a more nuanced account of reading comprehension and its underlying processes, the original broad SVR model depicting decoding and oral language as the primary contributors to reading comprehension remains intact. Moreover, this finding is generally found using samples from a variety of k – 12 ages. For instance, Vellutino et al. (2007) added a host of cognitive processes predicting decoding and oral language and termed their model the Convergent Skills Model. Utilizing two samples of younger (grades two and three) and older (grades six and seven) students, the researchers were able to explain an impressive 80% of variance in reading comprehension in the younger group and 77% of variance in the older group. Though the inclusion of many cognitive processes explained a large amount of variance, ultimately, the authors found "…these data can be taken as additional confirmation for…Gough and Tunmer's (1986) Simple View model…" (Vellutino et al., 2007, p. 26). Additionally, the researchers obtained some unexpected results regarding relationships among the cognitive processes. Therefore, the addition of several cognitive processes may not be viewed as advantageous to teachers. Rather, a simpler model that allows them to focus on the two primary processes may be preferred.

Similar to Vellutino et al. (2007), Kershaw and Schatschneider (2010) further confirmed the contributions of word reading and oral language in older students (grades three, seven, and ten). However, they found an additive model (as opposed to a multiplicative interaction) fit the data better. Regardless of the specific nature of the relationship, however, the contributions of word reading and oral language were uncompromised. Both studies also found longitudinal asymmetry in the contributions of word reading and oral language, with oral language becoming the stronger predictor as students matured and word reading approached automaticity. The above studies may appear to improve upon the original SVR, but these improvements appear to be relatively minor in light of the major contributions of word reading and oral language. As Tunmer and Chapman (2012) noted, the intent of the original SVR was to provide a simple, coarse understanding of reading comprehension.

Another aspect of reading that has received attention in the literature related to the SVR is reading fluency. The prevailing notion is if a child is able to read quickly, she is able to devote fewer cognitive resources to decoding words, thus allowing more cognitive resources to focus on comprehending the written text. Three studies (Adlof , Catts, & Little, 2006; Kershaw & Schatschneider, 2010; Silverman, Speece, Harring, & Ritchey, 2013) used SEM to examine the potential for adding reading fluency in the SVR as an individual component. Adlof et al. (2006) examined this both cross-sectionally and longitudinally in a sample of 604 children assessed in second, fourth, and eighth grades. Reading fluency was not found to be significantly and uniquely related to reading comprehension in any of their analyses. However, their longitudinal analyses shed light on the development of fluency. In second grade, reading fluency was not statistically different from word reading. Thus, any

variance in reading comprehension that was explained by fluency was shared with word reading. Additionally, oral language was not a strong predictor of reading comprehension. This finding is in line with previous research demonstrating a stronger link between word reading and reading comprehension in younger grades. By fourth and eighth grades, fluency emerged as a separate latent factor, but still did not explain variance in reading comprehension once word reading and oral language were controlled. Longitudinally, second grade fluency did not predict fourth grade reading comprehension and fourth grade fluency did not predict eighth grade reading comprehension over and above word reading and oral language.

However, the findings of Adlof et al. (2006) conflict with results found by Kershaw and Schatschneider (2010) and Silverman et al. (2013). Kershaw and Schatschneider (2010) found fluency consistently predicted concurrent reading comprehension in third, seventh, and tenth grades over and above word reading and oral language. However, these authors used three times the number indicators of reading fluency compared to Adlof et al. (2006) and the measures were qualitatively different. Kershaw and Schatschneider (2010) included nine measures using entire passages whereas Adlof et al. included two measures of single word reading and one measure of passage fluency. Similarly, Silverman et al. (2013) included a greater number of measures of reading fluency compared to Adlof et al. (2006). Treating reading fluency as a latent factor, Silverman et al. (2013) found it to be a statistically significant predictor of reading comprehension, though, from a practical standpoint, it only explained 4.7% of the variance over and above decoding and reading comprehension. However, these researchers went one step further and found fluency mediated the relationship between decoding and reading comprehension in their fourth-

grade sample. The authors noted this finding was in line with other research suggesting fluency acts as a bridge between decoding and reading comprehension.

There may be more than one way to explain the differences across these studies. First, the different ways in which each group of researchers operationalized and measured fluency may be pivotal. Cutting and Scarborough (2006) found using different measures of reading comprehension resulted in different results, even within the same group of children. It would not be unreasonable to posit this may also be the case for measures of reading fluency (each group of researchers also used different measures of reading comprehension). Second, the dynamic nature of reading fluency found by Adlof et al. (2006) may explain the difference with Silverman et al. (2013). The latter authors did not examine fluency longitudinally, but Adlof et al. (2006) found fluency was not a distinct construct until fourth grade, which may suggest Silverman et al. (2013) could have found different results had they examined younger students. Whether fluency has a place in the SVR as an independent construct remains unanswered and should be a topic of ongoing research.

Though the primary aim of the above studies was to alter the original SVR, a secondary finding that was consistently gleaned from them is that word reading and oral language, broadly conceived, remain the primary contributors to reading comprehension across time. Thus, the present study retains the original conceptualization of the SVR in terms of the predictive capacities of word reading and oral language to reading comprehension. The current study does not seek to examine the interaction of word reading and oral language (multiplicative versus additive), nor does it examine the contribution of reading fluency. Rather, this study focuses on the potential for differential predictive effects in subgroups of Spanish-speaking ELLs.

The SVR in Spanish-speaking LM and ELLs Using SEM

The SVR in English has garnered support from many empirical studies using a variety of methodological approaches. Few would argue against the need to identify a theoretical model of reading development in populations other than native-English speaking monolinguals. Given that different languages are composed of different orthographies and rely on different grammatical structures, it is possible for reading acquisition processes to vary across multiple languages. Identifying similarities and differences in reading development in multiple languages can facilitate better understanding of effective methods to teach reading. Further, reading acquisition processes may differ between students who navigate between two or more languages and students who speak and read in only one language. For instance, the literature presented above consistently demonstrated oral language skills become a stronger predictor of reading comprehension as students mature. For ELLs who are learning a new language, oral language skills will likely lag behind their monolingual peers. However, before undertaking comparisons between different types of learners, it is first important to establish a model of reading development in children for whom English is not the first language.

The SVR has been found to be tenable in samples of early and middle elementary Spanish-speaking ELLs (Gottardo & Mueller, 2009; Proctor, Carlo, August, & Snow, 2005). Proctor et al. (2005) examined reading acquisition in Spanish-speaking ELLs, but limited their investigation to measures assessed in English. Their results demonstrated similar relationships among the SVR components found with English monolinguals. Importantly, approximately two-thirds of their fourth grade sample initially received reading instruction in Spanish, which was later followed by instruction in English (the rest received reading

instruction only in English). Given their findings aligned with the SVR, the authors posited that similarities between bilingual and monolingual reading instruction may be facilitative of English reading acquisition in Spanish-speaking ELLs. Gottardo and Mueller (2009) included Spanish measures of decoding and oral language in addition to English, but their outcome remained English reading comprehension only. These authors found only English predictors were significantly related to English comprehension, which was not surprising given their sample of first and second graders were instructed in English. Kieffer and Vukovich (2012) longitudinally tested the SVR in a sample of LM (i.e. not ELL) learners, but included an interaction between decoding and oral language, which was missing from the prior two studies. They found the interaction using first and second grade measures predicted third grade English reading comprehension over and above the contributions made by decoding and oral language, which was supportive of the SVR. Additionally, they compared the LM sample to native-English speakers from the same low-SES background and found the SVR applied well to both groups.

These three studies provide the basis for using the SVR with Spanish-speaking ELLs in the present study. Taken together, they tested most aspects of the SVR in native-Spanish speakers learning English. That is, the predictors were measured in both Spanish and English, and the original three paths between word reading, oral language, their interaction, and reading comprehension were all examined. However, there are few studies in the literature that investigate Spanish reading comprehension. If there are indeed enough similarities between these languages to allow analogous instructional approaches to be effective in both languages, then one might expect the magnitude of the relationships between word reading, oral language, and reading comprehension to be comparable across

languages. This study begins to examine this question by including measures of Spanish not only for word reading and oral language, but also for reading comprehension. Thus, this study will be able to examine full SVR models in both languages and will be able to make comparisons.

Given evidence of the viability of the SVR in native-Spanish speaking students (in terms of English reading comprehension), a subsequent question might seek to examine if the asymmetric predictive capacities of decoding and oral language apply to samples of Spanish-speaking ELLs. In English monolinguals, decoding has been found to be a stronger predictor of reading comprehension in early elementary while oral language becomes a stronger predictor in middle elementary and later years. The extant literature base concerning Spanish-speaking ELLs and LM students, however, has not reached a consensus. Native-Spanish speaking readers who are simultaneously learning English may be unique in that they may not acquire the necessary vocabulary or academic language skills to facilitate reading comprehension at the same rate as students whose native language is English. Thus, expected developmental shifts in decoding and oral language may not take place at the same time as English monolinguals, if at all. Furthermore, it is likely that any developmental shift in ELLs is dependent on English oral language proficiency rather than a particular age band.

Lesaux, Crosson, Kieffer, and Pierce (2010) studied a sample of upper elementary Spanish-speaking ELLs in fourth and fifth grades. At these grade levels, the original SVR would posit *both* decoding and oral language would independently explain variance in reading comprehension, but oral language would be the stronger predictor. Compared to national norms, their sample displayed average word reading skills, but below average oral language and reading comprehension in English. In fact, word reading did not significantly

predict any variance in English reading comprehension, which was a striking finding and conflicts with research on English monolinguals. However, the findings also suggested poor reading comprehension resulted from poor oral language skills, which does align with past research concerning English monolinguals (e.g. Kendeou, van den Broek, White, Lynch, 2009; Kershaw & Schatschneider, 2010; Vellutino, Tunmer, Jaccard, & Chen, 2007). On the other hand, Kieffer and Vukovich (2012) observed early elementary word reading was not predictive of later reading comprehension as the SVR would suggest. These authors followed a group of Spanish-speaking LM students from first through third grade. They found first and second grade oral language and the interaction between oral language and word reading were predictive of third grade English reading comprehension. Skills related to word reading, however, did not have a unique role in predicting reading comprehension. Along with Lesaux et al. (2010), this might highlight the prominence of oral language skills in Spanish-speaking LM students and ELLs. Specifically, across the elementary years, oral language was consistently the stronger predictor relative to word reading. While it would be expected to find oral language significantly predicts reading comprehension in fourth and fifth grades (as in Lesaux et al., 2010), the fact these studies did not find word reading to be statistically significant across the elementary years is surprising.

Yet, as stated earlier, the research has found conflicting results. The opposite pattern has also been reported in the research literature. Namely, studies have shown English word reading is indeed the stronger predictor of English reading comprehension in Spanish-speaking ELLs. For instance, Mancilla-Martinez and Lesaux (2010) followed 173 Spanish-speaking children from preschool through fifth grade. They used latent growth curve modeling to examine predictive abilities of initial word reading and oral language skills (age

4.5) and their rate of change (age 4.5 to 11) in relation to reading comprehension at age 11. No developmental shift occurred as would be expected and word reading maintained a stronger relationship with later reading comprehension than oral language. In other words, both the initial status of word reading and its rate of change were more highly predictive of later reading comprehension than the initial status of oral language and its rate of change. Nakamoto, Lindsey, and Manis (2008) reported similar findings. These authors conducted a SEM in which sixth grade English and Spanish reading comprehension were predicted by third grade decoding and oral language in both languages. Within each language, third grade decoding had a stronger predictive relationship with reading comprehension than oral language. Perhaps even more surprising, Mancilla-Martinez et al. (2009) found the greater effect of word reading compared to oral language continued into middle school. These results led Mancilla-Martinez et al. (2009) to hypothesize that reading development in older ELL and LM students might be analogous to reading development in younger monolingual students. While this may be plausible, it does not reconcile the findings of those studies that found oral language to be the stronger predictor in elementary school (Kieffer & Vukovich, 2012; Lesaux et al., 2010).

The way in which the SVR components have been operationalized in Spanish-speaking ELL and LM students has been offered as one explanation for conflicting results (e.g. Mancilla-Martinez et al., 2009). Another explanation might be that dynamic processes actually do apply to this population, but they may be conditioned on other variables such as relative language proficiency. For instance, it may be possible some students in the above studies did experience developmental shifts, but these individual differences were masked

when averaged over the entire sample. Thus, heterogeneity in relative language proficiency in either Spanish or English may play a key role in explaining these inconsistencies.

Kieffer (2008) examined a national longitudinal dataset that followed 17,385 students from Kindergarten through fifth grade. He compared developmental reading achievement trajectories of three groups of students classified in Kindergarten as native English speakers, LM students with full English proficiency, or LM students with limited English proficiency[2]. Thus, he accounted for heterogeneity in oral English language skills, albeit with a single observed variable. Results showed those LM students with full English proficiency developed at a rate strikingly similar to their native-English speaking counterparts. Those with limited English proficiency, however, performed dramatically lower than both of the other groups of students. While this study was not directly related to the SVR, the results may allow researchers to infer the effects of oral language proficiency likely apply to Spanish-speaking ELLs. The bulk of the studies included in this review so far have treated their samples as homogenous groups. But it is likely their samples did not consist solely of students with equivalent levels of English language proficiency. As mentioned earlier, the present study seeks to explicitly model linguistic heterogeneity in a sample of Spanish-speaking ELLs. This may have the potential to uncover differentiated predictive effects of word reading and oral language on reading comprehension.

Studies Modeling Heterogeneous Readers

In comparison to the amount of research that has been conducted with the SVR, there is a paucity of studies that have attempted to classify students into latent subgroups based on reading skills. Of those studies that have done so, only two used Spanish-speaking ELLs and

[2] As the data were drawn from a national dataset, LM students spoke multiple languages, so the results are not specific to Spanish-speaking ELLs.

only one of those used the SVR as a backdrop. This study intends to use a regression mixture model to create latent classes of Spanish-speaking ELLs based on language proficiency in Spanish and English. Whereas Kieffer (2008) used a single indicator of oral language proficiency, the present study will use multiple indicators of language proficiency based on the components of the SVR. As it is axiomatic that students are highly variable in acquiring reading and comprehension, mixture models seem to be an ideal tool for studying such heterogeneity. Mixture models may be gaining in popularity in the social sciences, but to date only a handful have been conducted in order to investigate reading acquisition, development, and comprehension.

Most of the studies presented here do not directly address the SVR, nor do they inform the theoretical framework used in this study. However, they are briefly reviewed here as applications demonstrating the potential utility of mixture models in reading research.

Three studies have used latent transition analysis (LTA) to track the longitudinal development of reading. LTA is considered the longitudinal extension of latent class (LCA) and latent profile analysis (LPA; more detail on LPA is provided in the Methods section). In these studies, multiple latent classes are estimated at each time point and subjects are placed into the one class to which they have the highest probability of belonging. As latent classes are categorical, this has been referred to as a type of stage-sequential analysis (Kaplan, 2008).

Kaplan and Walpole (2005) used a national longitudinal dataset to investigate transitions between latent classes from fall of kindergarten to spring of first grade. Using five indicators of early reading skills at four timepoints, these authors found five latent classes fit the data at all timepoints. The five indicators represented mastery of reading skills

at increasing levels of difficulty. The emergent classes were ordered in terms of students' probabilities of demonstrating mastery on successive indicators. Results further demonstrated students transitioned from lower achieving classes to higher achieving ones over time, but increases in poverty were associated with less advancement into higher achieving classes.

Compton, Fuchs, Fuchs, Elleman, and Gilbert (2008) and Catts, Compton, Tomblin, and Bridges (2011) both used LTA to identify students with reading disabilities. In the study by Compton et al. (2008) two latent classes representing typically developing readers and students with reading disabilities, were found at first and fourth grades. Results showed the majority of students remained in their respective latent class over time, but 7% of typically developing readers transitioned into the reading disabled class. Catts et al. (2011) identified four latent classes representing normal readers, readers with word reading disabilities, readers with comprehension disabilities, and readers with disabilities in both. This study spanned four timepoints from second through tenth grade. As with Compton et al. (2008), the majority of students remained in the same latent class over time, but some did transition from normal readers into one of the reading disabled classes. Both of these studies identified heterogeneous groups of English monolinguals. Results from both of these studies found evidence for late-emerging reading disabilities. Since the proportions of students in the reading disabled latent classes were much smaller than those in the higher achieving latent classes, these groups of researchers may not have been able to identify them had they averaged results over the entirety of their respective samples.

Boscardin, Muthén, Francis, and Baker (2008) used growth mixture modeling (GMM) to examine differential developmental patterns of early reading (i.e. phonological

awareness and word recognition) skills in students from kindergarten to second grade. A separate GMM was conducted for each of the two variables. Five classes were identified in each model. These models were then combined to create a total of ten latent classes based on longitudinal profiles of both phonological awareness and word recognition. Overall findings demonstrated depressed phonological awareness skills in kindergarten were directly related to slower development in word reading in first and second grades. Furthermore, the authors demonstrated that there were additional subtypes of readers within the subgroup of poor readers. This study, in particular, highlights the importance of exploring heterogeneity in samples of readers.

While the presence of reading disabilities has been acknowledged for decades, there have remained debates about how best to identify these students. Many traditional methods relied on cutoff scores. The studies reviewed here, however, reveal these students can be identified empirically. While these methods are not able to be conducted with individual students in a classroom, there is the potential for mixture models to more accurately describe learning profiles of students with reading disabilities. This may eventually lead to more nuanced and descriptive profiles of these learners, which may enable earlier identification and service delivery. Hence, utilizing mixture models to understand heterogeneity within samples of readers has the potential to directly impact identification procedures and instructional practices.

Ford, Cabell, Konold, Invernizzi, and Gartland (2013) conducted a study to examine whether heterogeneous patterns of early literacy skills existed in a sample of 2,351 Kindergarten Spanish-speaking ELLs. However, they used traditional cluster analysis instead of a mixture model. While cluster analysis is a technique related to mixture

modeling, there are some advantages associated with mixture modeling that are not available in cluster analysis (see the Methods section for more information). Second, these authors sought to identify if distinct clusters were related to literacy achievement at the end of kindergarten and the beginning of first grade. This study did not examine the SVR specifically, but did include measures of prereading skills that are known to underlie the components of the SVR. The measures were phonological awareness, alphabet knowledge, and orthographic knowledge. The authors found four distinct clusters that were generally ordered in terms of achievement. Specifically, there was a high achieving cluster, a low achieving cluster, and two average achieving clusters. The two average clusters were characterized as having average phonological awareness, but were differentiated by either strength or weakness in alphabet knowledge and orthographic skills. The authors interpreted this as highlighting the roles of these two variables in terms of explaining heterogeneity among their sample. Additionally, phonological awareness may be necessary but not sufficient in terms of early reading development. The two clusters that performed the lowest at the beginning of kindergarten continued to struggle with spelling into the fall of first grade. As one of the first research articles identified that examined heterogeneity in Spanish-speaking ELLs, this study may prove fruitful in encouraging researchers of Spanish-speaking ELLs to take explicit steps to explore heterogeneity in their samples. The findings of this study suggested there are particular subgroups of young Spanish-speaking ELLs who are at increased risk of reading difficulties compared to their peers.

The only identified study that examined heterogeneity in Spanish-speaking ELLs and addressed components of the SVR was conducted by Guzman-Orth (2013). However, her focus was on distinct latent classes of oral language skills rather than all components of the

SVR. Decoding was included as a covariate and reading comprehension was included as a distal outcome. Additionally, this study used LTA to examine if students transitioned between latent classes over the course of two years. Indicators used for the latent profile analyses consisted of measures of receptive and expressive vocabulary as well as syntax knowledge, all measured in Spanish and English. At each of two timepoints, she found four latent classes best represented the heterogeneity in the data. Three classes consisted of ordered bilinguals (high, average, low) and the fourth demonstrated dominance in English relative to their Spanish scores. With respect to the LTA, there was high stability across time with the largest transition occurring for 19% of the low bilingual class who transitioned into the average bilingual class. This aspect of her study clearly demonstrated heterogeneity in Spanish-speaking ELLs' oral language proficiency. That is, oral language does not function in one singular manner in this population of students. This finding lends further validity to the grouping variable used in Kieffer (2008). In terms of later reading comprehension, results were generally intuitive as mean scores were ordered according to the order of the latent classes. The English dominant class performed slightly higher than the average bilingual class, but lower than the high bilingual class. This study extends the findings of Ford et al. (2013) by demonstrating heterogeneity in reading skills (as opposed to precursors to reading) and linking this heterogeneity to a distal outcome. Furthermore, the intuitive findings associated with the distal outcome provide validity to the latent classes and evidence they are not simply statistical artifacts.

There were only two identified studies that explicitly investigated heterogeneity among Spanish-speaking ELLs and only a handful that did so with English monolinguals. Clearly, this is an emerging area that certainly warrants much further research. The present

study takes a step in this direction. Classroom teachers certainly understand they instruct learners of multiple abilities and skill levels. Yet, the vast majority of research concerning the English reading skills of Spanish-speaking ELLs has ignored this variability. If the goal of reading research is to inform teaching practices and educational policy concerning this population, then researchers must acknowledge the heterogeneity that exists within it. The potential consequences of misguided or mistaken research findings may result in a disservice to the very population these researchers are attempting to serve.

The research questions presented in Table 1 are designed to further our knowledge about the functioning of the SVR in subgroups of Spanish-speaking ELLs. By using latent profile analysis to empirically derive classes using multiple indicators, the findings should be more robust compared to studies that use a single grouping variable or define achievement by percentiles or cutoff scores. Subsequently, the regression mixture model will examine whether the SVR functions similarly for all latent classes or if there are important differences that should be accounted for when designing reading instruction for this group of students.

The Present Study

Much of the research literature presented here sought to understand relationships between Spanish-speaking and English-speaking readers and whether component reading processes in one language influence analogous processes in the other language. Extant literature has found that well-developed literacy skills in a person's native language can transfer to her literacy learning in a second language (e.g. August et al., 2006; Cummins, 1979; 2001). Additionally, if continued development of the native-language skills does not occur, these skills may decline even while the second language skills are developed (Laija-

Rodríguez, Ochoa, & Parker, 2006). Thus, one native-Spanish speaker may not always remain fluent in Spanish as she learns English while another may experience continued development in both languages. Given that there may be a developmental relationship between both languages as native-Spanish speakers learn English, the present study seeks to examine heterogeneous language profiles using measures of both Spanish and English word reading and oral language skills as well as measures of Spanish and English reading comprehension. The SVR is used as a theoretical framework informing the examination of heterogeneity among components of reading comprehension and their associations both within and across languages.

Table 1

Research Questions and Measures to Examine Heterogeneity in the Functioning of the SVR among Spanish-Speaking English Language Learners

Question	Measures	Model component
How many latent classes of Spanish-speaking English language learners will emerge based on components of the SVR measured in second grade?	2nd grade English Letter-Word ID 2nd grade Spanish Letter-Word ID 2nd grade English Word Attack 2nd grade Spanish Word Attack 2nd grade English receptive vocabulary (PPVT) 2nd grade Spanish receptive vocabulary (TVIP) 2nd grade English expressive vocabulary (EOWPVT) 2nd grade Spanish expressive vocabulary 2nd grade English Passage Comprehension 2nd grade Spanish Passage Comprehension	Latent profile analysis (emergent classes)
What measures will differentiate the second grade latent classes and how will the latent classes be characterized?	2nd grade English Letter-Word ID 2nd grade Spanish Letter-Word ID 2nd grade English Word Attack 2nd grade Spanish Word Attack 2nd grade English receptive vocabulary (PPVT) 2nd grade Spanish receptive vocabulary (TVIP) 2nd grade English expressive vocabulary (EOWPVT) 2nd grade Spanish expressive vocabulary 2nd grade English Passage Comprehension 2nd grade Spanish Passage Comprehension	Latent profile analysis (class-specific item means)
Does membership in the latent classes moderate the predictive relationships specified in the SVR? Specifically, does the	3rd grade English Letter-Word ID 3rd grade Spanish Letter-Word ID 3rd grade English Word Attack 3rd grade Spanish Word Attack	Regression paths in the regression mixture model

capacity of word-level reading and oral language skills predicting reading comprehension function comparably across latent classes?	3rd grade English receptive vocabulary (PPVT) 3rd grade Spanish receptive vocabulary (TVIP) 3rd grade English expressive vocabulary (EOWPVT) 3rd grade Spanish expressive vocabulary 3rd grade English Passage Comprehension 3rd grade Spanish Passage Comprehension 3rd grade English expressive vocabulary (EOWPVT) 3rd grade Spanish expressive vocabulary 3rd grade English Passage Comprehension 3rd grade Spanish Passage Comprehension	
Do the relationships between word reading, oral language, and reading comprehension vary by language, or will the patterns of relationships be comparable across languages?	3rd grade English Letter-Word ID 3rd grade Spanish Letter-Word ID 3rd grade English Word Attack 3rd grade Spanish Word Attack 3rd grade English receptive vocabulary (PPVT) 3rd grade Spanish receptive vocabulary (TVIP) 3rd grade English expressive vocabulary (EOWPVT) 3rd grade Spanish expressive vocabulary 3rd grade English Passage Comprehension 3rd grade Spanish Passage Comprehension 3rd grade English expressive vocabulary (EOWPVT) 3rd grade Spanish expressive vocabulary 3rd grade English Passage Comprehension 3rd grade Spanish Passage Comprehension	Regression paths in the regression mixture model

Chapter 3

Methods

The data for this study is based on a four-year longitudinal study funded by the U.S. Department of Education, Cognition and Student Learning (USDE R324A090092), Institute of Education Sciences[3].

Participants

Students were recruited from six elementary schools in Southern California. School district data (California English Language Development Test) were used to identify students as Spanish-speaking. Additionally, teachers were asked to identify students for whom they knew English was a second language. Parents were sent a consent form written in both Spanish and English. Children of parents who returned the form indicating consent were included in the study.

Using a cohort sequential design, a cohort of 500 first (n = 163; 32.6%), second (n = 153; 30.6%), and third (n = 184; 36.8%) graders were followed for three years beginning in 2009-2010. This resulted in data from grades one through five at the end of the three years. For this study, only data from grades two and three from the first two years of the study were utilized yielding a sample size of N = 316. Thirty students were missing data on all variables in second grade and 63 students were missing data on all variables in third grade yielding usable data from n = 286 and n = 253 students, respectively. Of the sample of N = 316, 51.4% were female and 48.6% were male and all students identified as Latino/a.

---

[3] This study does not necessarily reflect the views of the U.S. Department of Education or the participating school districts. Special appreciation is given to Danielle Guzamn-Orth, Joseph Rios, Elizabeth Arellano, Nicole Garcia, Alfredo Aviles, Steve Gómez, Paula Aisemberg, Valerie Perry, Loren Albeg, Dennis Sisco-Taylor, Wenson Fung, and School District Laison and Consultant: Erin Bostick Mason for data collection and/or analysis. A special thanks goes to Dr. Cathy Lussier who directed all aspects of this project over the 4 year period.

Procedures

All participants were administered a battery of tests consisting of cognitive and academic assessments by trained undergraduate and graduate research assistants. Tests were administered in Spanish and English with presentation and language order counterbalanced. Testing occurred in a quiet area outside of the classroom. Prior to beginning the test battery, testers built rapport with students by asking them short questions regarding age, grade, and home language. This was done in the same language as the first test battery (i.e. if the English test battery was administered first, English was used to build rapport and vice versa).

Measures

For this study, all measures were administered in both second and third grades.

Letter-Word Identification. The Letter-Word Identification subtests of the Woodcock-Muñoz Language Survey-Revised (Woodcock, Muñoz-Sandoval, & Alverado, 2005) were administered in both Spanish and English. Students were presented with a list of real words and were asked to read the words in order as the list presents increasingly difficult words. The technical manual reports a reliability of .98.

Word Attack. The Word Attack subtest of the Woodcock Reading Mastery Test (Woodcock, 1998) was administered in English. This measure assesses a student's ability to read pseudowords, which are not real words, but are still read by combining the sounds made by the individual letter of which they are composed. The technical manual reports internal reliability to be .88. A corresponding Spanish version was developed and administered using the same administration rules as the English version. To ensure effective

translation, specific rules concerning the English and Spanish letters were followed. For example, *ift* in English was translated to *iyo* in Spanish.

Peabody Picture Vocabulary Test (PPVT). The PPVT (Dunn & Dunn, 1981) was used to assess English receptive vocabulary. A child was presented with four pictures while the tester said a word corresponding to one of the pictures. The child was asked to identify the correct picture. The technical manual reports a reliability of .91.

Test de Vocabulario en Imagenes (TVIP). This test is similar to the PPVT, but is administered in Spanish (Dunn, Lugo, Padilla, & Dunn, 1986). Children must choose one picture that correctly corresponds to a word stated by the tester. The technical manual reports a split-half reliability of .91 - .94.

Expressive One-Word Picture Vocabulary Test Spanish-Bilingual Edition. The Expressive One-Word Picture Vocabulary Test Spanish-Bilingual Edition (Brownell, 2001) was used as a measure of expressive vocabulary and was administered in both Spanish and English. Children were presented with a series of pictures and were asked to name each picture in Spanish followed by English. The pictures were arranged in order of increasing difficulty. If the child achieved a ceiling in one language, that language was discontinued, and the rest of the test was conducted in the other language until a ceiling was also achieved. The manual reports a correlation between item order and item difficulty of .95.

Spanish and English Passage Comprehension. Reading comprehension was measured by the Spanish and English versions of the passage comprehension subtest from the Woodcock-Muñoz Language Survey-Revised (Woodcock, Muñoz-Sandoval, & Alverado, 2005). Item formats change as the test progresses with early items utilizing rebuses (i.e. pictures), in which the examiner says a word or string of words and the student must point to a

corresponding picture. Mid-level items utilize a cloze format in which a picture is provided along with a sentence missing a word. Students must provide a word to fill in the blank. The most difficult items provide short passages using a cloze format without pictures.

Data Analysis Plan

The proposed model in this study consists of multiple components and was implemented in a series of steps to ensure correct model specification. The following subsections describe each of these steps in the order they were conducted. All models were conducted using M*plus* 7.3 (Muthén & Muthén, 1998-2012). Full information maximum likelihood estimation was utilized in all models as the observed variables are continuous. Additionally, this estimator assumes missing data are missing at random (MAR; Little & Rubin, 1990). Furthermore, this estimator allows for item-level missingness. That is, students can be included in the analyses if they have data on at least one of the observed variables.

Grade 2 z-scores. As the measures utilized in this analysis varied substantially in the range of possible scores (e.g. English and Spanish comprehension range from 0 – 33 while PPVT ranges from 0 – 228), all variables were standardized to z-scores using the following equation:

$$z = \frac{(x - \bar{x})}{\sigma} \tag{1}$$

where x is the observed score of a given variable, $\bar{x}$ is the mean of the given variable, and σ is the standard deviation of the given variable. It should be noted the use of the z-score metric was simply to facilitate interpretation and the metrics of the individual tests are not

inherently problematic when conducting a latent profile analysis (see below). The transformation does not change the distance between scores and did not alter the results.

Confirmatory Factor Analysis. This study examines the SVR by utilizing a latent variable framework. Thus, it was necessary to ensure the components of the SVR were first adequately measured by the observed variables. To this end, confirmatory factor analysis (CFA) was employed. Exploratory factor analysis was not considered as a first step because a strong a priori theoretical framework describing the relationships between the observed variables exists (e.g. Gough & Tunmer, 1990; Proctor et al., 2010). In a CFA, individual scores on a particular outcome variable, *y*, are a function of factor loadings, an individual's level on the latent factor, and measurement error. This model is depicted by the following equation (Brown, 2006; Raykov & Marcoulides, 2008):

$$y = \Lambda_y \eta + \varepsilon \tag{2}$$

where y is an observed variable, Λ_y is a factor loading for variable y, η is the factor with a mean of 0, and ε is measurement error.

The SVR was modeled using grade three variables. Specifically, the latent factor *Spanish Word Reading* was measured by Spanish WMLS-R Letter-Word ID and Spanish WMLS-R Word Attack. The latent factor *English Word Reading* was measured by English WMLS-R Letter-Word ID and English WMLS-R Word Attack. The *Spanish Oral Language* factor was measured by the TVIP and the Spanish version of the Expressive One-Word Picture Vocabulary Test. The *English Oral Language* factor was measured by the PPVT and the English version of the Expressive One-Word Picture Vocabulary Test. Spanish reading comprehension and English reading comprehension were not treated as latent factors, but, rather, they were observed measures from the WMLS-R as described above.

Commonly used model fit statistics were examined to judge the adequacy of the CFA. Specifically, these included the chi-square test of model fit, root mean squared error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). Fit indices were interpreted in accord with the guidelines set forth by Hu and Bentler (1999). A non-significant chi-square value was interpreted as indicative of good fit. However, this fit index is known to be sensitive to sample size (Brown, 2006), which necessitates the use of the other fit statistics. RMSEA values less than .08 were considered an indication of adequate fit and values < .05 as an indication of good fit (Brown, 2006). MacCallum, Browne, and Sugawara (1996) also suggest RMSEA values between .08 - .10 may indicated mediocre fit. However, the confidence interval was also evaluated to examine whether the upper bound was inclusive of the aforementioned values. The CFI and TLI statistics were interpreted similarly to each other. Specifically, values >.90 were considered indicative of adequate fit and values > .95 indicative of good fit. Finally, SRMR values of < .08 were considered indicative of adequate fit while values of < .05 indicative of good fit. It is important to emphasize Hu and Bentler's (1999) warned against interpreting any single fit statistic as definitive evidence to support or reject a model. This study examined these fit indices holistically and in tandem with one another to judge the adequacy of the models.

Structural Equation Model. Following the CFA, a SEM was fit in which Spanish reading comprehension was regressed on the *Spanish Word Reading* and *Spanish Oral Language* factors. An analogous model using the corresponding English measures was also conducted. The fit of the models was assessed using the same fit statistics and in the same fashion as the CFA. The regression coefficients were examined for statistical significance, but, at this

stage, were not of primary interest. That is, a non-significant regression coefficient did not preclude this study from progressing as the non-significant coefficient may be a result of averaging the regression coefficients across subgroups (i.e. latent classes) of students. Rather, the regression coefficients specific to individual latent classes (see below) form the basis for this study's research questions. This step of the analysis, then, was to identify whether word reading and oral language skills were predictive of concurrent reading comprehension for the sample as a whole before dividing the sample into multiple latent classes.

Latent Profile Analysis. Next, a latent profile analysis (Gibson, 1959; Vermunt & Magidson, 2002) was conducted using the measures administered in second grade in order to explore potential linguistic heterogeneity among this sample of Spanish-speaking ELLs. In this study, latent profile analysis (LPA) was considered appropriate because it places students into latent classes, or subgroups, based on their patterns of responses to multiple variables. The variables included in this model were intended to differentiate students according to the SVR components.

LPA is a model-based technique that uses a categorical latent variable to characterize the structure of the data. Since this study uses continuous indicators, the data structures being modeled were the means and covariances. Since LPA assumes multiple normal distributions underlying the overall sample distribution, mean scores are class-specific and are used to characterize the latent classes. It follows, then, there are a finite number of latent classes and these are considered mutually exclusive and exhaustive. The LPA equation, as provided by Vermunt and Magidson (2002) is:

$$f(y_i \mid \theta) = \sum_{k=1}^{K} \pi_k f_k(y_i \mid \theta_k) \tag{3}$$

where y_i represents an individual's score on an observed indicator, K is the number of latent classes, π_k is the prior probability of an individual to belonging to latent class k, and θ is a given model parameter. It can be seen an individual's value for a given parameter is a function of the probability of belonging to a particular latent class. As the model is probabilistic, individuals may have non-zero probabilities of belonging to more than one class, but these probabilities must sum to one. The latent class with which an individual has the highest probability of belonging is the class to which the individual is assigned. Probabilities of belonging to other latent classes are then treated as classification error.

While LPA is a clustering technique that may be considered similar to traditional cluster analysis, LPA provides researchers with some advantages over traditional cluster analysis (Vermunt & Magidson, 2002). First, the criterion used to cluster individuals in LPA is less arbitrary than in traditional non-hierarchical cluster analysis, in which a user chooses from a variety of distance measures (e.g. single linkage, average linkage), with each potentially leading to varying cluster patterns. This is because LPA attempts to maximize a log-likelihood function. This leads to a second advantage. Namely, the maximization of the log-likelihood function allows the model to be tested for goodness-of-fit to the data, which is not possible with traditional cluster analysis. Additionally, LPA parameters can be constrained to particular values (or equality across classes) and these constraints can also be tested for their validity. Finally, LPA is flexible in terms of being able to model variables with different, and perhaps, complex distributions as well as different scales. For example, a single LPA model can include variables that are continuous, ordinal, and binary.

A series of LPA models were conducted iteratively beginning with a one-class model, and increasing the number of classes by one in subsequent model runs until non-convergence was achieved, classes appeared to be redundant, or classes did not appear to be substantively meaningful in terms of explaining the heterogeneity in the sample. Each model was then compared to the previous model (with one less class) using a variety of fit indexes as no single fit index has been shown to consistently identify the optimal model (Nylund, Asparouhov, & Muthén, 2007). The model with the greater number of classes was chosen if it was supported by fit statistics and if the latent classes were substantively meaningful. Models were compared using commonly accepted fit statistics in mixture modeling. For the Bayesian Information Criterion (BIC; Schwarz, 1978) and Adjusted BIC (ABIC), lower values indicate a preferred model. Additionally, two information-heuristic measures of fit derived from the BIC were utilized: the Bayes Factor (BF) and correct model probability (cmP; Masyn, 2013). The BF allows for pairwise comparisons between two competing models (a model with $k - 1$ classes and a model with k classes) and calculates a ratio consisting of the probabilities of each of the models being correct. This ratio is assessed according to Jeffery's Scale of Evidence (Wasserman, 2000) where values for the $k - 1$ class model that are between 1 and 3 are considered weak evidence for the $k - 1$ class model, values between 3 and 10 are considered moderate evidence for the $k - 1$ class model, and values greater than 10 are considered strong evidence for the $k - 1$ class model. Additionally, values for the $k - 1$ class model that are below .10 are considered strong evidence for the k class model, values between .10 and .33 are considered moderate evidence for the k class model, and values between .33 and 1.00 are considered weak evidence for the k class model (Wasserman, 2000). The cmP compares a single model to the

entire set of models under consideration. The model with the highest probability is considered the preferred model (Masyn, 2013). In addition to these four indices, this study also utilized two likelihood based indices, the Lo-Mendell-Rubin Likelihood Ratio Test (LMR) and the Bootstrap Likelihood Ratio Test (BLRT). Both of these tests provide a *p*-value comparing a model with *k* classes to the previous model with *k* – 1 classes (Nylund et al., 2007). In each test, a non-significant *p*-value indicates the model with *k* – 1 classes is preferred. Essentially, the additional class (i.e. the model with *k* classes) does not provide a significantly better model fit compared to the model with *k* – 1 classes if the *p*-value is non-significant for the *k* class model. Finally, while not considered a fit statistic, a measure of entropy was utilized. Entropy provides a numeric summary of the accuracy of classification of individuals across all of the latent classes (Ram & Grimm, 2009). Entropy values range from 0 – 1 where higher values are preferred and a value of 1 indicates perfect classification. In this study, models with entropy values of .80 or greater were considered to have strong classification (Ram & Grimm, 2009). Once all models were conducted, the item profile plots for each were visually inspected to ensure the chosen model had a substantively meaningful interpretation.

Regression Mixture Model. Once the optimal number of latent classes was chosen, the final step of this analysis was conducted, which integrated the SEM and the LPA. Specifically, the predictive relationships between the *Word Reading* and *Oral Language* factors and reading comprehension were examined by latent class. This model is commonly referred to as a regression mixture model (Van Horn et al., 2009; Van Horn et al., 2014). Regression mixture models allow researchers to test differential effects of predictors on outcome variables across latent classes. One of the implicit assumptions in these models is

that the relationship between predictors and outcomes is heterogeneous across subjects (Van Horn et al., 2014). That is, research hypotheses should be based on an expectation of such heterogeneity. In the case of this study, the expectation of heterogeneity in the predictive effects of the SVR components was motivated by similar findings in samples of English monolinguals (e.g. Catts, Adlof, & Weismer, 2006; Catts, Hogan, & Fey, 2003). The basic regression equation in this study takes the form:

$$\text{ReadingComprehension}_i = \beta_{0k} + \beta_{1k}(\text{WordReading}) + \beta_{2k}(\text{OralLanguage}) + \varepsilon_{ik} \qquad (4)$$

where β_{0k} is the class-specific average level of reading comprehension when W*ord Reading* and *Oral Language* skills equal 0, β_{1k} is the class-specific weight of the relationship between W*ord Reading* and reading comprehension, β_{2k} is the class-specific weight of the relationship between *Oral Language* and reading comprehension, and ε_{ik} represents class-specific error. This model was applied to each of the latent classes as denoted by the *k* subscript. To identify whether the SVR functions similarly or variably across latent classes, magnitudes and patterns of regression coefficients, as well as statistical significance, were then be compared.

It has been documented the inclusion of auxiliary variables, such as covariates and distal outcomes, can influence the class enumeration process (Asparouhov & Muthén, 2013; Nylund-Gibson, Grimm, Quirk, Furlong, 2014; Vermunt, 2010). That is, class sizes, item probabilities, the number of emergent classes, and the qualitative characteristics of the latent classes found in an unconditional model (i.e. a model with no auxiliary variables) are all subject to change once auxiliary variables are included. This is generally viewed as an unwanted result as researchers want the measurement of the latent classes to be independent of the auxiliary variables (Asparouhov & Muthén, 2013; Nylund-Gibson, Grimm, Quirk,

Furlong, 2014; Vermunt, 2010). In other words, if auxiliary variables are allowed to influence the class enumeration process, the latent classes would be based on the heterogeneity in the indicator variables *in addition to* the heterogeneity in the covariates and/or distal outcomes. As this has been found to be the case with covariates and distal outcomes, similar consequences would be expected when including an SEM into a LPA when creating the regression mixture model. In the context of this study, the aim was to examine whether the SVR in third grade varies according to latent class membership in second grade only. If an approach accounting for the influence of the third grade measures is not utilized, then latent class membership would be based on second *and* third grade measures.

The approach utilized in this study is known as the three-step approach (Nylund-Gibson, Grimm, Quirk, Furlong, 2014; Vermunt, 2010). This approach is currently preferred because it ensures the emergent latent classes are not influenced by auxiliary variables or, in this case, a SEM. This method is enacted in three steps. In this study, the first step was to conduct the series of LPAs to find the optimal number of classes. Second, individuals were assigned to the latent class to which they had the highest probability of belonging. At this step, classification error was explicitly modeled to account for the non-perfect assignment of individuals to latent classes since LPA models are probabilistic. Finally, the SEM was included while holding individuals constant in their class assignment. This allowed the regression coefficients in the third grade SVR to take on class-specific values enabling comparisons. This model was then examined simultaneously in Spanish and English. A diagrammatic representation of the final regression mixture model appears in Figure 1.

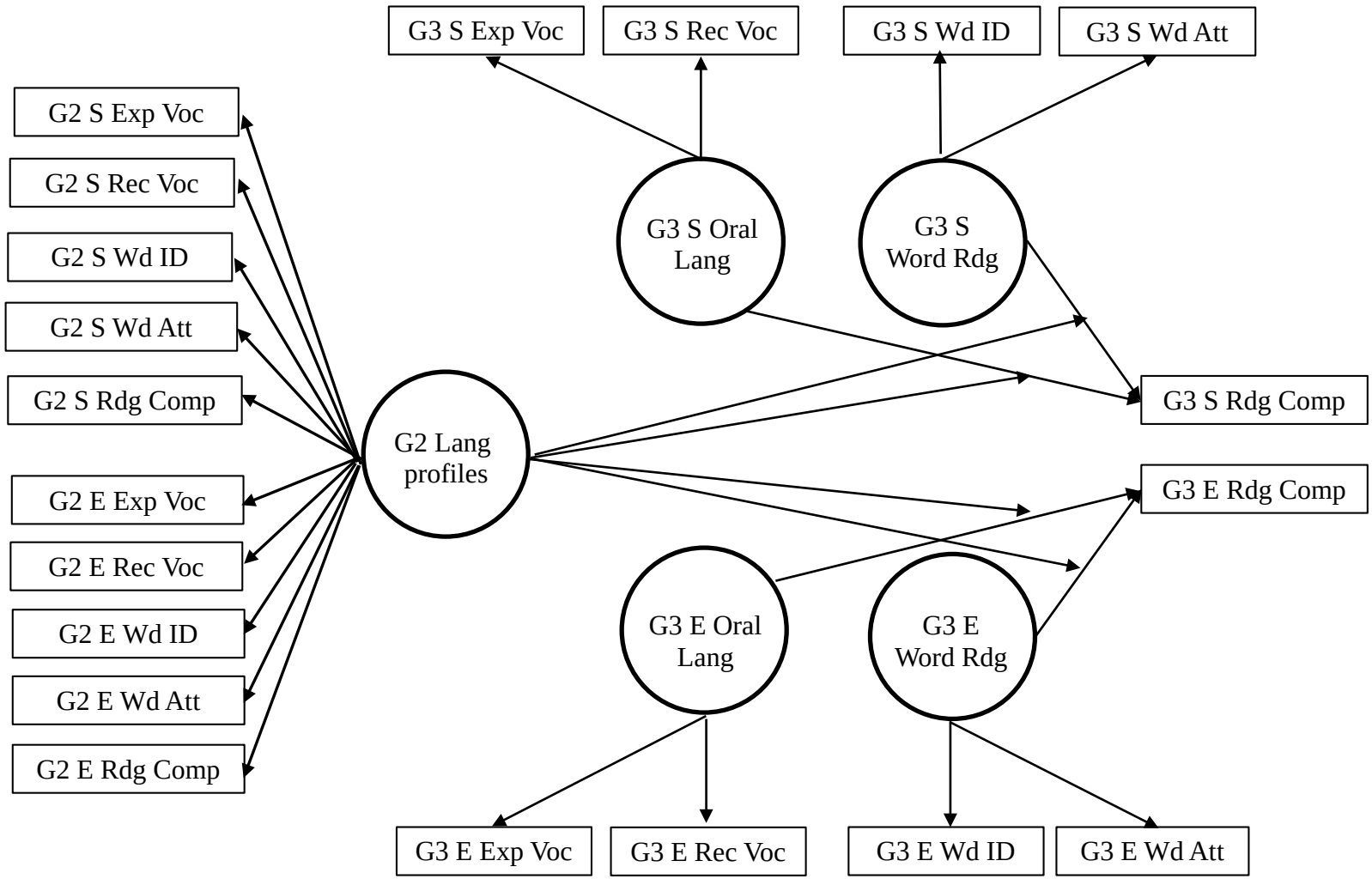

Figure 1. Final regression mixture model with Spanish and English. G2 = grade 2; G3 = grade 3; S = Spanish; E = English; Exp Voc = Expressive Vocabulary; Rec Voc = Receptive Vocabulary; Wd ID = Letter-word Identification; Wd Att = Word Attack; Rdg Comp = Reading Comprehension

Chapter 4

Results

The results are divided into subsections in accordance with each step of the analysis as outlined above. First, results of the descriptive statistics are presented. Next, results of the CFA and SEM are discussed concomitantly as each represents a component of a single model (i.e. measurement and structural parameters, respectively). Third, the LPA results based on second grade language variables are presented. Finally, the results of the regression mixture model, which combines the LPA and SEM, are presented. All analyses were conducted using M*plus* 7.3 (Muthén & Muthén, 1998-2012).

Descriptive Statistics

Means and standard deviations of all observed variables are presented in Table 2. As would be expected, the mean scores for all of the variables increased from second to third grade. However, there was little change in the amount of variation between the two grades. Thus, while students generally improved on these measures, the achievement gap between high and low performers did not become narrower. The mean scores of all assessments were higher on the English versions than the Spanish versions in both grades. This may be a result of English being the language of instruction for this sample. This study does not examine comparisons between assessments as each assessment used a different metric.

Correlations between all variables are presented in Table 3. While the majority of the correlations were statistically significant, some non-significant correlations are notable. As expected, both Spanish receptive and expressive vocabulary measures were unrelated to their English analogs in both grades with one exception. Third grade Spanish receptive

vocabulary was significantly, but weakly (r = .14) correlated with third grade English receptive vocabulary.

Table 2

Descriptive Statistics of All Observed Variables

Variable	*M*	*SD*
Grade 2		
English Letter Word ID	34.69	9.67
Spanish Letter Word ID	26.06	12.22
English Word Attack	14.90	10.41
Spanish Word Attack	10.20	10.63
English PPVT	97.69	17.86
Spanish TVIP	43.16	14.84
English EOWPVT	49.11	11.38
Spanish EOWPVT	26.82	16.26
English Reading Comp	13.59	4.24
Spanish Reading Comp	6.88	3.91
Grade 3		
English Letter Word ID	40.49	9.93
Spanish Letter Word ID	29.78	12.63
English Word Attack	19.28	11.23
Spanish Word Attack	13.15	10.76
English PPVT	109.85	18.16
Spanish TVIP	48.75	13.17
English EOWPVT	57.21	11.15
Spanish EOWPVT	32.39	18.00
English Reading Comp	15.99	4.15
Spanish Reading Comp	8.61	4.30

Note. PPVT = Peabody Picture Vocabulary Test; TVIP = Test de Vocabulario en Imagenes; EOWPVT = Expressive One-Word Picture Vocabulary Test

Table 3

Correlations Between All Observed Variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. E LWID2	-												
2. S LWID2	.54	-											
3. E WDAT2	.82	.56	-										
4. S WDAT2	.67	.84	.69	-									
5. E PPVT2	.40	.19	.34	.23	-								
6. S TVIP2	.13	.39	.13	.33	.05†	-							
7. E EOW2	.47	.17	.43	.25	.68	-.01†	-						
8. S EOW2	.22	.46	.22	.44	.07	.59	.07†	-					
9. E PCOM2	.78	.34	.65	.46	.47	.09†	.55	.13	-				
10. S PCOM2	.51	.79	.49	.80	.16	.45	.20	.52	.33	-			
11. E LWID3	.84	.53	.74	.65	.36	.12	.41	.21	.73	.51	-		
12. S LWID3	.53	.84	.53	.83	.18	.42	.17	.53	.35	.78	.57	-	
13. E WDATT3	.80	.53	.75	.65	.32	.16	.41	.25	.64	.55	.81	.56	-
14. S WDATT3	.67	.76	.68	.85	.23	.30	.25	.42	.50	.73	.73	.83	.74
15. E PPVT3	.37	.14	.29	.18	.66	.06†	.58	.12	.45	.15	.41	.22	.38
16. S TVIP3	.22	.45	.15	.41	.10	.62	.02†	.65	.14	.48	.25	.53	.25
17. E EOW3	.39	.11†	.31	.16	.59	-.02†	.63	.04†	.43	.13	.37	.16	.36
18. S EOW3	.27	.48	.27	.46	-.01†	.58	.02†	.74	.16	.53	.28	.55	.30
19. E PCOMP3	.71	.38	.56	.46	.40	.13	.44	.21	.70	.39	.74	.41	.66
20. S PCOMP3	.54	.70	.49	.74	.17	.47	.17	.58	.36	.76	.56	.79	.59

Table 3 (cont.)

	14	15	16	17	18	19	20
14. S WDATT3	-						
15. E PPVT3	.28	-					
16. S TVIP3	.41	.14	-				
17. E EOW3	.22	.65	.01†	-			
18. S EOW3	.49	.08†	.64	.03†	-		
19. E PCOMP3	.53	.49	.25	.47	.26	-	
20. S PCOMP3	.76	.25	.56	.19	.63	.47	-

Note. E = English; S = Spanish; 2 = Second grade; 3 = Third grade; LWID = Letter-Word Identification; WDATT = Word Attack; PPVT = Peabody Picture Vocabulary Test; TVIP = Test de Vocabulario en Imagenes; EOW = Expressive One-Word Picture Vocabulary Test Bilingual Edition; PCOMP = Passage Comprehension. All correlations significant at $p < .05$ except †*ns*

This makes intuitive sense as a student's oral proficiency in one language would not be expected to indicate oral proficiency in another language. All of the word reading measures were moderately to strongly correlated across both languages. This finding is not surprising given both languages use similar orthographies. Correlations between word reading and oral language measures were slightly more nuanced, but followed an expected pattern. Specifically, while correlations were statistically significant, the magnitude of a particular correlation was language dependent. For example, English oral language measures were more strongly related to English word reading measures than the corresponding Spanish word reading measures. Similarly, Spanish oral language measures were more strongly related to Spanish word reading measures than the English versions. Moreover, these patterns were stable across both second and third grade. Similar results were also found regarding passage comprehension in both languages. Within-language correlations between passage comprehension and the word reading and oral language measures were stronger than those across languages. Finally, the two word reading measures exhibited stronger relationships with passage comprehension than did the oral language measures in both languages and both grade levels. This finding is in line with previous research showing a stronger relationship between word reading and reading comprehension during the early elementary years (e.g. Adlof et al., 2006; Kendeou et al., 2009; Tunmer & Chapman, 2012).

Confirmatory Factor Analyses

Separate CFAs were conducted for each grade to examine whether the observed variables adequately measured their respective SVR components, which were treated as latent factors. Though the second grade SVR factors were not treated as latent factors in the latent profile analysis or regression mixture model, a second grade CFA was conducted in

order to assess whether these observed variables could sufficiently reflect the SVR components as theorized. That is, this study sought to explore heterogeneous subgroups based on the SVR components in second grade. Thus, it was deemed necessary to first ensure the variables used to empirically identify these subgroups were indeed reflective of the SVR components. The fit statistics for each model can be seen in Table 4.

Table 4

Fit Statistics of the CFA Models

Grade	χ^2	*df*	*p*	RMSEA (90 %CI)	CFI	TLI	SRMR
2	23.985	14	.046	.050 (.007 - .083)	0.992	0.985	.026
3	53.024	14	<.001	.105 (.076 - .136)	0.968	0.936	.038

Note. χ^2 = chi-square test of model fit; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root-mean square error of approximation; SRMR = standardized root mean square residual.

In grade 2, the model fit the data well according to all fit indices. In grade 3, however, the chi-square and RMSEA indices suggested poor fit to the data while the CFI, TLI, and SRMR suggested good fit. The chi-square test statistic is known to be influenced by sample size (e.g. Brown, 2006) and Chen, Curran, Bollen, Kirby, and Paxton (2008) caution the RMSEA test statistic performs variably under different modelling conditions. Findings from Chen et al. (2008) led these authors to conclude use of a single RMSEA cutoff score (e.g. .05 or .10) should not be recommended nor should the RMSEA test statistic be used as a single indicator of model support or rejection. Rather, the model should be assessed globally using multiple fit indices. Moreover, Hu and Bentler (1999) noted the RMSEA tended to overreject models when sample sizes were relatively small (i.e. $N < 250$), which is the case in this study.

Factor loadings and factor correlations were examined in addition to fit indices to assess both models. Factor loadings are presented in Table 5 and factor correlations are presented in Table 6. All indicators loaded strongly on their respective factors (ranging from .70 - .99).

Table 5

Factor Loadings for the CFA Models in Grades 2 and 3

	Factor			
Indicator	E Word Reading	E Oral Language	S Word Reading	S Oral Language
E Letter Word ID	.90/.90			
E Word Attack	.90/.90			
E PPVT		.74/.83		
E EOWPVT		.92/.78		
S Letter Word ID			.85/.85	
S Word Attack			.99/.98	
S TVIP				.70/.75
S EOWPVT				.86/.86

Note. Grade 2 loadings are before the backslash and Grade 3 loadings are after the backslash. E = English; S = Spanish; PPVT = Peabody Picture Vocabulary Test; TVIP = Test de Vocabulario en Imagenes; EOWPVT = Expressive One-Word Picture Vocabulary Test Bilingual Edition. All loadings significant at $p < .001$.

Table 6

Factor Correlations for Grades 2 and 3

Factor	1	2	3	4
1. E Word Reading	-	.54	.76	.27
2. E Oral Language	.52	-	.28	.08†
3. S Word Reading	.82	.32	-	.51
4. S Oral Language	.40	.14†	.61	-

Note. Grade 2 correlations are presented above the diagonal and grade 3 correlations are presented below the diagonal. All correlations significant at $p < .001$ except †*ns*

The magnitudes of the factor loadings were generally consistent across grades with the greatest change occurring with English expressive vocabulary. This loading decreased by .14, but the third grade estimate of .78 is still considered a strong loading. Thus, the items appear to adequately measure the latent factors in both grades. Regarding factor correlations, a similar pattern was found in both grades. Most factors were well distinguished as evidenced by low to moderate correlations between factors. The one exception occurred between the English and Spanish *Word Reading* factors, which were strongly related in both grades. This may be statistical evidence that these two factors should be collapsed into one though from a theoretical standpoint, this study views these factors as distinct. In the interest of being conservative and thorough, a subsequent model was conducted for each grade in which all Spanish and English word reading items were specified to load onto a single *Word Reading* factor. This modification resulted in substantially worse fit statistics in grade 2 ($\chi^2_{(17)}$ = 229.171; RMSEA = .209; CFI = .836; TLI = .731; SRMR = .085) and grade 3 ($\chi^2_{(17)}$ = 186.140; RMSEA = .198; CFI = .862; TLI = .773; SRMR = .073). Moreover, chi-square difference testing revealed the modified models were statistically significantly worse in grade 2 ($\Delta\chi^2_{(3)}$ = 205.186) and grade 3 ($\Delta\chi^2_{(3)}$ = 133.116). Therefore, the original model in which *English Word Reading* and *Spanish Word Reading* were specified as distinct factors was retained.

Even though the CFI, TLI, and SRMR values supported the third grade model, and even though there are noted limitations with the chi-square and RMSEA tests, sources of local misfit were explored via modification indices. The only modification that converged, resulted in admissible factor loadings, and showed an improvement in model fit was allowing the indicator Spanish word attack to load on both the *English Word Reading* and

Spanish Word Reading factors simultaneously, typically known as a crossloading[4]. This resulted in the following fit statistics: $\chi^2_{(13)} = 18.532$, $p = .138$; RMSEA = .041 (90% CI = .000 - .080); CFI = .995; TLI = .990; SRMR = .020. While this model shows excellent statistical fit, it is difficult to consider this model theoretically viable. First and foremost, allowing a Spanish language variable to load onto an English language factor would suggest that a student's English word reading abilities are a cause of her ability to decode pseudowords in Spanish. Though the two languages share similar orthographies, they are indeed distinct. Second, the fact that the variable that crossloads is a measure of Spanish pseudowords (i.e. not real words) may suggest this is a statistical artifact. This may be because the task is based on combining individual sounds that follow a regular pattern rather than reading whole words that may contain irregular letter combinations and sounds. That is, the task, as administered in third grade, may be tapping more strongly into an underlying ability to connect individual sounds regardless of the language of presentation. Third, only Spanish word attack was able to crossload onto the English and Spanish factors. Spanish letter-word identification did not crossload and led to nonconvergence. It would be reasonable to expect that if one Spanish word reading indicator crossloads, the other would do so as well. However, since Spanish letter-word identification is a task composed of real words, this may strengthen the argument the crossloading is a statistical artifact arising from the pseudowords on Spanish word attack. Additionally, it was shown above that allowing both Spanish letter-word identification and Spanish word attack to load onto a single factor with the English analogs resulted in a significantly worse model. Finally, the loadings of

[4] A subsequent model was tested in which Spanish letter-word identification was allowed to crossload instead of Spanish word attack. This model resulted in a standardized factor loading greater than 1.0 for Spanish letter-word identification, which was interpreted as evidence against this model.

Spanish word attack on the *English Word Reading* (.43) and *Spanish Word Reading* (.59) factors decrease considerably compared to when it loads solely on the *Spanish Word Reading* factor (.98).

For both these reasons and the evidence for the original third grade model provided by CFI, TLI, and SRMR values as well as factor loadings and factor correlations that were consistent with the second grade model, the original third grade factor model was retained. However, it may be worth further speculation as to why the crossloading was only identified in the third grade model and not the second grade model. It may be possible that as students transitioned from second to third grade, their growth in knowledge of individual letter sounds became less language-dependent as Spanish and English share similar orthographies. Knowledge of individual sounds would not necessarily translate into reading words as a whole or, perhaps more pertinent, words with irregularities, thereby explaining why there was no crossloading for the letter-word identification measures. Furthermore, the variable loaded higher on the *Spanish Word Reading* factor than the *English Word Reading* factor indicating it was indeed more strongly related to Spanish. However, as students received a year of English instruction between assessments, there may have been greater growth in English reading skills than Spanish reading skills (e.g. Laija-Rodríguez et al., 2006). Thus, English decoding may have become dominant, thereby being able to explain decoding skills in two languages that share similar letter sounds and pronunciations.

SVR Structural Equation Model

After confirming the second and third grade measurement models, the structural parameters of the SVR were tested by regressing observed third grade English passage comprehension on the third grade *English Word Reading* and *English Oral Language* factors

and observed third grade Spanish passage comprehension on the third grade *Spanish Word Reading* and *Spanish Oral Language* factors. The same fit statistics used to assess the CFAs were also used to judge the fit of the SEM. The resulting fit statistics were as follows: $\chi^2_{(26)} = 92.149$, $p < .001$; RMSEA = .100 (90% CI = .079 - .123); CFI = .963; TLI = .936; SRMR = .035. Again, the RMSEA value was not ideal and a subsequent SEM allowing Spanish word attack to crossload on both *Word Reading* factors was conducted. Fit statistics were again improved ($\chi^2_{(25)} = 58.096$, $p < .001$; RMSEA = .072 (90% CI = .048 - .097); CFI = .982; TLI = .967; SRMR = .029), but for the reasons mentioned above, the original model was retained. The regression parameters of the original SEM can be seen in Figure 2. The factor correlations are not presented in Figure 2 as they were nearly identical to those in the grade 3 CFA model.

In both languages, oral language and word reading skills significantly predicted passage comprehension. However, word reading was consistently the stronger predictor. This aligns with previous research that has found word reading to be the stronger predictor in early elementary grades. While third grade may be considered a transition period between early and later elementary grades, it may be this sample of Spanish-speaking ELLs had English word reading skills comparable to those of younger native English-speaking students.

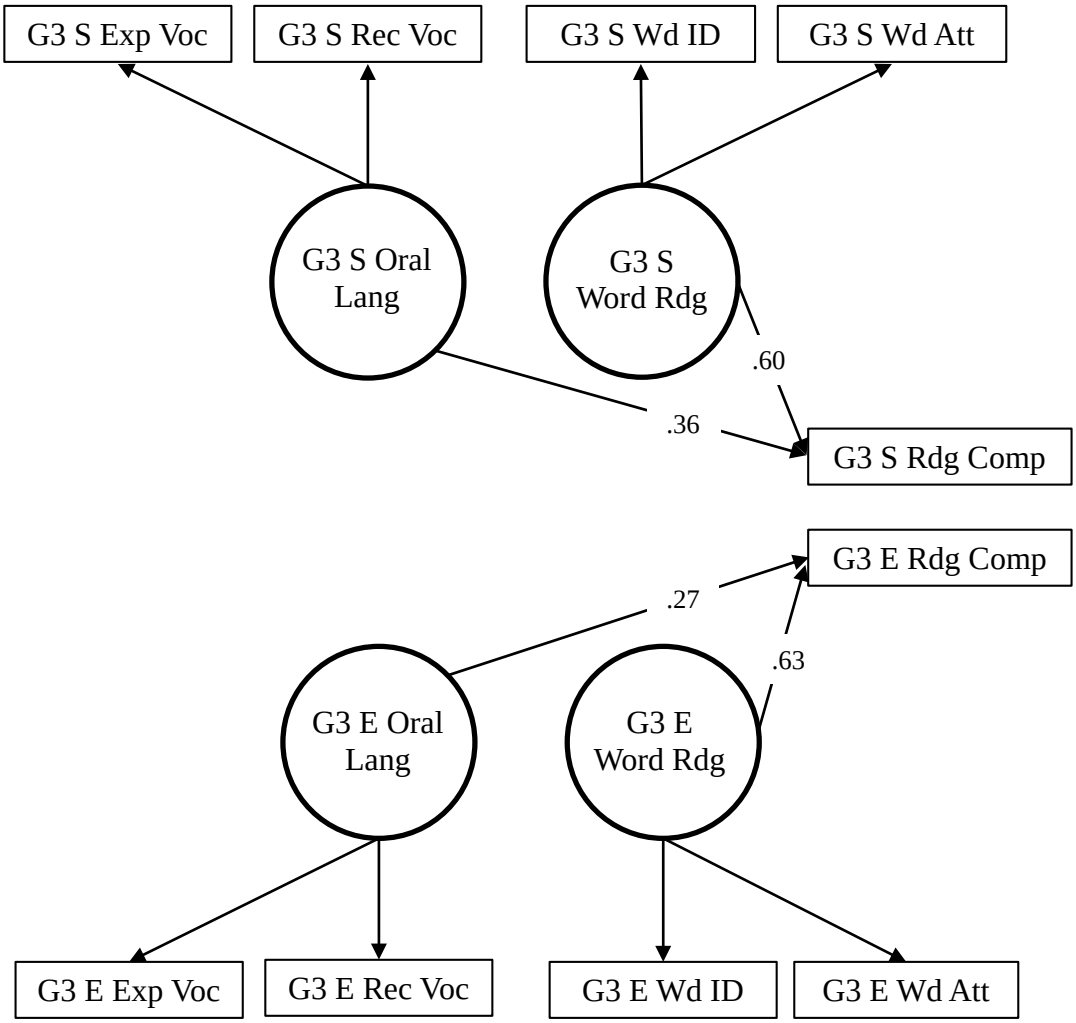

Figure 2. Grade 3 SVR structural equation model. G2 = grade 2; G3 = grade 3; S = Spanish; E = English; Exp Voc = Expressive Vocabulary; Rec Voc = Receptive Vocabulary; Wd ID = Letter-word Identification; Wd Att = Word Attack; Rdg Comp = Reading Comprehension. All regressions significant at $p < .001$.

Latent Profile Analysis

A series of latent profile analyses was conducted to explore heterogeneous language profiles within the larger sample of Spanish-speaking ELLs. Profiles were based on second grade variables that represented the components of the SVR. All second grade variables were converted to z-scores as they utilized substantially different scales. The z-scores did not affect the class enumeration and were used simply to foster interpretation of the profile

plots. The process began with a 1-class model then increasing the number of classes by one in each subsequent iteration until non-convergence occurred. This resulted in six models and the fit statistics are presented in Table 7.

The BIC and ABIC never reached minimum values and there was not a clear indication of either leveling off, so these were not used to inform the model selection. Similarly, the BF never reached a value greater than 1.0 and the cmP did not reach a probability that would be suggestive of a preferred model. The BLRT never became non-significant. However, the LMR became non-significant beginning with the 4-class model, which is interpreted as indicating the 3-class model as preferred. Essentially, adding a fourth class did not significantly improve the model's representation of the data.

Since the LMR was the only fit index to point to the 3-class model, item profile plots

Table 7

Fit Statistics of the LPA models in Second Grade

Number of classes	LL	BIC	ABIC	LMR *p*-value	BLRT *p*-value	BF	cmP	Entropy
1	-4055.61	8224.34	8160.92	-	-	-	-	
2	-3573.97	7323.28	7224.98	<.001	<.001	<.001	<.001	0.95
3	-3370.00	6977.55	6844.36	**0.019**	<.001	<.001	<.001	0.92
4	-3258.48	6816.73	6648.67	0.075	<.001	<.001	<.001	0.93
5	-3182.45	6726.87	6523.92	0.377	<.001	-	-	0.91
6	Model did not converge							

Note. Bold values indicate the preferred model for the given index. LL = Log-likelihood; BIC = Bayesian Information Criterion; ABIC = Adjusted Bayesian Information Criterion; LMR = Lo-Mendell-Rubin Likelihood Ratio Test; BLRT = Bootstrapped Likelihood Ratio Test; BF = Bayes Factor.

were then examined for both the 3-class and 4-class models. This allowed for substantive interpretation to inform the model selection process rather than relying solely on fit statistics. Examination of the 4-class plot revealed the fourth class was redundant with

classes two and three on half of the ten items. The 3-class model contained ordered classes, but demonstrated intuitively meaningful patterns of language profiles and, thus, was chosen as the preferred model informed by both fit statistics and substantive reasoning.

Figure 3 presents the item profile plot that was used to label and interpret the latent classes in the 3-class model. The class at the top of the plot demarcated by a solid line with triangle markers scored highest on all measures except English reading comprehension. Additionally, students in this class consistently scored higher on the Spanish measures than the equivalent English measures. Thus, this class was labeled *High/Spanish Dominant* and consisted of 19.9% of the sample. The middle class demarcated with a dashed line and square markers has scores that hovered near 0 (i.e. average) relative to the other classes. Additionally, students in this class consistently scored higher on the English measures than the Spanish measures. Thus, this class was labeled *Average/English Dominant* and consisted of 35.3% of the sample. Finally, students in the class at the bottom of the plot consistently scored lower on all measures with no readily discernible pattern in terms of relative language proficiency. This class was labeled *Low* and consisted of 44.8% of the sample. It may be worth noting that while the *High/Spanish Dominant* class tended to achieve more highly than the *Average/English Dominant* class, both classes scored fairly similar on the English vocabulary measures and nearly identical on the English reading comprehension measure. Indeed, there was only a difference of .01 z-score units between the two classes on English reading comprehension. The entropy value for the 3-classs model was .92, which is considered a high value (Ram & Grimm, 2009).

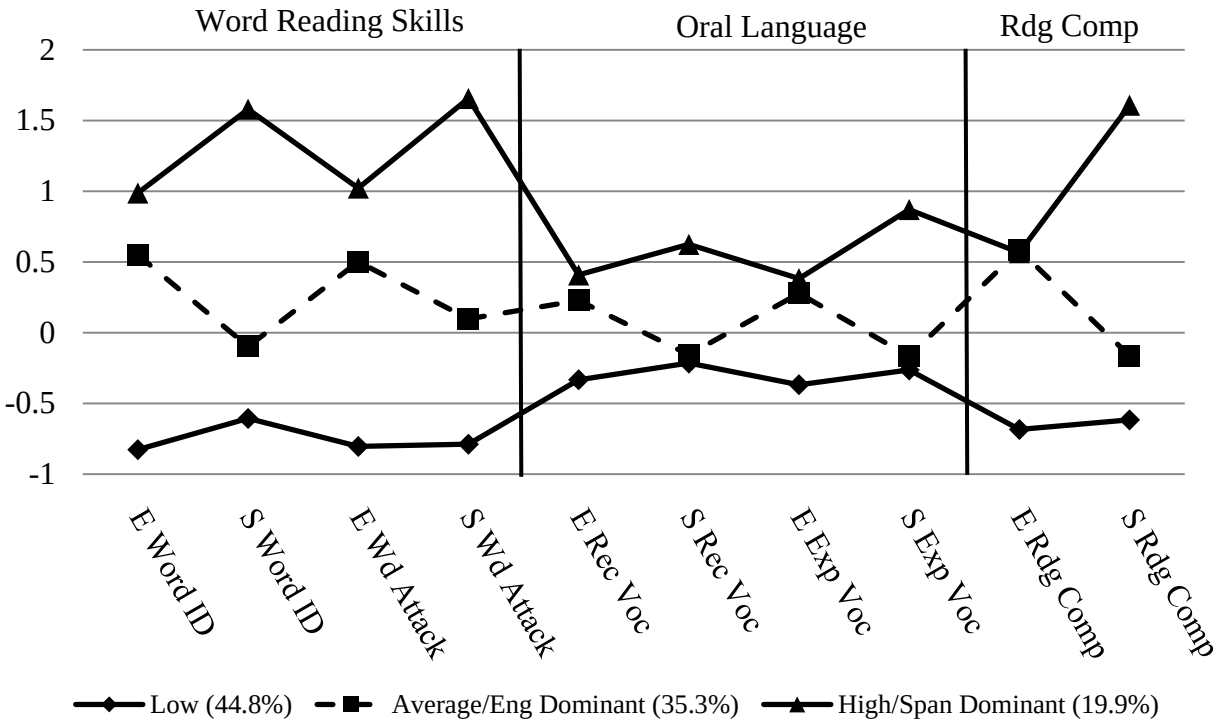

Figure 3. Item-profile plot of second grade language profiles based on the SVR. E = English; S = Spanish; Word ID = Letter-word identification; Wd Attack = Word attack; Rec Voc = Receptive vocabulary; Exp Voc = Expressive vocabulary; Rdg Comp = Reading comprehension.

Regression Mixture Model

After choosing the preferred unconditional (i.e. without auxiliary variables such as the third grade SVR variables) LPA model, the SVR SEM was included in the model using the three-step approach described in the previous chapter. This approach was utilized to avoid shifts in class enumeration that could have potentially occurred when including the SVR SEM. In this modeling context, each latent class was allowed to have its own set of coefficients for the regressions between the Spanish and English *Word Reading* and *Oral Language* factors and Spanish and English reading comprehension measures. Furthermore, this model allowed for class-specific correlations among the latent factors as well as class-

specific factor means. This facilitated an examination of differing patterns of relationships among the latent factors across classes as well as an empirical comparison of Spanish and English *Word Reading* and *Oral Language* skills. Results from the regression mixture model are presented in Table 8.

Class-specific regressions between SVR components. The top panel of Table 8 displays the class-specific regression coefficients of third-grade Spanish and English reading comprehension on third-grade Spanish and English *Word Reading* and *Oral Language* factors. This component of the analysis permitted an examination of whether the SVR functions similarly across the latent classes based on relative language skills in Spanish and English.

For the *Low* class, *Word Reading* and *Oral Language* were both significant predictors of reading comprehension in both languages. However, *Word Reading* was a considerably stronger predictor in both languages, particularly in English. Thus, for this group of students, reading comprehension in both languages was more strongly related to their ability to read individual words rather than their oral language proficiency. The *Average/English Dominant* class displayed the opposite pattern. For this group, *Word Reading* did not significantly predict reading comprehension in either language. Rather, *Oral Language* was the statistically significant predictor across languages. Interestingly, this group's second grade profile demonstrated greater strength in English measures relative to the Spanish measures, but the relationship between *Oral Language* and reading comprehension was greater in Spanish (β = .852) than in English (β = .417). Finally, for the *High/Spanish Dominant* class, a third distinct pattern was found. For this group, *Word Reading* was a significant predictor of reading comprehension, but only in English.

Table 8

Class-specific Parameter Estimates from the Regression Mixture Model

	2nd grade language profile		
3rd grade parameter	Low (44.8%)	Avg/Eng Dominant (35.3%)	High/Span Dominant (19.9%)
Regressions			
E Rdg Comp on E Word Rdg	.792**	.247	.466**
E Rdg Comp on E Oral Lang	.180*	.417**	.252
S Rdg Comp on S Word Rdg	.499**	.054	.151
S Rdg Comp on S Oral Lang	.219*	.852**	.799**
Correlations			
E Word Rdg with E Oral Lang	.337**	.416*	.544**
S Word Rdg with S Oral Lang	.564***	.552**	.532***
E Word Rdg with S Word Rdg	.353	.758***	.100
E Word Rdg with S Oral Lang	.124	.136	-.032
E Oral Lang with S Word Rdg	.025	.021	.182
E Oral Lang with S Oral Lang	-.076	-.007	.453*
E Rdg Comp with S Rdg Comp	.172	.203	.338
Factor Means			
E Word Rdg	-4.276***	-1.083***	0.000
S Word Rdg	-6.049***	-3.580***	0.000
E Oral Lang	-0.867***	0.025	0.000
S Oral Lang	-1.681***	-1.195***	0.000

Note. Factor means for the *High/Spanish Dominant* class are set to 0 by default and are used as a reference for the other two classes. E = English; S = Spanish; Rdg Comp = Reading Comprehension; Word Rdg = Word Reading; Oral Lang = Oral Language.
*$p < .05$, **$p < .01$, ***$p < .001$

Conversely, *Oral Language* was the only significant predictor of reading comprehension in Spanish. This pattern may be reflective of this group's greater proficiency in Spanish than English. As oral language proficiency has historically been more strongly associated with reading comprehension in readers with more advanced skills, the strength of the relationship between Spanish *Oral Language* and reading comprehension may be a result of their greater

Spanish proficiency. Similarly, in English, *Word Reading* may be the stronger predictor of reading comprehension due to their relatively weaker English skills.

Class-specific correlations among SVR components. Having found differential predictive relationships across classes, it was next of interest to assess whether the SVR components showed differential associations across languages and latent classes. The second panel of Table 8 presents these results.

The within-language associations of *Word Reading* with *Oral Language* were consistent across classes though the estimates were generally in the moderate range. Specifically, for all three latent classes, *Word Reading* and *Oral Language* were positively correlated for each language. This was not the case, however, across languages. That is, English *Word Reading* was not significantly correlated with Spanish *Oral Language* and vice versa for any of the latent classes. This finding makes intuitive sense as skills in reading words in English would not necessarily be expected to influence a student's development of Spanish vocabulary. English and Spanish *Word Reading* were not correlated for either the *Low* or *High/Spanish Dominant* classes. However, they were strongly correlated ($r = .758$, $p < .001$) for the *Average/English Dominant* class. Furthermore, English and Spanish *Oral Language* were not correlated for either the *Low* or *Average/English Dominant* classes, but they were significantly correlated ($r = .453$, $p < .05$) for the *High/Spanish Dominant* class. Taken together, these findings might reflect the overall reading abilities of these two classes. For instance, the *High/Spanish Dominant* class scored higher on all measures except English reading comprehension compared to the *Average/English Dominant* class. Thus, the *High/Spanish Dominant* class likely possessed stronger reading skills overall compared to the latter group. Since increased oral language proficiency is associated with stronger

reading comprehension in more advanced readers, this may explain the correlation of the *Oral Language* factors across languages for the *High/Spanish Dominant* group. On the other hand, word-level reading skills are associated with reading comprehension in less advanced readers such as those in the *Average/English Dominant* group. This may explain the strong correlation across languages in *Word Reading* for these students. Finally, the Spanish and English reading comprehension measures were not significantly correlated for any of the latent classes.

Class-specific factor means. The final step in this component of the analysis was to compare class-specific means of Spanish and English *Word Reading* and *Oral Language* factors for significant differences. When specifying latent factors for multiple groups, it is necessary to fix the factor means for one group to zero in order to estimate the factor means for the other groups. In this study, the factor means for the *High/Spanish Dominant* class were set to zero. The factor means of the other classes were then interpreted in reference to the *High/Spanish Dominant* class. A statistically significant factor mean value indicates that factor mean is significantly different from zero. In the present study, if a latent class' factor mean was statistically different from zero, then it was interpreted as being significantly different from the factor mean of the *High/Spanish Dominant* class. All factor means were standardized to allow for comparisons. The bottom panel of Table 8 presents the results of these comparisons.

Comparing the *Low* class to the *High/Spanish Dominant* class, all four factor means were significantly lower. Of these, the largest difference occurred with Spanish *Word Reading* while the smallest difference occurred with English *Oral Language*. This latter difference may be a result of English being the language of instruction, which could have

enabled the *Low* class to perform more similarly to their higher-achieving counterparts in terms of English vocabulary. Lower and higher achieving students within a classroom would often be exposed to the same academic vocabulary. Thus, even though the *Low* class demonstrated lower achievement, simple exposure to similar language as their peers may have helped them perform more similarly to the *High/Spanish Dominant* class on the vocabulary measures in this study. On the other hand, word-level reading is a skill that must be explicitly taught and is not acquired through mere exposure to everyday language. This may explain why the *Low* class' factor means on the *Word Reading* factors in both languages are considerably lower compared to the *Oral Language* factor means.

Comparing the *Average/English Dominant* class to the *High/Spanish Dominant* class, three of the factor means were significantly lower. As with the *Low* class, the lowest factor mean for the *Average/English Dominant* class was the Spanish *Word Reading* factor. Interestingly, the Spanish *Oral Language* factor mean was the next lowest mean following the Spanish *Word Reading* factor mean for the *Average/English Dominant* class. This may be reflective of this class' dominant English skills compared to their Spanish skills. There was a non-significant difference between the *Average/English Dominant* and *High/Spanish Dominant* classes on the English *Oral Language* factor. This may again be a result of exposure to English as the language of instruction. Indeed, Figure 3 shows both classes scored similarly on the English vocabulary measures, which is likely contributing to the similarity in factor means.

Chapter 5

Discussion

The primary aims of this study were to empirically identify latent subgroups of Spanish-speaking ELLs based on their word reading and oral language skills in both Spanish and English and use these groupings to subsequently test the SVR across subgroups and languages. Results from the current analyses may shed light on conflicting results found in previous studies (e.g. Kieffer & Vukovich, 2012; Lesaux et al., 2010; Mancilla-Martinez et al., 2009). There are a number of findings from this study that merit attention.

Full Sample SVR Structural Equation Model

Results from the SVR SEM using the full sample were consistent with previous research findings regarding Spanish-speaking ELLs and LM learners and the SVR (Gottardo & Mueller, 2009; Kieffer & Vukovich, 2012; Proctor et al., 2005; Proctor et al., 2010). Within each language, both the *Word Reading* and *Oral Language* factors significantly predicted reading comprehension. Proctor et al. (2010) specifically tested a model in Spanish in addition to English (whereas the other studies focused on English) and found that Spanish oral language and Spanish alphabetic knowledge were key components of Spanish reading comprehension. The current study extends these findings to both English and Spanish simultaneously with Spanish-speaking ELLs. However, Proctor et al. (2010) conducted their study with fourth graders and found oral language was a stronger predictor than alphabetic knowledge. Contrary to these findings, the current study identified *Word Reading* as a stronger predictor of reading comprehension and this was true in both languages. There may be a few reasons for the differing results. Proctor et al. (2010) measured alphabetic knowledge using a single pseudoword reading task. This study utilized

a measure of letter-word identification in addition to pseudoword reading to create the *Word Reading* latent factor. Thus, the different ways in which word reading and alphabetic knowledge were operationalized may have contributed to conflicting findings. Alternatively, the sample in the Proctor et al. (2010) study had received or was receiving literacy instruction in Spanish, which was not the case with the sample in this study. Thus, the sample in their study may have developed Spanish academic language skills that may have been stronger, yielding a Spanish reading profile that mirrored more advanced English monolinguals whereas the profile in the current sample aligned better with less advanced English monolinguals. Third, while the samples in both studies were close to each other in terms of age, they were not identical. The students in this study were one year younger (i.e. third grade versus fourth grade), so they may not have developed comparable oral language skills in either English or Spanish such that the *Oral Language* factor would have been a stronger predictor than *Word Reading*. Regardless of these differences, the results found here support previous research identifying the SVR as a viable model with Spanish-speaking ELLs.

Another interesting finding regards the fit of the measurement and structural models in second and third grades. Both the measurement and structural models fit the data very well in second grade, but fit declined in third grade, even though overall fit was still acceptable. This may be a result of the dynamic nature of the SVR, especially given that fit improved when Spanish word attack was allowed to crossload on both Spanish and English *Word Reading* factors. Perhaps as students advanced from second to third grade, their knowledge of letter-sound correspondences developed to the point where they were able to generalize across languages. This may be especially true given that Spanish and English are

both typologically and orthographically similar. Thus, the amount of variation in the word attack measures was similar for both languages, leading to Spanish word attack being able to crossload in third grade whereas it distinctly loaded on the Spanish *Word Reading* factor in second grade. Furthermore, in third grade, both *Oral Language* factors were well-defined by their respective indicators, which suggests the English and Spanish vocabulary measures were less related across languages than the *Word Reading* measures. This is further confirmed in Table 3 in which the correlations between *Word Reading* indicators are considerably greater in magnitude compared to the *Oral Language* indicators when examined across languages.

Even though the *Word Reading* factors were stronger predictors of reading comprehension in both languages, this finding may represent the beginning of a transition from relying on word-level skills to utilizing oral language skills for reading comprehension. Students are generally expected to have a relatively firm grasp of word-level reading skills by third grade as they transition from "learning to read" to "reading to learn." However, students were assessed at the beginning of third grade in this study, so they may not have received enough English instruction for *Oral Language* to become the stronger predictor of reading comprehension. Alternatively, for this particular sample as a whole, *Word Reading* may have been the stronger predictor because, as ELLs, their English oral vocabulary may not have developed to a point comparable to typically developing English monolingual peers.

While both arguments may be plausible, the theme of this study is to explicitly consider heterogeneity within the sample and its effects on the empirical findings. As such, the results from the regression mixture model suggest the above arguments are not true for

all of the students in this study. Rather, they may be true for many, and possibly a majority, of the students, but patterns in the data – especially regarding the *Average/English Dominant* class – suggest such broad generalizations would ignore the academic performances of a sizable proportion of the students. Heterogeneity within the sample is discussed in further detail below.

Heterogeneous Language Profiles

Regarding the first and second research questions, the second-grade language profiles demonstrated there were indeed three discernible latent classes of Spanish-speaking ELLs based on relative language and reading skills in both Spanish and English using the SVR components as a theoretical framework. The vast majority of the extant research literature has ignored the potential for linguistic heterogeneity within this population (for exceptions, see Ford et al., 2013; Guzman-Orth, 2013; Kieffer, 2008). Moreover, the three latent classes identified in this analysis were well-differentiated from each other as evidenced by the LPA's high entropy value. While this study is exploratory, the entropy value may suggest the emergence of the latent classes is not simply a statistical artifact, and they (LIKELY??) represent qualitatively different types(SUBGROUPS??) of Spanish-speaking ELLs.

Teachers and other professionals who work with Spanish-speaking ELLs understand linguistic differences exist among their students. This is also recognized at the district, state, and national levels as assessments are routinely administered with the explicit goal of categorizing ELLs into heterogeneous groups based on English language proficiency. While such heterogeneity has long been recognized, this study is one of the first to empirically identify subgroups of Spanish-speaking ELLs based on language proficiency in both

Spanish and English. Furthermore, this study extends language proficiency to include word-level reading skills and reading comprehension. Doing so provided a more nuanced perspective – particularly with respect to reading comprehension – than if oral language proficiency was considered alone. The findings in this study suggest it is imperative researchers begin to acknowledge and explicitly model such heterogeneity. Doing so may enable researchers to clarify, support, or refute previous findings obtained from research treating this population as a single entity.

The language profiles were not ordered simply in terms of high, medium, and low achievement. While the latent classes were indeed ordered, there were important within-class patterns that emerged across languages further suggesting the need to account for a student's language and reading skills in both languages as skills in the first language can transfer to the second (August et al., 2006; Cummins, 1979; Proctor, August, Carlo & Snow, 2006). For instance, students in the highest achieving class (*High/Spanish Dominant*) performed highest on both Spanish and English measures, though within their own group, they consistently performed better on Spanish measures than English analogues. Therefore, their Spanish skills were dominant, and may have enabled them to perform better on the English measures compared to their peers. The opposite pattern was found for students in the *Average/English Dominant* class, while those in the *Low* class generally performed only slightly better on the Spanish measures.

These differences were most prominent with respect to the *Word Reading* measures, which addresses this study's second research question. That is, the *Word Reading* measures clearly delineated the latent classes whereas there was less distinction with the *Oral Language* measures. This may be a function of the age of the students in this sample. The

research literature has long identified greater amounts of variation in word-level reading skills compared to oral language skills during the early elementary years (e.g. Tunmer & Chapman, 2012; Tunmer & Gough, 1990). Concerning the *Oral Language* measures, the *High/Spanish Dominant* and *Average/English Dominant* performed similarly on the English assessments. Moreover, the *Average/English Dominant* class performed similarly to the *Low* class on the Spanish assessments. Thus, the two higher achieving classes performed at a similar level as the next lower achieving class in their *non-dominant* language on measures of oral language proficiency. This finding makes intuitive sense, but its importance may lie in extending its logic to comparisons of readers at different levels of oral language proficiency. For instance, Mancilla-Martinez et al. (2009) theorized that older Spanish-speaking ELLs with less developed English oral language proficiency may perform similarly to younger English monolinguals. A similar argument may be made concerning the results in this study. Since students in this sample performed more similarly to lower achieving students when assessed in their non-dominant language, this may be viewed as evidence in support of Mancilla-Martinez et al. (2009). Furthermore, in terms of reading comprehension, the *High/Spanish Dominant* and *Average/English Dominant* classes performed equally on English reading comprehension. This suggests that Spanish word-level skills did not give the *High/Spanish Dominant* class an advantage in terms of English reading comprehension.

While this study did not initially intend to examine findings related to linguistic skills transferring between languages, this deserves consideration given the emergence of the *High/Spanish Dominant* class. Cummins (1979) hypothesized that skills developed in the native language could transfer to a student's second language, which was termed the Linguistic Interdependence hypothesis. Findings from this study would initially seem to

support such a notion, but with a more nuanced perspective that would better align with the findings of Proctor et al. (2010). Proctor et al. characterized this interdependence between languages as a continuum and that interdependence was not equal across different reading skills. Specifically, they theorized interdependence would be strongest among alphabetic knowledge given that Spanish and English are orthographically similar. Second, they hypothesized the strength of the interdependence between reading comprehension in both languages would be moderate, while the weakest interdependence would be between Spanish oral language and English reading comprehension. The profile plot in Figure 3 appears to support at least their first contention. Across the three latent classes, students who performed at a particular level on a *Word Reading* measure in one language could be expected to perform at a relatively similar level (compared to their peers) in the other language. For example, those in the *Low* class performed below their peers on both English and Spanish letter-word identification and there did not appear to be overlap with the adjacent *Average/English Dominant* class. Additionally, correlations across languages (see Table 2) were strongest for the observed measures of word-level skills compared to oral language measures and reading comprehension for the sample as a whole. However, as the focus of this study is to examine class-specific associations between observed and latent variables, the topic of linguistic interdependence is revisited below using class-specific findings from the regression mixture model.

Class-Specific SVR Findings

The regression mixture model was used to answer this study's third and fourth research questions. With respect to the third research question, latent class membership was indeed found to moderate the predictive relationships specified by the SVR. While both

Word Reading and *Oral Language* were predictive of reading comprehension across languages for the full sample, this was not the case for all three of the latent classes. With respect to the fourth research question, each latent class demonstrated distinct patterns of language-specific relationships between the SVR components. These findings are discussed in further detail in the following sections.

Class-specific regressions. Each of the latent classes exhibited differential relationships between reading comprehension and its predictors in both languages. For the *Low* class, *Word Reading* was a stronger predictor than *Oral Language* regardless of language. Typically, word-level reading skills are associated with reading comprehension skills early in a child's academic career. Since this subgroup of students performed below their peers, their reading development appears to parallel that of younger students. This finding may help explain the surprising results found by Mancilla-Martinez et al. (2009), in which decoding skills were found to be a stronger predictor of reading comprehension than oral language skills in a sample of adolescent Spanish-speaking LM students. Though the present study utilized a much younger sample, the findings confirm those results for some native Spanish-speakers. The sample in Mancilla-Martinez et al. (2009) excluded recent immigrants and consisted of students in mainstream classrooms. They note that nearly half of their sample was previously classified as Limited English Proficient, suggesting their sample was considered to be fluent in English at the time of their study. Yet, it is the language profile of the *Low* group in this study that most resembles their findings. One reason may be their sample was drawn from an impoverished urban public school district. As these districts tend to suffer from poor academic achievement outcomes, the reading development of the students in their study may have been similar to the lowest achieving

students in the present study. While their sample likely contained heterogeneous language profiles, it is possible the majority had poor reading skills, and, when averaged over the whole sample, a profile akin to the *Low* subgroup found here emerged. The present study supports their findings that language development in non-native English speakers may parallel that of younger English monolinguals, but this may only apply to those learners who are already lagging in reading and language skills relative to their peers regardless of language.

For the *Average/English Dominant* class, only *Oral Language* was a significant predictor of reading comprehension across languages. Furthermore, the standardized coefficient in Spanish was more than twice as large as the coefficient in English. This may appear to be a somewhat puzzling finding given that this latent class showed stronger English than Spanish skills across all observed measures. However, the regression coefficients do not necessarily reflect achievement on the language measures, only the magnitude of the predictive relationship between the predictors and outcome. A close examination of the item-profile plot in Figure 3 shows that the Spanish vocabulary measures are at nearly the same level as the Spanish reading comprehension measure for this class. On the other hand, the English reading comprehension measure is at a higher level than the English vocabulary measures. Therefore, the predictive relationship in Spanish appears to align more closely with a linear trend, which would explain the larger coefficient.

That said, the more interesting finding appears to be the pattern of the predictive relationships. This latent class may be most similar to English monolinguals at a similar age. Typically, third grade is the developmental point at which one would expect oral language skills to emerge as the stronger predictor of reading comprehension in English

monolinguals. This finding makes intuitive sense as it readily aligns with that expectation. Since these students have stronger English skills than Spanish skills, it follows their English reading comprehension would be more strongly related to their English oral proficiency. However, this is also the case in Spanish. It may be there was some degree of linguistic interdependence in terms of language for this subgroup. Laija-Rodríguez et al. (2006) found a weak, but statistically significant, relationship between academic language measures in Spanish and English using a sample of similarly-aged students. Consistent with their finding, this particular latent class may have been able to draw on their English oral language skills to support their Spanish reading comprehension.

An alternative explanation may simply be this class had little variation in terms of their *Word Reading* skills in both languages. Though they were not the highest performing class on these measures, they may simply have all performed similarly at an average level. If so, then there would not be enough variation within these measures to significantly predict their reading comprehension scores. This should not be interpreted as these students having mastered word-level reading in both languages since the *High/Spanish Dominant* subgroup still scored higher across the *Word Reading* and *Oral Language* measures than the *Average/English Dominant* subgroup. For these students, it may be more important for teachers to focus on developing their oral language skills while continuing to monitor their word-level reading skills.

The third latent class, *High/Spanish Dominant* showed an altogether different pattern of predictive relationships that appeared to be language-dependent. Specifically, for these students, Spanish reading comprehension was significantly predicted by Spanish *Oral Language* while English reading comprehension was significantly predicted by English

Word Reading. For these students, relative reading and language proficiencies are directly associated with their reading comprehension skills. For instance, though this latent class scored higher than the other classes across the observed measures, their Spanish skills were considerably stronger than their English skills. Thus, in English, they appear to mirror younger, less advanced English monolinguals. However, in Spanish, their skill profile is comparable to older, more advanced English monolinguals. Unlike the *Average/English Dominant* class, if there was any linguistic interdependence occurring with this class, it did not seem to affect English reading comprehension. Specifically, their advanced (compared to the other latent classes) skills in Spanish do not appear to have advantaged them in regards to English reading comprehension compared to the *Average/English Dominant* class. Examining both the *Average/English Dominant* and *High/Spanish Dominant* classes, it is clear they were not well-differentiated in terms of English vocabulary measures (see Figure 3). However, there was clear delineation in the Spanish vocabulary measures and Spanish and English word-level reading measures. Perhaps linguistic interdependence for the *High/Spanish Dominant* class was limited to the word-level reading measures.

Synthesizing results across the *Average/English Dominant* and *High/Spanish Dominant* classes provides some support for Kieffer's (2008) findings. His study found that non-native English speakers who were fully English proficient experienced longitudinal development of reading skills similarly to English monolinguals. Furthermore, those who were not fully English proficient performed dramatically lower than either of the other two groups. There are parallels that can be drawn with this study. As in Kieffer (2008), English oral language proficiency appears to have played a role in this study. However, this study extends this to include both Spanish and English oral language proficiency relative to each

other. A consistent finding was the oral language of dominance was predictive of reading comprehension within that language. Though the *Average/English Dominant* class was lower achieving overall, their English dominance may have enabled them to perform similarly to the higher achieving *High/Spanish Dominant* class in terms of English reading comprehension and vice versa. These latent classes may have been similar to the LM learners who were fully English proficient in the Kieffer (2008) study. The *Low* class in this study may have been comparable to those learners who were not classified as English proficient in Kieffer (2008).

Factor correlations. The results for the class-specific factor correlations generally followed expected patterns across classes. The *Word Reading* and *Oral Language* factors were correlated within each language for all three latent classes. However, two factor correlations warrant further discussion. For the *Average/English Dominant* latent class, English and Spanish *Word Reading* were highly and significantly correlated. This provides further evidence there was little variation in the *Word Reading* factors thereby leading to *Oral Language* being the stronger predictor of reading comprehension for this subgroup.

Second, for the *High/Spanish Dominant* subgroup, the English and Spanish *Oral Language* factors were significantly and moderately correlated. This was somewhat surprising given this class scored higher on the Spanish vocabulary measures than the English vocabulary measures. However, the correlation was moderate and was limited to oral language. Perhaps these students were more often exposed to navigating environments that required oral facility in both languages compared to their peers. For example, they may have had more experiences as oral translators for family members or lived in neighborhoods where Spanish was the dominant spoken language, but English was also necessary. There

are likely a number of other sociological scenarios that would explain this finding and this is certainly deserving of further research.

Factor means. Factor means followed all expected patterns. Specifically, the *Low* class had lower factor means for all four factors compared to the *High/Spanish Dominant* class and these differences were statistically significant. Compared to the *High/Spanish Dominant* class, the *Average/English Dominant* class had lower factor means on English and Spanish *Word Reading* and Spanish *Oral Language*, but not English *Oral Language*. This is not surprising given both classes scored similarly on the English vocabulary measures and English comprehension. This provides further evidence for the results discussed in more detail above.

Implications for Practice

Educators and other professionals who work directly with Spanish-speaking ELLs understand there are individual linguistic differences among these students. Educational policy regularly assesses and categorizes these students according to their English oral language proficiency in an effort to match students with services deemed most appropriate. However, research has been slow to recognize such heterogeneity and this study is a small step in this direction.

The results found here suggest that assessments seeking to classify Spanish-speaking ELLs solely according to English oral language proficiency may not be fully capturing the entirety of students' language skills. For instance, if this study had only assessed the English vocabulary and reading comprehension measures, it would have been difficult to discern between the *High/Spanish Dominant* and *Average/English Dominant* subgroups. Including Spanish measures, however, differentiated these subgroups and demonstrated those who had

greater Spanish skills obtained generally higher achievement across the majority of the measures. This is a particularly salient finding in light of research concerning transference of literacy skills in a student's first language to her second language. Taken together, this suggests English language programs that ignore Spanish literacy development may not be taking advantage of an entire set of skills that can accelerate English reading acquisition and comprehension. As August et al. (2006) stated, adequate literacy skills in the first language must exist prior to transference. This study found evidence of this in both the *Average/English Dominant* and *High/Spanish Dominant* subgroups. For the former, word-level reading skills were related in both languages, while oral language skills were related in the latter subgroup. These relations were not identified in the *Low* subgroup suggesting their Spanish literacy skills may not have been well-developed, so linguistic transference may not be an advantage for this subgroup at this point in their learning.

Assessments that include measures of native language proficiency in addition to English language proficiency may better inform teaching practices. If included, these assessments may be able to provide a more nuanced picture of students' relative strengths and weaknesses. Teachers may then be able to build on students' native-language literacy skills if those skills are found to be sufficiently developed. For instance, teachers may be better able to group students according to proficiency levels in both their native and second languages to provide targeted support to students at similar achievement levels. In this study, teachers may be able to capitalize on linguistic transference in terms of word-level reading skills with the *Average/English Dominant* subgroup as these were strongly correlated across languages. Indirect strategies such as drawing connections between typologically similar letters may be sufficient to foster continued development of decoding skills. Vocabulary and

oral language development, however, may require more explicit instruction as this subgroup did not demonstrate an oral language correlation across languages. For instance, it may not be sufficient to simply introduce definitions of words. Individual vocabulary words may need to be taught using multiple methods such as pre-teaching the same word in both languages followed by discussing them in text, and finally, having students use the word(s) directly in their own work and in both languages.

The *High/Spanish Dominant* subgroup may benefit from inferring connections between vocabulary words in both languages as oral language skills were correlated across languages for these students. For example, it may be easiest for this subgroup to learn new English words by using the Spanish equivalents. In cases where there is not a direct equivalent, teachers may need to introduce new English words using multiple methods such as those with the *Average/English Dominant* subgroup. They may also have students define the word more fully using Spanish and then attempt to transfer these descriptions into English allowing a more in-depth exploration of the word. Since this subgroup performed the best on measures of decoding, teachers may not need to focus instruction on these skills and, instead, focus on vocabulary with some periodic reinforcement of decoding skills as words become increasingly complex or irregular.

The *Low* subgroup may require more explicit instruction in both decoding and vocabulary until they are more fully developed to a point allowing for linguistic transference. It may be most beneficial to utilize curricula that explicitly teach skills in both languages for this subgroup. This might increase the opportunities for linguistic transference. Additionally, this subgroup may benefit more from explicit and intense instruction more than the other two subgroups. A student's membership in any of these

particular subgroups likely changes over time. Teachers should carefully and repeatedly monitor these students' development as their skills may advance to a point consistent with the higher-performing students and make instructional and grouping adaptations as necessary.

Directions for Future Research

One of the methodological strengths of the approach utilized in this study is mixture models are model-based, which allows for replication with independent samples. Replicating the findings from the present study is one area of further research. But as this study is one of the first to utilize this methodological technique with Spanish-speaking ELLs, there is much room for further investigations extending this framework to varying ages, languages, demographic variables, etc. Moreover, researchers should design and test the efficacy of targeted interventions based on subgroup membership in terms of relative language proficiency. Researchers should also track the longitudinal developmental profiles of the subgroups of Spanish-speaking ELLs. Varying trajectories may identify particular subgroups at increased risk and/or whether any of the subgroups close achievement gaps over time. These are only a few suggestions for further research, but as heterogeneity in this population is studied further, many more questions are sure to arise.

Limitations

While this study is one of the first to empirically identify heterogeneous groups of Spanish-speaking ELLs and the moderating effect of group membership, it is not without limitations. First, this study only examined grades two and three. Relative language proficiency is not static and can change due to a variety of factors such as language instruction. Therefore, the number and types of emergent latent classes may be different at

other ages and/or grades. Further, the relations between latent classes and the varying predictive capacities of word reading and oral language may not remain constant as Spanish-speaking ELLs' relative language proficiency changes. Thus, even if the same latent classes were to emerge at other ages and/or grades, the class-specific results with the SVR found here may be different.

This study did not examine cross-language predictive relationships for each of the latent classes. It may be possible that Spanish *Word Reading* and/or *Oral Language* significantly predicts English reading comprehension and vice versa for one or more of the latent classes. Furthermore, this study did not assess whether cross-linguistic interactions between *Word Reading* and *Oral Language* were related to Spanish or English reading comprehension for any of the latent classes. This salient question also deserves more attention by researchers, especially given the findings of Proctor et al. (2006) who found cross-linguistic effects and interactions. These investigations were beyond the scope of this study, but they certainly warrant further research.

Finally, the sample size in this study may be considered relatively low given the complexity of the model. Van Horn et al. (2014) recommend this technique be utilized with large samples. However, we believe the findings in the present study are viable and defensible for three reasons. First, the latent classes were well-defined as evidenced by the high entropy value of the final model (.94). Second, the preferred unconditional model consisted of only three latent classes, which were substantively meaningful. Finally, there were a small number of regressions in the final model and the results of the differential regression effects across classes made sense on an intuitive level and aligned with previous theoretical considerations.

Research addressing learning to read in a second language has evolved over the past 40 years. It is now generally agreed upon that reading and language skills in a student's native language can be an asset when learning corresponding skills in a second language. Thus, modern research should consider proficiency across both languages when designing future studies. Furthermore, researchers should take explicit steps to recognize and account for heterogeneity within populations of second-language learners as has long been done for English monolinguals. The relatively recent emergence of advanced methodological techniques will allow researchers to craft increasingly complex questions that will be able to shed light on the nuanced and intricate processes that comprise well-developed reading skills. This study has begun to use this methodology to take such steps, but a larger research agenda is both warranted and timely.

References

Aaron, P.G., Joshi, R.M., Gooden, R., & Bentum, K.E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading: An alternative to the discrepancy model of LD. *Journal of Learning Disabilities, 41,* 67-84. doi: 10.1177/0022219407310838

Abedi, J., Hofstetter, C.H., & Lord, C. (2004) Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74,* 1-28. doi: 10.3102/00346543074001001

Adlof, S.M., Catts, H.W., & Little, T.D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19,* 933-958. doi: 10.1007/s11145-006-9024-z

Asparouhov, T., & Muthen, B. (2013). *Auxiliary Variables in Mixture Modeling : 3-Step Approaches Using Mplus* (pp. 1–48). Retrieved from http://www.statmodel.com/examples/webnotes/webnote15.pdf

Batalova, J., & McHugh, M. (2010). *Top languages spoken by English language learners nationally and by state.* Washington, DC: Migration Policy Institute. Retrieved from: http://www.migrationpolicy.org/research/top-languages-spoken-english-language-learners-nationally-and-state

Boscardin, C.K., Muthén, B., Francis, D.J., & Baker, E.L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology, 100,* 192-208. doi: 10.1037/0022-0663.100.1.192

Brown, T.A. (2006). *Confirmatory factor analysis for applied researchers.* New York, NY: The Guilford Press.

Brownell, K. (2001). *Expressive One-Word Picture Vocabulary Test* (3rd ed.). New York, NY: Academic Therapy.

Bureau of Labor Statistics (2014). *The labor force participation rate since 2007: Causes and policy implications*. US Department of Labor. July 2014. Retrieved from https://www.whitehouse.gov/sites/default/files/docs/labor_force_participation_report.pdf

Catts, H.W., Adlof, S.M., Weismer, S.E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49,* 278-293. doi: 10.1044/1092-4388(2006/023)

Catts, H.W., Compton, D., Tomblin, J.B., & Bridges, M.S. (2011). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/a0025323

Catts, H.W., Fey, M.E., Zhang, X., & Tomblin, B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3,* 331-361. doi: 10.1207/s1532799xssr0304_2

Catts, H.W., Hogan, T.P., & Fey, M.E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36*, 151-164. doi: 10.1177/002221940303600208

Compton, D.L., Fuchs, D., Fuchs, L.S., Elleman, A.M., & Gilbert, J.K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences, 18,* 329-337. doi: 10.1016/j.lindif.2008.04.003

Connor, C.M., Alberto, P.A., Compton, D.L., & O'Connor, R.E. (2014). *Improving reading*

outcomes for students with or at risk for reading disabilities: A synthesis of the contributions from the Institute of Education Sciences Research Centers (NCSER 2014-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/.

Cutting, L.E. & Scarborough, H.S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10,* 277-299. doi: 10.1207/s1532799xssr1003_5

Dunn, L. M., & Dunn, L. M. (1981). *The Peabody Picture Vocabulary Test–Revised.* Circle Pines, MN: American Guidance Service.

Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario Imagenes Peabody* [The Peabody Picture Vocabulary Test]. Circle Pines, MN: American Guidance Service.

Farnia, F. & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading, 36,* 389-421. doi: 10.1111/jrir.12003

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*, 889-912. doi: 10.1007/s11145-012-9397-0

Fry, R. (2007). *How far behind in math and reading are English language learners?* Washington, DC: Pew Hispanic Center

Gándara, P. & Contreras, F. (2009). *The Latino education crisis: The consequences of failed*

social policies. Cambridge, MA: Harvard University Press

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives*, *11*(36). Retrieved from http://epaa.asu.edu/epaa/v11n36/

Gibson, W.A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, *24*, 229-252. doi: 10.1007/bf02289845

Gottardo, A. & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *101*, 330-344. doi: 10.1037/a0014320

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

Guzman-Orth, D. (2013). *When is an English language learner not an English language learner? Exploring individual differences in developmental language and literacy acquisition for at-risk learners: A latent transition approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3545051)

Hemphill, F.C., and Vanneman, A. (2011). *Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress* (NCES 2011-459). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Herbers, J.E., Cutuli, J.J., Supkoff, L.M., Heistad, D., Chan, C-K., Hinz, E., & Masten, A.S. (2012). Early reading skills and academic achievement trajectories of students facing

poverty, homelessness, and high residential mobility. *Educational Researcher, 41,* 366-374. doi: 10.3102/0013189X12445320

Hernandez, D.J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation.* NewYork: The Annie E. Casey Foundation. Retrieved from http://www.aecf.org/resources/double-jeopardy/

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127-160. doi: 10.1007/bf00401799

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Convention criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,*1-55. doi: 10.1080/10705519909540118

Jones, S.M., Brown, J.L., & Aber, J.L. (2011). Two-year impacts of a universal school based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development, 82,* 533-554. doi: 10.1111/j.1467-8624.2010.01560.x

Kao, G. & Thompson, J.S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology, 29,* 417-442. doi: 10.1146/annurev.soc.29.010202.100019

Kaplan, D. (2008). An overview of Markov Chain methods for the study of stage-sequential developmental processes. *Developmental Psychology, 44,* 457-467. doi: 10.1037/0012-1649.44.2.457

Kaplan, D. & Walpole, S. (2005). A stage-sequential model of reading transitions: Evidence from the Early Childhood Longitudinal Study. *Journal of Educational Psychology, 97,* 551-563. doi: 10.1037/0022-0663.97.4.551

Kendeou, P., Savage, R., & van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology, 79*, 353-370. doi: 10.1348/978185408X369020

Kendeou, P., van den Broek, P., White, M.J., & Lynch, J.S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*, 765-778. doi: 10.1037/a0015956

Kershaw, S. & Schatschneider, C. (2012). A latent variable approach to the simple view of reading. *Reading and Writing, 25*, 433-464. doi: 10.1007/s11145-010-9278-3

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*, 851–868. doi: 10.1037/0022-0663.100.4.851

Kieffer, M. J. (2010). English proficiency, socioeconomic status, and late-emerging reading difficulties. *Educational Researcher, 39*, 484–486. doi: 10.3102/0013189x10378400

Kieffer, M. J. (2012). Early oral language and later reading development in Spanish speaking English language learners: Evidence from a nine-year longitudinal study. *Journal of Applied Developmental Psychology, 33*, 146-157. doi: 10.1016/j.appdev.2012.02.003

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*, 433-452. doi: 10.1177/0022219411432683

Kirby, J.R. & Savage, R.S. (2008). Can the simple view deal with the complexities of reading? *Literacy, 42*, 75-82. doi: 10.1111/j.1741-4369.2008.00487.x

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher, 31,* 3-12. doi: 10.3102/0013189x031001003

Lesaux, N.K., Crosson, A.C., Kieffer, M.J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology, 31*, 475-483. doi: 10.1016/j.appdev.2010.09.004

Lesaux, N. K., Koda, K., Siegel, L. S., & Shanahan, T. (2006). Development of literacy of language minority learners. In D. L. August & T. Shanahan (Eds.), *Developing literacy in a second language: Report of the National Literacy Panel* (pp. 75−122). Mahwah, NJ: Lawrence Erlbaum.

Little, R. L., & Rubin, D. B. (1990). *Statistical analysis with missing data*. New York, NY: Wiley.

Lonigan, C.J., Farver, J.M., Nakamoto, J., & Eppe, S. (2013). Developmental trajectories of preschool early literacy skills: A comparison of language-minority and monolingual English children. *Developmental Psychology, 49*, 1943-1957. doi: 10.1037/a0031408

Mancilla-Martinez, J., Kieffer, M.J., Biancarosa, G., Christodoulou, J.A., & Snow, C.E. (2009). Investigating English reading comprehension growth in adolescent language minority learners: Some insights from the simple view. *Reading and Writing, 24,* 339-354. doi: 10.1007/s11145-009-9215-5

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners.

Journal of Educational Psychology, 102, 701–711. doi: 10.1037/a0019135

Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 375–393). Oxford, UK: Oxford University Press.

Miles, S.B. & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development, 77,* 103-117. doi: 10.1111/j.1467-8624.2006.00859.x

Muthén, L.K., & Muthén, B.O. (1998-2012). *Mplus user's guide.* Los Angeles: Muthén & Muthén

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading & Writing: An Interdisciplinary Journal, 20,* 691–719. doi: 10.1007/s11145-006-9045-7

Nakamoto, J., Lindsey, K.A., & Manis, F.R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading, 12,* 351-371. doi: 10.1080/10888430802378526

Nation, K. Cocksey, J., Taylor, J.S.H., & Bishop, D.V.M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *The Journal of Child Psychology and Psychiatry, 51,* 1031-1039. doi: 10.1111/j.1469-7610.2010.0225.x

Nation, K., & Snowling, M.J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27,* 342-256. doi: 10.1111/j.1467-9817.2004.00238.x

National Assessment of Educational Progress. (2013). *NAEP data explorer*. Retrieved from http://nces.ed.gov/nationsreportcard/naepdata/

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

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nylund, K.L., Asparouhov, T., & Muthén, B.O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535-569. doi: 10.1080/10705510701575396

Nylund-Gibson, K., Grimm, R., Quirk, M., & Furlong, M. (2014). A latent transition mixture modeling using the three-step specification. *Structural Equation Modeling, 21*(3), 439-454. doi: 10.1080/10705511.2014.915375

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading & Writing, 23,* 189-208. doi: 10.1007/s11145-008-9159-1

Pasquarella, A., Gottardo, A., & Grant. A. (2012). Comparing factors related to reading comprehension in adolescents who speak English as a first (L1) or second (L2) language. *Scientific Studies of Reading, 16,* 475-503. doi: 10.1080/10888438.2011.593066

Proctor, C.P., August, D., Snow, C., & Barr, C.D. (2010). The interdependence continuum:

A perspective on the nature of Spanish-English bilingual reading comprehension. *The Journal of the National Association for Bilingual Education, 33,* 5-20. doi: 10.1080/15235881003733209

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,* 246-256. doi: 10.1037/0022-0663.97.2.246

Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with Grade 2 achievement. *Early Childhood Research Quarterly, 28,* 437-449. doi: 10.1016/j.ecresq.2012.11.002

Ram, N. & Grimm, K.J. (2009). Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development, 33,* 565-576. doi: 10.1177/0165025409343765

Raykov, T. & Marcoulides, G.A. (2008). *An introduction to applied multivariate analysis.* New York, NY: Routledge.

Reardon, S.F. & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal, 46,* 853-891. doi: 10.3102/0002831209333184

Reynolds, A.J., Temple, J.A., Robertson, D.L., & Mann, E.A. (2002). Age 21 cost-benefit analysis of the Title I Chicago child-parent centers. *Educational Evaluation and Policy Analysis, 24,* 267-303. doi: 10.3102/01623737024004267

Roth, F.P., Speece, D.L., & Cooper, D.H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research, 95,*

259-272. doi: 10.1080/00220670209596600

Rumberger, R. W., & Anguiano, B. A. (2004). *Understanding and addressing the California Latino achievement gap in early elementary school.*Retrieved from University of California Linguistic Minority Research Institute, University of California website: https://escholarship.org/uc/item/65d6v84n

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461-464.

Silverman, R.D., Speece, D.L., Harring, J.R., & Ritchey, K.D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading, 17,* 108-133. doi: 10.1080/10888438.2011.618153

Snow, C.E. (2001). *Reading for understanding: Toward an R & D program in reading comprehension.* Santa Monica, CA: RAND.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stanovich, K.E. (2008). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education, 189,* 23-55

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947. doi: 10.1037//0012-1649.38.6.934

Tunmer, W.E., & Chapman, J.W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities,45,* 453-466. doi: 10.1177/0022219411432685

U.S. Department of Education, National Center for Education Statistics (2014). *The Condition of Education 2014* (NCES 2014-083, English Language Learners.

Retrieved from:https://nces.ed.gov/programs/coe/indicator_cgf.asp

Van Horn, M.L., Jaki, T., Masyn, K., Howe, G., Feaster, D.J., Lamont, A.E.,…Kim, M. (2014). Evaluating differential effects using regression interactions and regression mixture models. *Educational and Psychological Measurement.* Advance online publication. doi: 10.1177/0013164414554931

Van Horn, M.L., Jaki, T., Masyn, K., Ramey, S.L., Smith, J.A., & Antaramian, S. (2009). Assessing differential effects: Applying regression mixture models to identify variations in the influence of family resources on academic achievement. *Developmental Psychology, 45*, 1298-1313. doi: 10.1037/a0016427

Vellutino, F.R., Tunmer, W.E., Jaccard, J.J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11,* 3-32. doi: 10.1080/10888430709336632

Vermunt, J. K. (2010). Latent Class Modeling with Covariates: Two improved Three-Step Approaches. *Political Analysis*, *18*, 450–469. doi:10.1093/pan/mpq025

Vermunt, J.K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars, & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89-106). Cambridge: Cambridge University Press.

Wasserman, L. (1997). *Bayesian model selection and model averaging* (Working Paper No. 666). Carnegie Mellon University, Department of Statistics

Woodcock, R. W. (1998). *Woodcock Reading Mastery Test–Revised* (*Form G).* Circle Pines, MN: American Guidance Service.

Woodcock, R. W., Muñoz-Sandoval, A. F., & Alverado, C. G. (2005). *Woodcock–Muñoz Language Survey.* Itasca, IL: Riverside.