1-1-2016

Science as a Second Language: Analysis of Emergent Bilingual Performance and the Impact of English Language Proficiency and First Language Characteristics on the Colorado Measures of Academic Success for Science

Joanna Kristine Bruno
University of Denver

Science as a Second Language: Analysis of Emergent Bilingual Performance and the Impact of English Language Proficiency and First Language Characteristics on the Colorado Measures of Academic Success for Science

Abstract

In an age when communication is highly important and states across the nation, including Colorado, have adopted Common Core State Standards, the need for academic language is even more important than ever. The language of science has been compared to a second language in that it uses specific discourse patterns, semantic rules, and a very specific vocabulary. There is a need for educators to better understand how language impacts academic achievement, specifically concerning Emergent Bilinguals (EBs). Research has identified the need to study the role language plays in content assessments and the impact they have on EBs performance (Abedi, 2008b; Abedi, Hofestter & Lord, 2004; Abedi & Lord, 2001). Since language is the means through which content knowledge is assessed, it is important to analyze this aspect of learning. A review of literature identified the need to create more reliable and valid content assessments for EBs (Abedi, 2008b) and to further study the impact of English proficiency on EBs performance on standardized assessments (Solorzano, 2008; Wolf, & Leon, 2009). This study contributes to the literature by analyzing EBs performance on a state-level science content assessment, taking into consideration English language proficiency, receptive versus productive elements of language, and students' home language. This study further contributes by discussing the relationship between language proficiency, and the different strands of science (physical, life, and earth) on the state science assessment. Finally, this study demonstrates that home language, English language proficiency, and receptive and productive elements of language are predictive of EBs' achievement on the CMAS for science, overall and by strand. It is the blending of the social (listening and speaking) with the academic (reading and writing) that is also important and possibly more important.

Document Type

Dissertation

Degree Name

Ph.D.

Department

Curriculum and Instruction

First Advisor

Maria del Carmen Salazar, Ph.D.

Second Advisor

Kathy Green

Third Advisor

Richard Kitchen

Keywords

Assessment, Emergent bilinguals, Middle school, Receptive and productive langauge, Science achievement, Second language acquisition

SCIENCE AS A SECOND LANGUAGE: ANALYSIS OF EMERGENT BILINGUAL PERFORMANCE AND THE IMPACT OF ENGLISH LANGUAGE PROFICIENCY AND FIRST LANGUAGE CHARACTERISTICS ON THE COLORADO MEASURES OF ACADEMIC SUCCESS FOR SCIENCE

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Joanna K. Bruno

November 2016

Advisor: Maria del Carmen Salazar

Author: Joanna K. Bruno
Title: SCIENCE AS A SECOND LANGUAGE: ANALYSIS OF EMERGENT BILINGUAL PERFORMANCE AND THE IMPACT OF ENGLISH LANGUAGE PROFICIENCY AND FIRST LANGUAGE CHARACTERISTICS ON THE COLORADO MEASURES OF ACADEMIC SUCCESS FOR SCIENCE
Advisor: Maria del Carmen Salazar
Degree Date: November 2016

Abstract

In an age when communication is highly important and states across the nation, including Colorado, have adopted Common Core State Standards, the need for academic language is even more important than ever. The language of science has been compared to a second language in that it uses specific discourse patterns, semantic rules, and a very specific vocabulary. There is a need for educators to better understand how language impacts academic achievement, specifically concerning Emergent Bilinguals (EBs). Research has identified the need to study the role language plays in content assessments and the impact they have on EBs performance (Abedi, 2008b; Abedi, Hofestter & Lord, 2004; Abedi & Lord, 2001). Since language is the means through which content knowledge is assessed, it is important to analyze this aspect of learning. A review of literature identified the need to create more reliable and valid content assessments for EBs (Abedi, 2008b) and to further study the impact of English proficiency on EBs performance on standardized assessments (Solorzano, 2008; Wolf, & Leon, 2009). This study contributes to the literature by analyzing EBs performance on a state-level science content assessment, taking into consideration English language proficiency, receptive versus productive elements of language, and students' home language. This study further contributes by discussing the relationship between language proficiency, and the different strands of science (physical, life, and earth) on the state science assessment. Finally, this study demonstrates that home language, English language proficiency, and receptive and productive elements of language are predictive of EBs' achievement on the CMAS for science, overall and by strand. It is the blending of the social (listening and speaking) with the academic (reading and writing) that is also important and possibly more important.

Acknowledgements

I must thank many people in my life for supporting me throughout this process. First, I would like to thank my family for listening, proofreading, encouraging me, and giving me the love that I needed to get this completed. Second, I would like to thank my amazing group of friends who, at times, talked me off the ledge, kept my body and mind fit, and provided the advocacy I needed to get through this project. Third, I would like to thank my colleagues who inspired me and provided guidance. I would also like to thank my advisor, Maria del Carmen Salazar, and my quantitative research methods professor, Kathy Green, who kept me focused when I wanted to make my project bigger and bigger, and for their professional knowledge throughout this journey.

In addition to my parents, I dedicate this to the two patriarchs in my family, Aniceto Lucero and Angelo Bruno, who taught me the value of hard work and perseverance.

Table of Contents

List of Tables

Chapter 1: Introduction

Colorado state law defines an English language learner as "a student who is linguistically diverse and who is identified pursuant to section 22-24-105 (2) as having a level of English language proficiency that requires language support to achieve standards in grade-level content in English" (CRS 22-24-103 (4)). Emergent Bilinguals (EBs) [1] are defined as students not yet proficient in English, but through school and acquiring English they become bilingual (Garcia, Kleifgen, & Falchi, 2008). The use of the term EB is intentional in order to showcase these learners from an asset orientation, as opposed to a deficit orientation. Garcia et.al. (2008) argue that the use of other terms to describe this population of students (e.g., language minority, limited English proficient) perpetuates the inequities and disadvantages that EBs encounter in their education and ignore their home language and cultural understanding.

The number of EBs nationally has increased 49% over the last ten years (NCELA, 2015). In Colorado, EBs represent the fastest growing student population with an increase of 21.3 % over the last ten years (CDE, 2014b). Meanwhile, the Colorado total student population increased by 6.8% within the same time frame (CDE, 2014b). In the 2013-2014 school year, EBs represented approximately 12% of the public school students in Colorado (CDE, 2014b) and 56.9 % of the EBs in Colorado were concentrated in the

[1] Emergent Bilinguals (EB) are synonymous with what the mainstream labels English Language Learners (ELL), English learner (EL), language minority (LM), or limited English proficient (LEP) students.

Denver metro area (CDE, 2014b). Over 83% of EBs are concentrated in 10 school districts (CDE, 2014b), and the six largest school districts in Colorado have EB populations ranging from 40% in Adams-Arapahoe County to 7.6% in Jefferson County (CDE, 2014b). Although the majority of large school districts are within the Denver metro area, many smaller districts are being impacted by the demographic shift with high numbers of EBs. For example, Adams 14 has 45.9% EBs and Yuma has 40.6% EBs (CDE, 2014b).

No Child Left Behind (NCLB) (2001) asserts that states and districts must hold the same high standards for EB students as they do for all other students, and that they are accountable for assuring that all students meet high academic expectations. NCLB requires that EB students be included in annual state assessments and subsequent accountability measures; federal and state legislation brings attention to the progress of EB students in English proficiency and in academic achievement. The assessment system serves as both a lever for, and an instrument of, reform; for accountability purposes, the assessment results assist in establishing goals and incentives for improvement. Moreover, as data, the results provide educators with information for assessing the success of their programs and planning instruction. The new federal legislation signed into law in December of 2015, titled the “Every Student Succeeds Act” (ESSA), allows for more state flexibility over assessment and accountability measures; however there are still federal requirements for both. Therefore, states are still obligated to have standards, assessments, and accountability measures for EBs. Colorado has more stringent laws for the assessment of EBs than the federal government. For example, the former federal legislation, NCLB, allowed for EBs to be exempt from assessment within their first year in the country. However, Colorado law (CRS 22-24-103 (4)) states that every child will take assessments, regardless of when they arrived. This has been especially problematic for students with limited proficiency in English because an assessment given in English is inherently an assessment of English, and these students do not have a strong enough command of English to do well on the science assessment.

Background and Context

Drawing upon personal experience as a classroom science teacher, English as a second language teacher, and a teacher educator, this dissertation study is framed using the conceptual lens of "educational equity." In this context, equity is defined by Blankstein & Noguera (2015) as "…a commitment to ensure that every student receives what he or she needs to succeed…" (p. 3). In the United States, the pursuit of equity in education began in 1983 with *A Nation at Risk* and there has been almost no progress around closing the achievement gap in almost 30 years (Blankstein & Noguera, 2015). In Colorado, the achievement gap (the disparity of performance among different groups of students) and the opportunity gap (the underlying causes to these disparities, such as the opportunity to learn) (Flores, 2007) is widening between the EB and non-EB populations of students, so the pursuit of educational equity for EB students is a moral imperative.

The need to investigate possible factors that may contribute to these gaps and to understand how language influences achievement was the central focus of this study. English language proficiency (ELP) is measured through four domains of language: listening, speaking, reading, and writing. Traditionally ELP is reported as an overall score and by individual domains. However, this study investigated the influence of the overall ELP scores, in addition to the combinations of the receptive (reading and listening) and productive (writing and speaking) elements of language on overall science achievement and by individual science strand (physical, life, and earth science). Data from eighth grade Colorado EB students, who took both the 2015 state-level English language proficiency and the 2015 state-level science achievement assessments, were investigated.

Research Questions

This study analyzed 8th grade EBs performance on the CMAS for science (overall score and content domain) by exploring the influence of linguistic factors of English proficiency, receptive and productive elements, and students' home language. Specific research questions were:

1. What factors predict EBs performance on the Colorado Measures of Academic Success (CMAS) for science assessment?
    a. To what extent did the students' primary home language predict performance on the 8th grade CMAS for science?
    b. Beyond the students' home language, to what extent did the level of English language proficiency predict performance on the 8th grade CMAS for science?
    c. Beyond students' home language and the level of English language proficiency, to what extent did receptive and productive elements of language predict performance on the 8th grade CMAS for science?
2. What factors predict EBs performance on specific content domains within the Colorado Measures of Academic Success (CMAS) for science assessment?

a. To what extent did the students' primary home language predict performance on the 8th grade CMAS for science?
b. Beyond the students' home language, to what extent did the level of English language proficiency predict performance on the 8th grade CMAS for science?
c. Beyond students' home language and the level of English language proficiency, to what extent did receptive and productive elements of language predict performance on the 8th grade CMAS for science?

Emergent Bilinguals. Garcia et al. (2008) discuss the importance of changing our discourse around how we categorize students whose primary home language is not English:

> English language learners are in fact *emergent bilinguals.* That is, through school and through acquiring English, these children become *bilingual,* able to continue to function in their home language as well as in English, their new language and that of school. When officials and educators ignore the bilingualism that these students can and often must develop through schooling in the United States, they perpetuate inequities in the education of these children. That is, they discount the home languages and cultural understandings of these children and assume their educational needs are the same as a monolingual child (p. 6).

Bilingualism and biliteracy are assets (Escamilla & Hopewell, 2010) that need to be acknowledged and affirmed. The Emergent Bilingual population is not homogeneous, thus the heterogeneity of the population presents many challenges in serving all learners equitably (Boyson & Short, 2012). Although the majority of EBs in the United States and in Colorado (75% and 80% respectively) are native Spanish-speakers, the group is not monolithic (Boyson & Short, 2012). There is great variety in age, previous schooling, and socioeconomic status of the Spanish-speaking EB population. Some EBs enter schools literate in their home language, yet others enter with no formal schooling and emergent literacies in their first language; most seem to fall somewhere in between. In Colorado, state law (CRS 22-24-103 (4)) defines an English language learner [2] as "a student who is linguistically diverse and who is identified pursuant to section 22-24-105 (2) as having a level of English language proficiency that requires language support to achieve standards in grade-level content in English". Federal and state laws emphasize growth in English proficiency and academic achievement.

Double the Work. Emergent bilingual students need to do double the work to acquire the language and skills not only to function socially and culturally, but also to succeed academically in elementary, secondary, and post-secondary education. Schools need to provide EBs with the opportunity to learn and develop English language skills that will afford them access to the same educational opportunities as their native-English speaking peers. One important factor in second language acquisition (SLA) research is that of Cummins (1979), he differentiates between basic interpersonal communication skills (BICS), the language of everyday communication, and cognitive academic language proficiency (CALP), the complex language and discourse demanded for academic success. Emergent bilingual students need to be afforded the opportunity to learn at the CALP level in order to be successful academically. Bailey (2007) extends Cummins work by dividing the CALP into two categories: school navigational language and curriculum content language to better capture the nuances of academic language.

[2] English language learner is used here since it is the language used within the Colorado Revised Statutes.

Beyond this basic distinction, however, there is only consensus around the idea that academic language is the language needed for academic success (Bailey & Huang, 2011).

Nationally and locally, schools have wrestled with the educational needs of EBs, and today they also confront the additional challenge of high school graduation requirements. Colorado's State Board of Education is currently promulgating rules regarding guidance for graduation guidelines for school districts; academic achievement, as demonstrated on state-level assessments, is a large component of this legislation. The Colorado Measures of Academic Success (CMAS) include state summative assessments in mathematics, reading writing, and communicating, science, and social studies[3]. Each CMAS content area has its own discourse, which include syntactic structures and discipline-specific vocabulary (Bailey & Huang, 2011; Cook, Boals, & Lundberg, 2011). Historically, both on the national and local levels, EBs have performed below par with their native-English speaking peers on large-scale standardized assessments (Abedi & Dietel, 2004; Cook et al., 2011; CDE, 2015a). Emphasizing the acceleration of academic language for EBs within the secondary system is critical. Using second language acquisition-based pedagogies to develop academic language could lead to the narrowing of the EB achievement and opportunity gaps.

Second language acquisition research shows that BICS is generally acquired within one to three years; however research continues to demonstrate that CALP requires between four to seven years (Brown, 2000; Cook et al., 2011; Cummins, 2008). There appears to be a research-to-practice gap that is widened by educational policies that are

[3] CMAS science is required for students once in high school, CMAS social studies is not required at the high school level, and CMAS mathematics, and reading, writing, and communicating are required at the 9th grade level).

disconnected from the reality of the public school system (Vanderline & van Braak, 2011). Emergent bilingual students will constantly be confronted with doing double the work due to the nature of learning English at the same time as learning complex concepts and skills within content areas.

Assessment of EBs. Colorado joined the World-Class Instructional Design and Assessment (WIDA) consortium in 2011 and thus, from 2011 to present, participated with other states in the development of the Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs) to determine English proficiency. The ACCESS has two parts: The placement assessment, which is given to all students who indicate a primary home language other than English, on a home language survey, is used to screen for language proficiency at in-take and determines individual initial placement. The proficiency assessment is a large-scale assessment given to all EBs[4] in January each year to assess their level of English proficiency (CDE, 2015b).

The ACCESS has six performance levels ranging from 1-6 (WIDA, 2015). Students who tested at level 6 are recommended to be re-designated from Limited English Proficient (LEP) to Fluent English Proficient (FEP) monitor year 1. Students would then progress through FEP monitor year 2 and then finally to Formerly English Language Learner (FELL) (CDE, 2015b). Formerly English Language Learners are students who may no longer receive direct support but who are still not performing at a level of their native-English speaking peers.

[4] Only those EBs designated as NEP and LEP, per Colorado law, are required to take the ACCESS. FEP students are not required to take the assessment since they are in “monitor” status prior to being re-designated.

In 2008, the Colorado legislature passed the *Colorado Achievement Plan for Kids* legislation focused on education reform in Colorado. The historic education reform initiative included, among other things, new content standards and new assessments aligned to the new standards. The development of the CMAS was a response to the call for education reform in Colorado. At the onset of full implementation of all the new standards and assessments, Colorado amended the legislation to limit the number of assessments given at the high school level. As of August 2015, students at the secondary level are assessed in grades 6 through 9 in mathematics and reading, writing, and communicating, and in grade 8 and once in high school in science (CRS 22-7-106). These academic assessments are considered high-stakes tests due to the state and federal accountability that accompanies them.

According to the Colorado Revised Statutes, all public school students enrolled in school, regardless of when they entered the school district, or in the case of EBs when they entered the United States, are to take the CMAS in grades 3 through high school (CRS 22-7-106). The Colorado Academic Standards (CAS) for science are integrated in grades Kindergarten through 8 and are divided into three strands (physical, life, and earth sciences) for high school. High school students are assessed on all three strands within the CMAS high school science assessment, regardless of whether or not they have had coursework in those areas. In keeping with the goal of "all students, all standards" (CDE, 2015c), EBs and students with disabilities take the state CMAS assessments. There are, however accommodations for the CMAS; common accommodations include word-to-word dictionaries, additional time for completion, and translated instructions (CDE, 2015d).

In April 2015, the new CMAS for science was given to students in Colorado using an on-line platform. On the 8th grade assessment there are 43 selected response items (multiple choice), 17 constructed response items (short answer), and 20 questions associated with simulations (CDE, 2015e). Test items were written by Colorado educators and assess evidence outcomes across the three strands, as well as nature of science components aligned to the Colorado Academic Standards.

Results on the CMAS for science are reported at various levels: state, district, school, class, and student. The reports include a student's various scale scores, associated performance level, and percent correct scores and are displayed on a four-page report along with comparative information related to the student's school, district, and state performance. Three types of aggregate reports are produced: Content Standards Report, School Performance Level Summary, and an Item Analysis Report. These reports provide summary information for a given school or district (CDE, 2015e).

EB Science Achievement/Opportunity Gap. Nationally, EBs do not perform as well as their native-English speaking peers on achievement assessments (Abedi & Dietel, 2004; Cook et al., 2011). The hard part about writing that statement is not that it is surprising, but rather the fact that any assessment in English is an assessment of English, making it very difficult to make true inferences about student achievement when by their very designation, EBs are not proficient in English.

In Colorado, the gap between EBs and their peers on the 8th grade CMAS science assessment can be seen in the scores and performance levels.[5] As reported by the Colorado Department of Education in 2015, statewide, 29% of students scored at the top

[5] The four CMAS science performance levels are: limited command, moderate command, strong command, and distinguished command (CDE, 2015e).

two performance levels on the 8th grade CMAS for science; however, only 11% of EBs scored within those two performance levels (CDE, 2015e). These results are summarized in Table 1. The following section utilizes this data to outline the research purpose and rationale of this study.

Table 1

2015 CMAS Science Results by Performance Level (%)

	Performance Level				
Students	% Limited Command	% Moderate Command	% Strong Command	% Distinguished Command	% Strong & Distinguished*
All	36.7	34.2	26.5	2.5	29.0
EBs**	61.1	27.8	10.8	1.4	6.1

Note. *Students scoring in either of these performance levels are considered to be on track for College and Career Readiness (CDE website, 2015); **Averages across all EB sub-categories (NEP, LEP, FEP, and FELL).

Research Purpose and Rationale

Educators and policy makers maintain that science education is central to both preparation to be socially responsible adults and academically successful in higher education (Bottia, Stearns, Mickelson,, Moller, & Parker, 2015; Colorado Education Initiative, 2015). There is a concern that the U.S. has a shortage of college graduates sufficient to meet the needs of the science and math professions (Colorado Education Initiative, 2015). In order to maintain a globally competitive workforce and to fulfill the needs of future employers, all students must receive a rigorous education in science and their strengths in these areas must be nurtured and developed early to ensure continued interest in STEM fields (Colorado Education Initiative, 2015). Today's EB students are the future's workforce, so understanding that these students are an asset, identifying opportunities to learn, and accurately assessing and guiding future instruction, will help the nation to prosper.

The educational landscape is changing for EBs with the onset of the Common Core State Standards (CCSS), the Next Generation Science Standards (NGSS), and the WIDA consortium; however, the need to examine the academic performance of EBs cannot wait because the need to increase the academic achievement of EBs is pressing. Colorado began educational reform in 2008 with the passage of Senate Bill 08-212 and adopted new, more rigorous academic standards in ten content areas in 2009. In 2010 Colorado adopted the Common Core Standards for Mathematics and English language arts (ELA) (CDE, 2015c). A new assessment aligned to the new science standards was created and the first year of implementation of the assessment for 8th grade was the spring of 2014. The new CMAS science assessment afforded the opportunity to analyze EBs performance on a rigorous content assessment in the context of an on-line platform, uniform EB designations and proficiency levels, uniform content standards, and an instrument designed to measure those specific content standards.

Analysis of eighth grade EB students' performance on the CMAS for science is a pressing matter due to the high stakes nature of the assessment with regards to accountability for schools and districts[6], and the increased focus on Science Technology Engineering and Mathematics (STEM) at the K-12 level. Choosing eighth grade students for this study was purposeful, for two reasons: (1) the linguistic complexity of the content, concepts, and skills in science increase at the middle school level, and (2) middle school is a gate keeper for students in science; either you hook them or lose them. This

[6] AYP, the accountability measure under the No Child Left Behind Act (NCLB), includes science as well as participation rates.

dissertation represents an effort to better understand the ways in which content knowledge[7] and language ability[8] interact with one another to affect assessment results used for high-stakes accountability purposes. EBs performance on the spring 2015 8th grade CMAS for science, and the influence of linguistic factors, was analyzed. Overall English Language Proficiency (ELP), receptive elements of language (listening and reading), productive elements of language (speaking and writing), and students' primary home language were the linguistic factors used to determine if there was an influence on achievement. In addition to exploring the influence on overall performance on the CMAS for science, there is a need to determine if these linguistic factors influence performance on individual science strands (physical, life, and earth).

The language of science tends to be highly technical; therefore, the strands of science may differ in terms of complexity of language. There may be more cognates, words that are similar in meaning between English and Spanish because they may share the same Latin root and are very similar, within the life sciences; therefore this may influence the difference in performance. It is important to understand that language acquisition takes time, especially with regards to academic language (Thomas and Collier, 1986).

Problem

The Second Language Acquisition research agrees that acquiring Cognitive Academic Language Proficiency takes longer than acquiring Basic Interpersonal

[7] Content knowledge is defined as the level of command a student has of science content knowledge as determined by the interpretation of scale scores on the stage science assessment aligned to the state science academic standards (CMAS Interpretation guide, 2015).

[8] Language ability is defined as the level of command a student has of the English language as determined by the interpretation of scale scores on an English language proficiency assessment that is aligned to state English language proficiency standards (ACCESS Interpretation guide, 2015)

Communication Skills (Cook et al., 2011; Cummins, 2008). It takes between 4-7 years to acquire academic language, however the U.S. school system does not allow for such a timeframe to acquire English (Abedi & Gándara, 2006; Collier, 1987, 1995). Results from high stakes assessments are used to make inferences about students and create policies around appropriate solutions to close the achievement gap; therefore, it is important to identify factors that could mitigate EB status on their performance.

One of the challenges of assessing EB students is the diversity of language, cultural and demographic backgrounds, and levels of English proficiency (Abedi, 2002; Solano-Flores, 2003). A student from a refugee camp, who is not literate in their primary home language, is very different from a student who has attended school in their home country and is literate in their home language. Another example is a second generation student, born in the United States, has been attending school, but speaks only their home language compared to a student whose primary language is non-alphabetic and not easily translated. All of these students are considered EBs and are assessed using the same assessment.

Difficulties also exist in disentangling measures of academic content knowledge, administered in English, from students' English language proficiency. An assessment given in English is inherently an assessment of English, so it could be difficult to tease out how much the student knows in science versus how well they can access the assessment. To be a scientifically literate, one must possess content knowledge, and be able to communicate their understanding effectively using technical language (National Academies Press, 2012). Therefore, language and content are inextricably linked, which makes interpretation of results more difficult. Subsequently, inferences drawn and decisions made from assessment results may not be supported due to a validity argument. Some have argued that assessment results for EB students are not valid due to a violation of construct validity, which brings into question the quality of measurement (Abedi, 2002). Construct validity is simply, the measurement of what is intended to be measured, that is, the science assessment is measuring science content knowledge. However, when the assessment is given in English and some students have not yet acquired English language proficiency, the assessment may not be measuring what it is intended to measure (Bailey et al., 2015).

Recent CMAS test results from 2015 provide dramatic examples of the urgent need to better understand and improve EB students' performance. For example, on the basis of assessment results in science on the CMAS for the 2014-2015 school year, the state reported that the overall EB students science proficiency level was 18% lower that the state average (CDE, 2015a). More specifically, results based on the level of English language proficiency EBs had were as follows: Non-English Proficient (NEP) 28.5% below state average, Limited-English Proficient (LEP) 27.7% below state average, and Fluent-English Proficient (FEP) scored 4 % below state average. These results indicate a need to explore causes, effects, and implications.

Effective policies and practices for reducing the achievement/opportunity gap first require the use of valid measures of EB students' achievement, including both their English proficiency and academic content proficiency. Accurate assessment must undergird any credible analyses of the complex relationships between English proficiency and academic achievement, which are essential in understanding and improving EB students' academic success. If a teacher focuses on what a student can do with regards to language, then they can begin to understand what they are capable of and support them academically. Grant et al., (2011) found that "…helping ELs with their technical and general language proficiency helps their mathematics achievement" (p. 1). The complexity of the interconnectedness of language and content is underscored by how students are supported linguistically and academically, based on the decisions made regarding assessment results. Unless educators are able to draw appropriate inferences, their ability to make decisions based on the results of EB students' performance is sharply reduced. Demographic and legal changes have created a greater sense of urgency around how EBs are assessed for achievement purposes and assessment validity.

There are numerous studies focused on EB student performance on assessments of academic achievement in public schools. Many educators believe that the lack of academic language in English negatively affects the EB student's scores on such assessments. Language assessment researchers (Bachman, 2002; Douglas, 2000) have been interested in the complex relationship between content knowledge and the language used to express that knowledge. This has become more and more relevant as large-scale assessments seem to be moving away from testing isolated content and skills toward a more integrated approach. Language is the primary vehicle through which knowledge is demonstrated, and in this case it is specifically content knowledge (Bachman, 1990). This is reflected on current assessments of Language Arts, Mathematics, and Science. For example, Wolf et al., (2008) found that standardized mathematics assessments in most U.S. states contain very few problems involving purely calculations; some degree of language is always involved in either presenting the problem or producing a solution.

This creates tension between the amount of language within state assessments, and the time it takes to acquire language proficiency.

Research states that acquiring English language proficiency for school-aged children can take approximately four to seven years (Abedi & Gándara, 2006). However, reaching a level of proficiency in language that would raise an EB students' academic achievement to the grade-level of their peers can take as long as eight to ten years (Collier, 1987, 1995). Differences in speed of English language acquisition may depend on a number of factors such as the age at which English is first taught, and how many years of formal schooling students received in their native language (Brown, 2000; Collier, 1995). This indicates that it can take many years for EB students to perform as well as native English speakers on a test written in English.

The use of national, state, and classroom assessments designed for native English speakers (non-EBs) have been shown to be problematic for EB students because they cannot access the test questions (Martiniello, 2008). If an assessment has unnecessary language, idioms, or words with double meanings associated with test questions, EBs may have a difficult time discerning what the question is asking. As a result, researchers express concerns about the accuracy of EB students' individual and group academic rankings for standardized math and science subject tests (Abedi, 2008; Abella, Urrutia, & Shneyderman, 2005; Solano-Flores & Trumbull, 2003). These rankings may be more reflective of an EB's English proficiency than their content knowledge.

Emergent bilingual students may find science tests difficult due to the academic language they contain; however, at this time, few studies have focused on EB students' science test performance in comparison to non-EBs' (Luykx et al., 2007; Wolf & Leon, 2009; Young et al., 2008). Studies comparing these two groups, specifically related to the influence of language proficiency, are essential because Colorado's goal is to increase interest and encourage pursuit of higher education in the fields of science, technology, engineering, and math (CEI, 2014) and many EBs are being left behind due to their lack of success on the state assessments. This study analyzed possible linguistic factors that may influence performance on the state science assessment.

Definition of Terms

The following definitions are offered to ensure understanding of the terms used within the context of this study.

"Academic language" is the language used in academic contexts.

The "ACCESS for ELLs," is an instrument developed by the WIDA consortium to measure English language proficiency. Colorado uses this instrument as their state English language proficiency assessment administered annually to EBs in January.

"Basic interpersonal communication skills," of BICS is the everyday social language as defined by Cummins (1979).

"Cognitive academic language proficiency," or CALP, is the language of school (Cummins, 1979).

The "Colorado Department of Education," or CDE, is the state governmental agency that oversees education.

The "Colorado Measures of Academic Success," or CMAS is the state summative academic assessment system. This study refers mostly to the CMAS for science.

"Constructed Response" or CR, is an item type composed of open ended, short answer questions that measure application-level cognitive skills as well as content knowledge.

An "Emergent bilingual," or EB, refers to a student whose primary home language is not English and who is not English proficient. In most literature, EBs are also referred to as English language learners (ELLs), English learners (ELs), limited English proficient (LEP), or language minorities (LM).

"L1" refers to first language.

"L2" refers to any language learned after the first (i.e., second, third, etc.).

"Productive language" refers to the speaking and writing domains of language.

"Receptive language" refers to the listening and reading domains of language.

"Second language acquisition," or SLA, refers to the subfield of linguistics that studies the acquisition of a language subsequent to the first language acquisition (Brown, 2000).

"Selected Response," or SR, is a multiple choice item type.

"Simulation" refers to an assessment item type where the student interacts with the item by observing and performing tasks associated with that item.

"Technology Enhanced Item," or TEI, is a computer-delivered item type and include specialized interactions for collecting response data. These include interactions and responses beyond traditional selected-response or constructed-response.

The "World-class Instructional Design and Assessment," or WIDA, is a consortium of US states and territories headquartered at the University of Wisconsin Center for Education Research (http://www.wida.us).

Chapter 2: Literature Review

This chapter provides a comprehensive overview of research and theory through the conceptual lens of equity, situated within second language acquisition, sociocultural, and assessment theories to approach the research questions (see Figure 1). The first section provides a brief overview of the three major schools of Second Language Acquisition (SLA). The focus then changes to Krashens'(1987) hypotheses, a conceptual framework on how language is acquired, and the hypothesis that contends there is critical period for SLA.

The second section provides an overview of sociocultural theory, specifically Vygotsky's (1978a) Zone of Proximal Development (ZPD). The focus then shifts to a review of the literature on academic language, the intersection of science, and the influences of a students' home language. It begins with Cummins' (1979) work around BICS and CALP, including the extension of this work by Bailey (2008) to include school navigational language and curriculum content language, and evolves into how academic language has been defined and characterized in the literature, specifically around science. Next is a discussion on the relationship between academic language proficiency and student achievement on high stakes assessments in science.

The third section introduces assessment theory, and the connection to sociocultural theory, focusing on issues of validity surrounding high-stakes assessments and EBs. It begins with a discussion of the appropriateness of assessing EBs with assessments normed on English proficient students, and ends with budding research around innovative technological item types that have the potential to level the playing field.

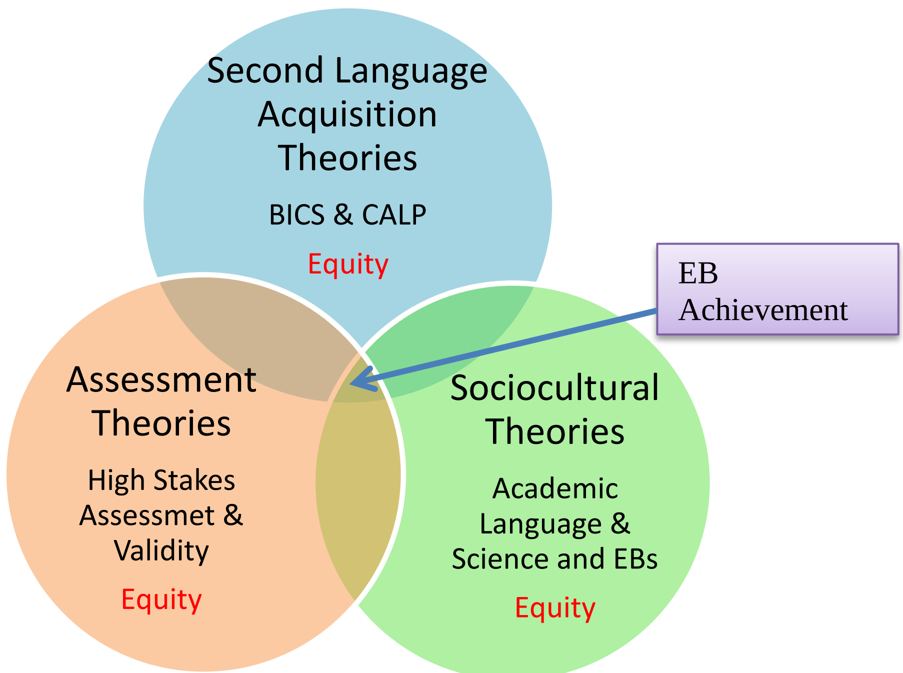

Figure 1: Theoretical Framework that illustrates the intersection of second language acquisition, sociocultural, and assessment theories using the conceptual lens of "equity" situated within each.

Second Language Acquisition

Second language acquisition is a very broad category of research that includes many disciplines. This part of the review will briefly describe the connections to first language acquisition, the main categories of SLA theory, the critical period hypothesis, and Krashens' (1987) hypotheses. Second language acquisition, although heavily researched, still evokes serious debate regarding certain ideas (i.e., explicit and implicit learning, intentional and incidental learning) (Brown, 2000).

Schools of Second Language Acquisition Theory. Schools of thought in second language acquisition have progressed from structural linguistics and behavioral psychology to generative linguistics and cognitive psychology to constructivism (Brown, 2000). These ideas are hardly new concepts; however, second language acquisition theories build upon those of first language acquisition and are broadly categorized into three main areas: behaviorism, innatism, and interactionism.

Behaviorism. The major theme behind behaviorism is that language is acquired through imitation and repetition. It is essentially learning language through habits. If students are provide with enough opportunities for imitation and repetition, they will learn the target language. This gave rise to the audio-lingual methods used in computer programs, such as *Rosetta Stone.* Although some aspects of behaviorism apply toward SLA, innatist and interactionist theories take precedence in the research community (Lightbown & Spada, 2004).

Innatism. The innatist school believes that humans are born with an innate ability to acquire language. Since human brains are hard-wired for language, learning a language is a naturally programmed ability. Language acquisition begins upon initial exposure, and then follows a prescribed sequence until language is acquired. Chomsky, a cognitive scientist, named an area of the brain hard-wired for language, the language acquisition device, which is widely accepted by Innatist and Interactionists (Brown, 2000).

Interactionism. Similar to the Innatist model, Interactionists believe in the innate ability to learn language; however more emphasis is placed on the environment in which the learning occurs. Piaget and Vygotsky's sociocultural theories of mental processing are drawn upon in this model (Lightbown & Spada, 2004). Interactionists propose that one's innate ability to acquire language is mediated by the interaction with others and enhanced by proficient speakers. Language within the environment can be linguistically modified to aid in acquisition; including contextual clues, slower speech rate, and total physical response (Lightbown & Spada, 2004).

Critical Period Hypothesis. Initially the critical period hypothesis was connected to first language acquisition (Lightbown & Spada, 2004); however, in recent years research has suggested that there is a connection of the critical period hypothesis to second language acquisition (Ioup, 2005; Moyer, 2004). The critical period hypothesis is the subject of a long-standing debate in linguistics and language acquisition over the extent to which the ability to acquire language is biologically linked to age. Lennenberg, an innatist, proposed the critical period, which states that if language acquisition does not begin before a certain age, full acquisition will not occur (Brown, 2000). Researchers, however, do not agree on when the critical period ends; the "classic" argument, one in which Lennenberg believes, is at the onset of puberty and others believe by the age of 5 (Brown, 2000).

Does this mean that second language acquisition would have the same or similar critical period? It is generally agreed that there is not a critical period for SLA, except when it comes to pronunciation (Collier, 1987a; Marinova-Todd et al., 2000) and accent

(Brown, 2000). To examine these further, researchers need to consider neurological and phonological issues (Brown, 2000); however, this issue has not been settled.

Collier (1987b) highlighted nine studies, conducted from 1962 to 1984, that demonstrated that EBs arriving between ages 8 and 12 acquired academic language in a second language faster than students who arrived younger. Older EBs may have an advantage over younger EBs in acquiring language at a faster rate because they have more fully developed first language upon which to build. Although older learners may acquire language at a faster rate, younger learners tend to be the ones who realize ultimate attainment (Brown, 2000; Long, 2007).

Krashen's Monitor Model based on five hypotheses. Steven Krashen (1987) developed a model of second language acquisition by building upon Chomsky's (1986) Language Acquisition Device and integrating his five hypotheses (Brown, 2000). Krashen's model begins with comprehensible input (input hypothesis), then applies a filter (affective filter hypothesis) prior to entering the language acquisition device. The language acquisition device processes the input to produce knowledge (natural order hypothesis), which is then monitored (monitoring hypothesis) using learned knowledge (acquisition-learning hypothesis) to produce linguistic output (Brown, 2000; Lightbown & Spada, 2004).

In the first hypothesis, acquisition-learning hypothesis, Krashen divides language into learned and acquired language for second language learners. Learned language is the result of a concerted effort and attention given to rules and form (Lightbown & Spada, 2004). In contrast, acquired language is the result of exposure, much like the natural process of acquiring a first language. Krashen believes that these two divisions of language are distinct.

The second hypothesis, the monitor hypothesis, involves productive language (speaking and writing) which is explicit and intentional learning. The learner has an internal monitoring system in which learned language "monitors" acquired language (Lightbown & Spada, 2004) and editing of one's output is consciously perceived (Brown, 2000).

The third hypothesis, natural order hypothesis, states that language rules are acquired in a predictable or "natural" order (Brown, 2000). In contrast to the monitor hypothesis, the natural order hypothesis does not use the learned language, only discusses the acquired language due to the natural processes that occur in acquisition.

The fourth hypothesis, the input hypothesis, involves receptive language (listening and reading). This hypothesis emphasizes comprehensible ($i + 1$) input and its importance in SLA, which is similar to Vygotsky's zone of proximal development (zpd) in first language acquisition (Lightbown & Spada, 2004). Since the CMAS for science is an assessment in English, comprehensible input contributes to EB performance on this assessment.

Krashen's fifth hypothesis, the affective filter hypothesis, states that the best acquisition will occur in environments where anxiety is low because other factors, such as mental dispositions, can raise barriers to language acquisition (Brown, 2000). This hypothesis has been criticized as untestable, especially with respect to establishing causality. However, motivation to learn, prior learning experiences impact how we learn language (Brown, 2000).

Sociocultural Perspectives on Learning

For schools to address the achievement gap, “proficient” must be defined in terms of the language demands of academic assessments and the lengthy process of becoming more able to meet those demands. Research suggests that the academic achievement of EBs in schools is inextricably tied to support for academic language development within socioculturally appropriate environments (Cook, et al., 2011).   Sociocultural theory is characterized by its consideration of individual engagement as being shaped by sociocultural processes on different planes of development (Nasir & Hand, 2006). Rogoff (1995) focused on sociocultural practices along three planes of analysis – participatory appropriation, guided participation, and apprenticeship – within three aspects of social interaction – personal, interpersonal, and community/institutional. Emergent bilingual students are constantly navigating their learning through multiple sociocultural planes, based on their level of English language proficiency, in order to move from peripheral participants in the community to more of a central membership (Lave & Wenger, 1991). As EBs become adept at fluidly moving between the three aspects of social interaction, they become more confident about their language learning, and subsequently, their participation in content classrooms.

Vygotsky’s (1978) sociocultural theory of learning posits that learning occurs through social interaction, and that the potential for cognitive development is limited to a “zone of proximal development” to explain that learning is not only a social experience, but one in which a student needs scaffolding to support their evolving cognitive understanding of concepts and skills. As EBs become more confident contributors within the content classroom, the scaffolding they receive needs to shift with the dynamic nature of language learning. As students develop the capacity to perform complex cognitive function with increasingly less reliance on external mediation, support for the development of their academic language need to increase (Bailey, 2007).

Academic Language. Educators are concerned about students' lack of academic language; however, many cannot define it, identify students who have it or do not have it, or provide specific ideas on how to help develop it (Freeman & Freeman, 2009). Schleppegrell (2004) states that not all features of academic language are present all the time. Rather, academic language has many features that constitute a language register (Schleppelgrell, 2004). In linguistics, language registers are varieties of languages used for a particular purpose or setting and are composed of lexical, grammatical, and discourse features that characterize specific uses of language (Schleppegrell, 2001; WIDA, n.d.). Although the research community does not have consensus on specific details, academic language is the register or discourse required for academic success (Abedi, 2008b; Bailey & Huang, 2011).

Sociocuturalists attribute differences in linguistic ability to students' access to styles of language use or registers related to their home language, emphasizing the role of socialization in academic language development (Gee, 1990; Rymes, 2010). Discourse, which is central to Gee's (1990) perspective, refers to "ways of being in the world or forms of life which integrate words…" (p. 142). Similar to registers, discourse has multiple, yet different meanings in linguistics. Gee (1997) divided discourse into two separate ideas: “little d” and “big D.” When discourse is in lowercase it refers to communicative language acts and Discourse, when it is capitalized, refers to a combination of language features and social practices used within a specific group or Discourse community. Other linguists however use discourse to mean any communicative act utilizing any and all language (Cazden, 2001).

Research suggests that the academic achievement of EBs in schools is inextricably tied to long-term support for academic language development within socioculturally appropriate environments (Bailey & Huang, 2011). More specifically, the research presented here has at least two implications: First, comparisons between English language proficiency and academic content proficiency measures must be part of the process that states use to define what English proficient means. Second, representations of the growth of English learners’ achievement must respect the fact that English learners grow at different rates. These growth rates are mediated by many factors; clearly, one is a students’ initial proficiency level. Research also points to other important variables that affect growth, such as student poverty and access to academic curriculum (Callahan & Gándara, 2004). Sociocultural perspectives are useful for interpretation of assessment results in that students’ abilities, as demonstrated on assessments, may not only be a reflection of their cognitive abilities, but also their socialization.

This study analyzed the relationship between academic language and content within the context of EBs’ performance on the CMAS for science. The ideas of register and discourse come into play while evaluating assessment items relative to English language proficiency levels and the four domains of language. The ability to communicate at high levels in science, demands a specific level of discourse, therefore, a strong command of academic language. One of the challenges educators face, especially as students get older, is recognizing that while some students who seem to speak English well are successful academically, others are not. The ability to distinguish social language from academic language is key, yet difficult. One of the consequences of advancing students who have a strong command of social English is they become long-term English learners (LTEL) (Batalova, Fix, & Murray, 2007; Fix, 2005, as cited in Freeman & Freeman, 2009) and these students have a difficult time catching up with their peers.

Cummins' BICS and CALP. Cummins created a theoretical framework to distinguish between social and academic language. He explained that he developed this "in order to draw educators' attention to the timelines and challenges that second language learners encounter as they attempt to catch up to their peers in academic aspects of the school language" (2008, p.71). Within his framework he discussed everyday conversational language which he named basic interpersonal communication skills. Cummins (2008) contrasted this with cognitive academic language proficiency to describe the language of school and further defined it as "the extent to which an individual has command of the oral and written academic registers of schooling" (as cited in Freeman & Freeman, 2009, p. 67). In other words, academic language is a specific language needed to understand and contribute successfully in a school setting.

Cummins does not suggest that academic language is superior to conversational language, just that they are distinctly different and a source of academic challenge for second language learners (Cummins, 2003). As Baker (2006) points out:

> School-based academic/cognitive language does not represent universal higher-order cognitive skills nor all forms of literacy practice. Different sociocultural contexts have different expectations and perceived patterns of appropriateness in language and thinking such that a school is only one specific context for "higher order" language production (as cited in Freeman & Freeman, 2009, p. 176).

Cummins further categorized language into four quadrants along with two intersecting continua to assist educators in conceptualizing the difference between BICS and CALP (Figure 2).

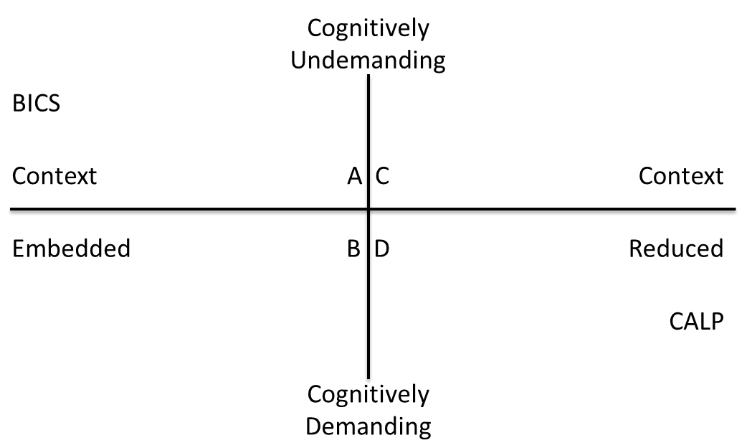

Figure 2. Cummins four quadrant Model. This model illustrates the intersection of context and cognitive demands of language. Source: Freeman & Freeman, 2009.

The horizontal line on the diagram represents a continuum that extends from context-embedded to context-reduced language. Cummins is explicit in his word choice and does not use *de-contextualized* for one end of the continuum since all language occurs in some context[9] (Freeman & Freeman, 2009). The vertical line represents the continuum of cognitively undemanding to cognitively demanding. BICS, therefore is contextualized and cognitively undemanding, and CALP is more cognitively demanding and less contextualized (context reduced). Boundaries are not always clear, as is the nature of continua (Bailey & Huang, 2011); so, it is important to keep in mind that there are various levels of attainment of BICS and CALP.

Bailey (2007) extended Cummins' work by discussing BICS as social language and CALP as academic language. She further differentiates AL into school navigational language and curriculum content language to better represent academic language as having some aspects of social language within the context of the school setting. She then applies a frame to include purpose, formality, context of use, context of acquisition, modality, teacher expectations, and grade-level expectations in order to clarify each category.

Characteristics of academic language.

> There is now general agreement that all students are learning to manage new sociocultural and language routines in classrooms and schools and that in each content area, students make use of specialized vocabulary, grammar, language functions and related discourse structures, and text types (WIDA, n.d.).

There is no question that academic language is one of the most important factors in school success (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006), and is a complex concept that is defined differently by researchers due to various philosophical and methodological perspectives (Gottlieb & Ernst-Slavit, 2009).

---

[9] There are some who have challenged the inclusion of context with language (Aukerman, 2007), claiming that it is deficit thinking and "ultimately destructive" (p.18).

Academic language can be considered through two different lenses: functional and communicative, and linguistic. The first lens can be categorized into general and discipline-specific language. General academic language is cross-cutting language that includes vocabulary and structures found within many disciplines (Schleppelgrell, 2001). Whereas, disciplines have their own particular academic discourse that includes content-obligatory vocabulary and syntactic structures (Bailey & Huang, 2011; Cook et al, 2011). The Colorado Department of Education uses the term “critical language” to represent both the general (academic) and discipline-specific (technical) language (CDEc, 2015). The Office of Standards and Instructional Support at the CDE highlights the importance of making explicit both types of language because learning content cannot be separated from academic discourse (Schleppegrell, 2001).

The second lens, linguistic, can be discussed at various dimensions: word/phrase (lexical), sentence (syntactic), and discourse (Bailey & Huang, 2011; Halliday & Hasan, 1989). Although presented separately, the dimensions overlap and influence each other.[10] At the word level, vocabulary can be divided into general, discipline-specific or specialized academic, and context-specific (Bailey & Huang, 2011; Scarcella, 2003). Context-specific academic vocabulary are words that have different meaning when used in various content areas – for example, the word *meter* has a different meaning (rhythmic structure) when used in music. There is a large body of research identifying academic language at the vocabulary level. Many educators are familiar with and use the Academic Word List, a compilation of general academic words developed by Coxhead

[10] Gottlieb & Ernst-Slavit (2009) share the example of “…a high school team debate. The specialized and technical academic words of the topic fold into specific grammatical structures, which in turn, shape the organization of a point-counter-point argument, backed by evidence, required of persuasion” (p.3).

(2000). There is disagreement, however, on the validity of the Academic Word List when used with school-aged children since it was developed using adult texts (Lawrence et al., 2010). In addition to just simply presenting words, consideration needs to be given to multiple meanings of words, nominalization, idiomatic expressions, and double entendres (Gottlieb & Ernst-Slavit, 2009).

At the sentence level, academic language is characterized by grammatical structures (syntax), conventions, and language forms, which are primarily found in assessments, textbooks, and classroom-based tasks (Bailey & Huang, 2011; Gottlieb & Ernst-Slavit, 2009). Since learning and understanding grammatical structures facilitates English language development (Fisher, Rothenberg, & Frey, 2007), it is important to consider that there are some aspects of academic language that are not intuitive and even illogical.[11] Other areas to consider are types of sentences (i.e., simple, compound, complex), types of clauses (i.e., relative, coordinate, embedded), and prepositional phrases (Gottlieb & Ernst-Slavit, 2009).

Snow and Uccelli's (2009) academic language model at a discourse level included discipline-specific genres, reasoning, taxonomies, and salient relations. At this level of academic language, it is important to include general language functions[12] (Chamot & O'Malley, 1994) and, those that are specific to science discourse[13] (Lemke, 1990). Schleppegrell (2004) states, that disciplinary-specific discourse is "a set of registers

---

[11] Think about the following examples: There is no "ham" in hamburger; a "slim" chance and a "fat" chance mean the same thing.

[12] Some language functions include: explaining, informing, classifying, debating, and evaluating.

[13] Science discourse examples: hypothesizing, questioning, designing, analyzing, and modeling.

through which students will be expected to learn and participate as they move through grades" (p. 411). Each academic discipline has specific discourse or communicative functions (Bailey, Butler, Stevens, & Lord, 2007; Schleppegrell, 2004), and each of these functions is related to grammatical and communicative rules and organizational patterns to accomplish content and specific purposes. When teachers make language functions explicit, they define more fully the tasks that students must be able to perform in the content areas (Chamot & O'Malley, 1994). The Colorado English Language Proficiency standards, the WIDA-created standards, do just this for teachers. They make explicit the language functions, forms, and conventions necessary to understand the content within a given context which contributes to students' language proficiency (Echevarría, Short, & Powers, 2006) and content area performance (Chamot, & O'Malley, 1994).

Students' language proficiency is measured within four domains of language: listening, speaking, reading, and writing. Traditionally, educators and researchers divide these four domains into two sub-categories, oral language (listening and speaking), and literacy (reading and writing). There are other ways to combine the language domains in order to make specific inferences about a students' language proficiency. The primary one being, "overall", which combines all four domains (either equally or weighted). This tends to be the marker that is used most on high stakes assessments for second language proficiency. Another way, although not as discussed in the literature is "receptive," (listening and reading), and "productive"[14] (speaking and writing). "Receptive" measures how well students receive and understand information, and "productive" measures how students produce and communicate that understanding. Researchers tend to focus on

[14] Productive is also referred to as "expressive" in the literature.

overall ELP, therefore there are not many studies focusing on the receptive and productive elements of language specifically. Hence the need for this study.

Science and Emergent Bilinguals. Short and Fitzsimmons (2007) state that "English language learners must perform *double the work* of native English speakers in the country's middle and high schools" (p.1). Emergent bilinguals are faced with learning English at the same time they are learning science. I would also say that in addition to learning English and content, students must also learn another language, the *language of science*. Norris and Phillips (2009) stated that teaching science content is not enough; teachers must focus on developing disciplinary literacy in science. Historically, discipline-specific literacy has been absent in the science classroom (Norris & Phillips, 2009) and science teachers often lack the skills to move beyond technical vocabulary and allow students to access scientific discourse (Snow, 2010). The CDE's inclusion of general academic language within their definition of "critical language," brings to the forefront the need for educators to be explicit in their teaching of both general and technical language in science. Snow (2010) supports the need to incorporate explicit instruction of general academic discourse into the science curriculum.

Scientific concepts and language are often abstract, and this level of abstraction goes beyond the word level to the sentence and discourse levels. In addition, scientific texts are difficult for EBs to access due to the inclusion of many prepositional phrases, noun phrases, and passive voice constructions (Bailey, 2007). However, with the allowance for time and support, science provides a good context to support academic language development (Hakuta, Butler, & Witt, 2000).

The Next Generation Science Standards (NGSS) (NGSS lead states, 2013) emphasize scientific discourse as a vehicle for understanding scientific ideas, and demand the use of science language to communicate understanding (Krajick, 2013; Lee, Quinn, & Valdés, 2013), including constructing explanations and engaging in argument from evidence. The linguistic demands of these standards are high for all students, especially for EBs. However, the intent of the NGSS is not to think of the language of science as a system of rules to be learned, but rather, as Lee et al. (2013) argued:

> …students must develop and understand the linguistic tools for meaning-making in science as comprising a unique linguistic register. This register provides tools for understanding what people are doing, what their relations are to each other, and how they are using language in the context of making scientific meaning (p. 226).

Meaningful understanding develops as students incorporate tools developed to assist with disciplinary literacy. Lemke (1990) argued that the language of science should be seen as "differentiated speech" and students need to build connections across differentiated speech forms, from everyday language to disciplinary discourse. By engaging students in sophisticated language functions (i.e., arguing from evidence, providing explanations) they are able to make growth in conceptual science and in language proficiency (Lee et al., 2013).

Generally there are two researched-based approaches that explicitly integrate content and language: content-based language instruction (CBLI) and sheltered English instruction (SEI). The intention of CBLI, originally situated in ESL classes to integrate the teaching of language and the teaching of academic subjects (Scarcella, 2003; Schleppegrell, 2004; Snow, 2001), was to provide students with increased opportunities to experience larger discourse-level features and social interaction patterns essential to language use (Lee et al., 2013). However, ESL teachers' content knowledge in multiple content areas was proving to be inadequate. More recently, CBLI shifted to a "sheltered" model (SEI) where content classes for EBs are taught by content-area teachers who have had some training in language pedagogies and use language objectives in addition to content objectives (Echevarria & Short, 2006; Echevarria & Vogt, 2008).

Lee et al., (2013) argue for two shifts away from this thinking,

> … (a) a shift away from both content-based language instruction and the sheltered model to a focus on language-in-use environments and (b) a shift away from 'teaching' discrete language skills to a focus on supporting language development by providing appropriate contexts and experiences (p. 226).

This new envisioning incorporates the demands for the NGSS and allows teachers to create classroom environments that are rich in scientific discourse where students engage in science and engineering practices, such are argumentation with evidence, explanations of phenomena, and the claims, evidence, and reasoning framework.

Students' Home Language. Research on the relationship between language use in the home and EBs is primarily around literacy development in their first or second language and is almost entirely correlational (August & Shanahan, 2006). However, most research in this area focuses mainly on elementary students in the area of reading. There have been a few studies (Brasel, 2008; Lindholm-Leary & Hernández, 2011) that investigate older students and the findings are mixed and suggest less parental influence. Parental influence may be stronger for younger than for older children because younger children generally cannot read the kind of text that will contribute to language development and thus are dependent on parents (Howard et. al., 2014). For older children, the quality and type of schooling may override parental influence.

Surprisingly, there are very few studies investigating differences between Spanish-speaking EBs and EBs who speak other home languages. The studies that include this type of variable are specifically looking at linguistic complexity or reading (Reese et. al., 2006). Also, as stated previously, the focus of research in this area is at the elementary level, not the middle school. Therefore, results may not be generalizable. There are two areas of research focusing on secondary level students and their home language as a variable. The first is around the idea of the “model minority” and the second is around the “Good Language Learner.”

Asian students, specifically those who are EBs, are termed the “model minority” in literature due to the nature of how high they score on standardized assessments versus other EBs with various language backgrounds. Speakers of Asian languages comprise the second largest group of EBs in the United States, eight percent of the EB population, while Spanish-speaking EBs comprise of the largest group of EBs at 80 percent (Goldenberg, 2008). Educational researchers continue to unravel the model minority stereotype (Conchas & Perez, 2003) citing studies in the anthropology of education about the stereotype, and engage with the discourse relevant to studies connecting students’ educational experiences with identity and achievement (Lee, 2009). Despite the various backgrounds and experiences of Asian students (i.e., Burmese refugees), this stereotype persists in schools and can not only impact non-Asian students who do not score high on high stakes assessments, but also the Asian students who do not score high as well.

Contrary to the model minority stereotype for Asian students, Spanish-speaking students get characterized as non-achieving due to lower scores on high stakes assessments. The fact that this subgroup of EBs is the largest in the nation is important to keep in mind, since Spanish-speakers in the United States tend to come from lower economic and educational backgrounds than other language minority populations (Goldenberg, 2008). Consequently, a large number of EBs are at risk for low assessment scores not only because of language, but also because of socioeconomic factors, immigration factors, or refugee status.

The idea of a "Good Language Learner" (Norton & Tooney, 2001) is grounded in sociocultural theory which assumes that "language and development occur as people participate in the sociocultural activities of their community" (Rogoff, 1994, p. 209). Vygotsky (1978) emphasized the importance of social contexts in the process of acculturation, whereby "intellectual tools of society" (Rogoff, Mosier, Mistry, & Göncü, 1993, p. 232) are brought to bear by more experienced participants in a culture to support less experienced members. There is much literature about motivation, parental education level, and parental involvement within the area of language development that could provide an interesting intersection here, however when people focus more on Spanish-speaking students and the achievement gap, it is important to consider that there may be more variables that come into play when thinking about the idea of a "Good Language Learner."

One such variable may be based on Ogbu's (1983) theory of voluntary versus involuntary minorities and ideas around motivation and acculturation. He used his idea of *community forces* to hypothesize why immigrant minorities did well in school and non-immigrant minorities did less well. Sociocultural adaptations seemed to be part of the reason for this. Non-immigrant minorities were "historically denied equal educational opportunities in terms of access to educational resources, treatment in school, and rewards in employment…" (Ogbu, 1983, p. 157). Multigenerational students seemed to have internalized this marginalization as demonstrated through their academic performance. Salazar (2008) eludes to this internalized marginalization through the use of the term *maleta* (suitcase). She asserts that students leave their *maletas*, filled with their language, culture, and ways of knowing, at the schoolhouse doors in favor of the dominant culture. This may be something to explore further when discussing primary home language.

Assessment Theory

There are many factors shaping educational assessment policy in the United States: standards-based education and the demand for accountability. Pellegrino (1999) asserts that American education is in a period of high expectations for all children, which inherently demands equity and excellence. Subsequently the demand for accountability follows. The current rhetoric around improving the educational system is driving this demand for accountability, and the response has been through assessment. If assessments are to assist in the improvement of the educational system, then, as Pellegino (1999) states, "If social and public goals regarding academic achievement are to be attained, then we must make more effort to improve assessment…" (p. 5). Thus, improving assessments may lead to an increased balance between accountability and assessments.

High Stakes Assessment. High stakes assessments are a large part of the educational landscape within the United States and are used as a means for accountability. Some proportion of the achievement/opportunity gap may be due not to an EB student's lack of content knowledge but to the content assessment's inability to accurately measure that knowledge when insufficient language proficiency stands in the way. Research has demonstrated that an assessment given in the English language is an assessment of English, even if it is a content assessment such as a science test (Abedi, 2004). This is the construct irrelevant variance that has been identified in research as a major threat to the validity and therefore potential usefulness of assessments of the content knowledge of EB students (Abedi, 2002; Haladyna & Downing, 2004). If the language proficiency level of an EB student is insufficient for the student to understand the language of a science assessment, for example, then the assessment may, in part, be measuring the wrong construct (i.e., measuring language knowledge rather than science knowledge). Therefore, one could say, that language has become construct-*relevant;* especially in the era of new content standards. Unfortunately, the distinction between communication and unnecessary linguistic complexity may have become less determinate with the new standards. Both the CCSS and the NGSS ask students to use various communication structures (e.g., argumentation with evidence) that increase linguistic complexity which could be problematic for developing assessments that truly measure the intended construct, content knowledge. Therefore, one could argue that when academic assessments are constructed in a way that language appears to become *construct-relevant,* then a validity argument must be made.

Issues for Emergent Bilinguals and Assessment. Prior to NCLB (2001) legislation, ELP assessments of EBs were solely intended for programmatic purposes (teaching English as quickly as possible with a focus on oral language development); not accountability, as they are today. The legislation asserts that states and districts must hold the same high standards for EB students as they do for all other students, and that they are accountable for assuring that all students meet high expectations. The assumption here is that the required ELP assessment and content-based assessment interact to produce the overall desired outcome of successful academic achievement for EBs. This assumption poses major consequences for the student and, under federal accountability measures, for the district and the school as well (Solorzano, 2008). Federal accountability, Annual Yearly Progress (AYP), only includes mathematics and English language arts, however, under Colorado law, districts are also held accountable for science through their accreditation measures.[15] This section discusses the assessment of EBs on ELP and content-based standardized tests. Issues of validity and reliability are discussed within the frame of construct-irrelevance due to potential differential item functioning. Bailey and Carroll (2015) state:

> States must expect that all educators will hold ELL students to high Academic content standards. However, when students are being assessed for content knowledge in a language they are still learning, fair and valid (i.e., meaningful) interpretations depend on clear measurement of the construct (e.g., avoiding irrelevant construct variance caused by measuring language abilities rather than the intended mathematics or science knowledge) and appropriately implementing testing accommodations (p. 255).

[15] Colorado District accreditation measures include the District performance frameworks (DPF) and School performance framework (SPF) (CDE website, 2014).

ELP assessment of EBs. The NCLB (2001) requires annual English language proficiency (ELP) assessment aligned to ELP standards, intended to measure an EB students' progress in learning English. The results of these assessments are intended to be at the "macro" level, since the assessment is a blunt instrument being used at one point in time, thus inferences drawn from these results may be inadequate for instructional purposes (Bailey & Carroll, 2015).

English language proficiency assessments have undergone considerable revision in recent years in order to create test items that reflect the academic uses of language at the K–12 level (Bailey & Carroll. 2015). Language proficiency is divided into four domains, listening, speaking, reading, and writing, and students acquire them interdependently, but at different rates and in different ways (Spolsky, 1989). Some people believe that academic language is composed mainly of the reading and writing domains of language due to the link to literacy and report them as a separate "comprehension" score on ELP assessments (WIDA, 2015). However, others argue that academic language is composed of all four domains, but may not contribute to proficiency equally (Sato, 2010). These differences in contribution may be considered independently and/or in pairs (i.e., receptive and productive, oral and comprehension). The added emphasis on communication within the CCSS and NGSS adds even more importance to the inferences drawn from ELP assessments to reflect these academic uses of language.

Content-based assessment of EBs. Historically, EBs have underperformed on standardized assessments compared to their native English speaking peers (Abedi, 2002, 2009; Abedi et al., 2004; Abedi & Dietel, 2004; Cook et al, 2011; Solorzano, 2008). Although this historical record paints a bleak picture of EBs, their performance may not be an accurate representation of their content knowledge (Abedi, 2002, 2008b; Abedi & Gandara, 2006; Martinello, 2008). Standardized assessments are normed on the majority, native English speakers, and as such become de facto assessments for content and academic language (Abedi, 2002; 2008b; Solorzano, 2008). Although developed to assess only one construct, content knowledge, the use of standardized assessments is being questioned based on a validity argument around accountability for EBs. The American Educational Research Association (2000) states, that "an assessment should not be used with a student who does not understand the language of the test" (as cited in Solorzano, 2008, p. 262). This statement was made prior to the enactment of NCLB, which goes to show that there is a research to policy gap.

One could assume that any student taking the standardized assessment who is designated as an EB indicates that the student does not understand the "language of the test," therefore could encounter potential problems in understanding the language of the assessment. Although research shows that other factors such as socioeconomic status and parent education level impact EBs achievement, language has the greatest impact (Abedi, 2002, 2008b, 2009; Abedi et al., 2004). As language demands increase in the new academic content standards (CCSS and NGSS), EBs may increasingly have problems accessing assessments at all three levels (word, sentence, and discourse). Characterizing these inherent language demands, although challenging, will be necessary to support instructional practices and align EB assessments to the new academic content standards (Bailey & Wolf, 2012).

Bailey and Carroll (2015) noted a recent approach to identify the "Key Practices and Disciplinary Core Ideas" in the new content standards and the receptive and productive language functions that likely will be required to carry out these practices (*ELPD Framework*, CCSSO, 2012). They noted, "This approach includes high-level descriptions of language uses rather than attempting to specify discrete language structures that provide a foundation for language" (p. 269). This means that students will need specific discourse skills (i.e., stating a claim, constructing an argument, etc.), critical language to support those skills, and knowledge of sentence structures to communicate their understanding. Assessment vendors need to take note of these ideas and incorporate them into new assessments that would be a fairer representation of an EBs academic achievement. Teachers will also need professional development to support instruction that addresses the tasks and uses of language that are more abstract and which focus on discrete knowledge.

Chapter 3: Method

This chapter describes the research methodology, including: data sources, participants, instruments, and method of analysis. This dissertation study focused on examining the factors that may influence academic achievement of Emergent Bilinguals (EBs) on a high-stakes science assessment in general and within specific content domains.

The first research question was simply: What factors predict EBs' performance on the Colorado Measures of Academic Success (CMAS) for science assessment? The variables, by matched student identifiers in Colorado schools for 8th grade EBs in a regression analysis, were CMAS science overall scale scores, ACCESS overall scale scores, computed receptive and productive scores, and students' home language. Hierarchical multiple regression was used to conduct the analysis: block one used SES as a control variable since research has identified this as a predictor of achievement, block two was students' home language, block three was English language proficiency, and block four was receptive and productive levels of proficiency. Colorado Measures of Academic Success science overall scale score was the dependent variable.

The second research question was: What factors predict EBs performance on specific content domains within the CMAS for science assessment? Colorado Measures of Academic Success for science scale scores for the three content domains (physical, life, and earth sciences), the WIDA-ACCESS overall scale scores and computed receptive and productive scores, and students' home language by matched student identifiers in Colorado schools for 8th grade English learners served as variables. Hierarchical multiple regression was used to conduct the analysis: block one used socioeconomic status as a control variable since research has identified this as a predictor of achievement, block two was students' home language, block three was English language proficiency, and block four was receptive and productive levels of proficiency. CMAS domain scores on physical, life, and earth sciences served as the dependent variables in three separate regression analyses.

Research Questions

My research sought to answer the following questions based on previous findings and recommendations:

1. What factors predict EBs performance on the Colorado Measures of Academic Success (CMAS) for science assessment?
    a. To what extent did the students' primary home language predict performance on the 8^{th} grade CMAS for science?
    b. Beyond the students' home language, to what extent did the level of English language proficiency predict performance on the 8^{th} grade CMAS for science?
    c. Beyond students' home language and the level of English language proficiency, to what extent did receptive and productive elements of language predict performance on the 8^{th} grade CMAS for science?

2. What factors predict EBs performance on specific content domains within the Colorado Measures of Academic Success (CMAS) for science assessment?
    a. To what extent did the students' primary home language predict performance on the 8th grade CMAS for science?
    b. Beyond the students' home language, to what extent did the level of English language proficiency predict performance on the 8th grade CMAS for science?
    c. Beyond students' home language and the level of English language proficiency, to what extent did receptive and productive elements of language predict performance on the 8th grade CMAS for science?

Data Sources

The Colorado Department of Education collects data from students and school districts in various ways throughout the year, only after being approved by an Educational Data Accountability Committee and following the rules of Family Educational Rights and Privacy Act. The data used for this analysis were obtained through a Student Biographical Data Grid used on the two state-level assessments, CMAS for science and ACCESS for ELLs. The individual data were masked using a unique student identifier to protect the identity of students. Matched student identifiers were used to conduct secondary data analysis. Data were received through a secure file transfer from the CDE.

Participants

Emergent bilingual students in Colorado are divided into subgroups: NEP (Non-English Proficient), LEP (Limited English Proficient), FEP (Fluent English Proficient), FELL (Former English Language Learner), and PHLOTE (Primary Home Language Other Than English). NEP, LEP, and FEP are a part of the Colorado Revised Statues as official language designations for students who are learning English as a second language and are receiving extra program support. FELL and PHLOTE are used for students who are not receiving extra support services, but whose language development is influenced by another home language other than English.

Eighth grade students who took the 2015 CMAS for science assessment are a part of the sample for this study (see Table 2). Additionally, from that group of students, EBs coded as NEP and LEP who took the ACCESS assessment are further analyzed. EBs selected for this study include students with special needs, students in gifted and talented programs, immigrant and migrant students, and second or third generation Americans.

Only EBs coded as NEP and LEP who took both the CMAS for science and the ACCESS assessments were included in the sample. It is important to note that, while this sample consists of Spanish-speaking EBs, different dialects of Spanish are represented within the sample but were not specifically identified for these analyses. Socio-economic status was used to control some of the variability within the group.

Table 2

Demographic Frequencies and Percentages of Sample Population

	2015 CMAS # (%)	*2015 ACCESS # (%)*
	N = 64,104	*N= 6,402*
Native American/Alaskan	513 (0.8%)	18 (0.1%)
Asian	2029 (3.2 %)	390 (0.6%)
Black African-American	2966 (4.6%)	226 (0.4%)
Hispanic Latino	20,959 (32.7%)	5860 (9.1%)
White	34,976 (54.6%)	163 (0.3%)
Hawaiian/Pacific Islander	153 (0.2%)	55 (0.1%)
Multi-Racial	2232 (3.5%)	19 (0.1%)
Unknown	276 (0.4%)	
Males	32,637 (51%)	3,672 (5.7%)
Females	31,467 (49%)	3,059 (4.8%)
FRL	26,579 (41.5)	
Non-FRL	37,525 (58.5)	
HL-Spanish	10,263 (16%)	5,884 (9.2)
HL-Other	53,841 (84%)	518 (0.1%)
EB	13,242 (20.7%)	6,402 (10%)
Non-EB	50,862 (79.3%)	0

Note. FRL = Free/Reduce Lunch; Non-FRL = No Free/Reduced Lunch; HL = Home Language, EB = Emergent Bilingual; NEP = Non-English Proficient; LEP = Limited English Proficient.

Instruments

Colorado requires all public school students to take a standards-based summative assessment in science at the 8th grade level. Every student, regardless of language background or ability, must be provided with the opportunity to demonstrate their content knowledge. In addition, Colorado law requires an annual assessment of English language proficiency for students identified as non-English proficient and limited English Proficient. This study includes two state-level high-stakes assessments, the Colorado Measures for Academic Success for science and the Colorado English Language Proficiency assessment (WIDA-ACCESS).

CMAS for science. The CMAS for science is Colorado's standards-based assessment designed to measure the Colorado Academic Standards in Science. Each assessment is comprised of three sections and all sections contain a combination of selected-response items, technology-enhanced items, and constructed-response items (CDE, 2015). A subset of the Science assessment includes simulation-based item sets, which are groups of items that all relate to a scientific investigation or experiment. Students use the information in the simulations and in the items to answer the questions or respond to the prompts. The simulation based items may be selected-response items, technology-enhanced items, and constructed-response items. Table 3 lists the item features of the 2015 assessment.

Table 3

CMAS Science Assessment Features for 2015

Item Features	# of Items
SR	28*
TEI	15
CR 2-point	14
CR 3-point	3
Total items	60

Note. *SR reported as 43; however TEI is 25% of total item count.

Assessment items are created by Colorado teachers and the assessment vendor, reviewed for alignment to state academic standards, undergo a review for bias and sensitivity by committee, are field tested, and are analyzed through item data review prior to administration in an operational assessment. Selected response and technology enhanced items each are worth one point, whereas constructed response items could be scored using a two-point rubric (scores ranging from 0-2) or a 3-point rubric (scores ranging from 0-3). Cronbach's alpha for the 2015 overall assessment was reported as 0.93 while alpha's for the different content domains: physical science, life science, and earth science were 0.82, 0.81, and 0.83 respectively. Interpretations of the CMAS scores were validated using various sources of validity evidence: evidence based on test content, evidence based on response processes, evidence based on internal structure, and evidence based on fairness (CDE, 2015).

A blueprint of the assessment was developed with specificity at multiple levels to optimally measure the CAS and each item underwent various levels of review to confirm alignment. In addition, field tests and Differential Item Function analyses were conducted to identify items that may be measuring a dimension unrelated to the intent of the construct (CDE, 2015). Cognitive labs and Adjudication were conducted to validate that students were responding as expected and items were being scored as expected (CDE, 2015). Factor analysis and scree plot examinations were conducted to identify the number of dimensions the assessment seemed to be measuring. Based on these results, a unidimensional Item Response Theory model was used in calibration and scaling. Lastly, as evidence for fairness, a practice environment was created with item types that would be on the assessment, so teachers and students could practice answering questions using an on-line platform. In addition, *Universal Design* principals were used during the creation of the assessment and differential item functioning analysis was conducted after the field tests to determine which items would be operational.

CMAS reports provide information on student performance in terms of scale scores, performance levels, and percent correct scores. This study investigated scale scores by the overall test, by content domain, and by item-type.

ACCESS for ELLs Tool. The overarching purpose of ACCESS for ELLs tool is to assess the developing English language proficiency of EBs in Grades K–12 in the United States following the English Language Development Standards (2012) of the multi-state WIDA Consortium (WIDA, 2015). ACCESS for ELLs Tool is an English language proficiency test designed to measure English language learners' social and academic language proficiency in English. It assesses social and instructional English as well as the language associated with language arts, mathematics, science, and social studies within the school context across the four language domains (Listening, Reading, Writing, and Speaking).

Performance indicators (PIs) describe the expectations for EB students for each of the five standards, at five different grade-level clusters, across four language domains, and at the five language proficiency levels. The ACCESS assessment is based on the 80 strands, containing 400 individual PIs, within the WIDA ELD Standards. Each selected-response item or performance-based task on the ACCESS assessment is carefully developed, reviewed, piloted, and field tested to ensure that it allows students to demonstrate accomplishment of the targeted PI. Figure 3 illustrates an example of an on-line writing released item shown through a series of screen shots.

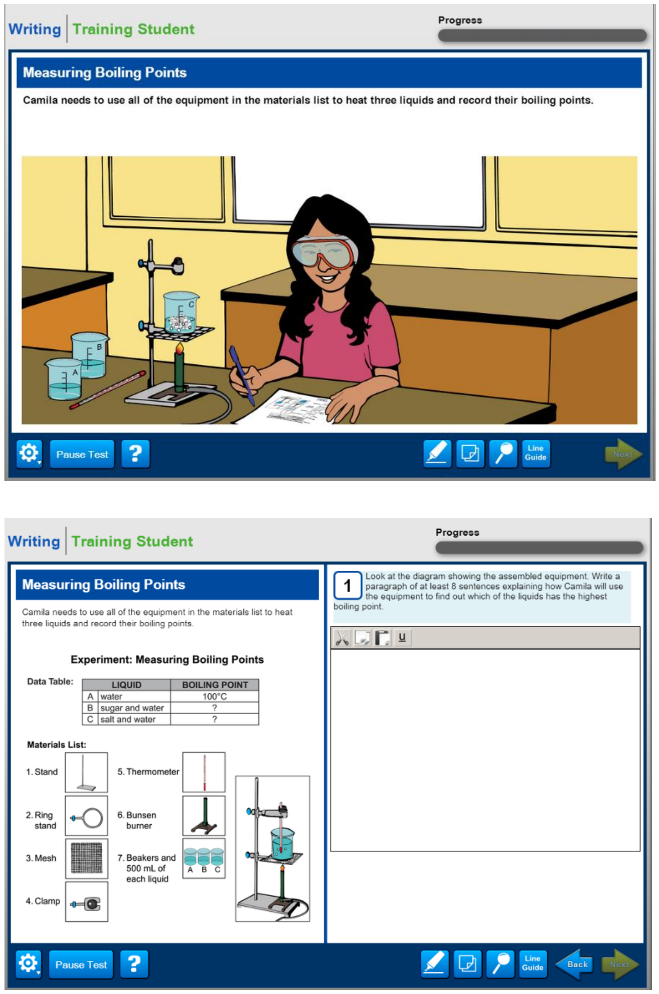

Figure 3. Sample on-line writing task for 6-8 grades (WIDA, 2015)

Because ACCESS for ELLs is a tiered test, each form in Tier A, B, or C targets only a certain range of the entire ability distribution, results for reliability on any one form, particularly for the shorter Listening test, may at times be lower than typically expected (WIDA, 2015). Cronbach's alpha was reported by form and by language domain. Table 3 reports the reliability coefficients for the 6-8 grade span assessment for individual language domains and form.

Table 4

Cronbach's alpha coefficients by language domains and assessment form

Language Domain	Form*		
	A	B	C
Listening	0.75	0.66	0.61
Speaking	0.90		
Reading	0.78	0.78	0.76
Writing	0.89	0.94	0.92

Note. *The Speaking domain only uses one form.

Four composite scores are also reported for the assessment: oral (listening and speaking domains), literacy (reading and writing domains), comprehension (listening and reading domains), and overall (listening, speaking, reading, and writing). A stratified Cronbach's alpha coefficient was used to compute, and weight by the contribution of each domain score to determine the composite. Table 4 lists the stratified Cronbach's alpha coefficients for the four composites.

Table 5

Stratified Cronbach's Alpha Coefficients for Four Composite Scores

Composite	
Oral (50% L and 50% S)	0.86
Literacy (50% R and 50% W)	0.90
Comprehension (30% L and 70% R)	0.83
Overall (15% L, 15% S, 35% R, & 35% W)	0.93

The ACCESS for ELLs assessment uses the CAL's Validation framework, which incorporates Evidence Centered Design and an Assessment Use Argument, an argument-based approach to addressing validity (Bachman & Palmer, 2010, as cited in WIDA, 2015). Figure 4 illustrates the validation framework used for this assessment, as documented in the 2015 WIDA Technical Manual.

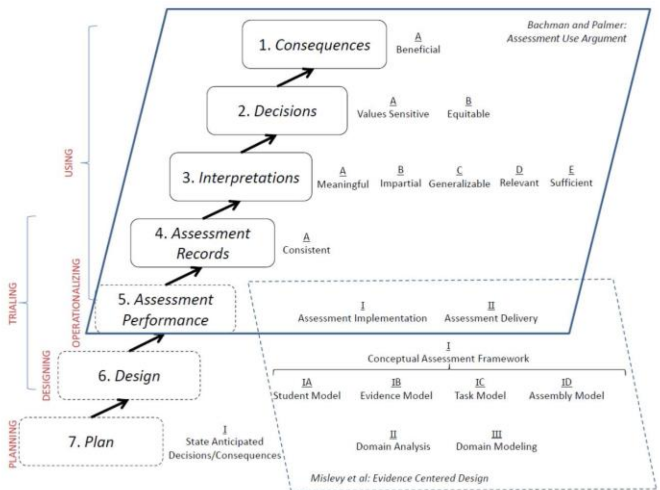

Figure 4. CAL's Validation Framework (based on Bachman & Palmer, 2010)

ACCESS reports provide information on student performance in terms of scale scores, proficiency levels, and percent correct scores. This study investigated scale scores overall, and by language domain and receptive and productive domains.

Method of Analysis

Variables. The analysis included performance and demographic variables. The dependent variable for research questions one and two was student performance on the 8th grade CMAS for science (overall scale score and scale score by content domain, respectively). The independent variables for both research questions was student performance on the ACCESS assessment for English language proficiency (overall scale score), student performance on the ACCESS assessment based on receptive (reading and listening) and productive (writing and speaking) domains of language (scale scores), and students' home language (coded as "1" for Spanish and "0" for all other languages). Student socio-economic status (FRL and Non-FRL) was included as a control variable.

Missing data (systematic or random). The data file received from the CDE was complete in that they only provided data, based on matched student identifiers, for those 8^{th} grade students who had complete records and took both the CMAS for science and ACCESS for ELLs.

Computed variables. There was one instance in which a variable was computed from the ACCESS data, specifically the "productive" composite variable (two of my research questions ask about both the receptive and productive elements of language). The WIDA reports the "receptive" composite variable as *comprehension*, but does not calculate or report the "productive" composite variable. Therefore, the variable was created by combining the speaking and writing domains of language based on the weights that WIDA used for each domain per their 2015 technical manual (speaking = 30%, and writing = 70%) (WIDA, 2015).

Statistical Analyses and Effects of Violations of Assumptions. Statistical analysis included descriptive statistics (frequencies, measures of central tendency, skewness and kurtosis), graphs, tests of assumptions, and hierarchical multiple regressions. SPSS 20 was used for all the analyses.

Hierarchical Multiple Regression. Hierarchical multiple regression is used to evaluate the relationship between a set of independent variables and the dependent variable, controlling for or taking into account the impact of a different set of independent variables on the dependent variable. Variables are entered into "blocks" in a fixed order of entry to control for the effects of covariates and to test the effects of certain predictors independent of the influence of others. For research questions one and two, block one was SES as a control variable since research has identified this as a predictor of achievement, block two was students' home language, block three was English language proficiency, and block four was receptive and productive levels of proficiency.

Assumptions for this statistical test are normality, linearity, homoscedasticity[16], independence of errors, and multicollinearity. The minimum sample size rule 5-to-1 was met; the sample was large. Normality was measured using the criteria of a -1.0 + 1.0 range for allowable skewness. After reviewing residual plots, the assumption of linearity was met. Outliers were checked by examining the standardized residuals and use the +/-3 rule examined the plot of these residuals to check for homoscedasticity. The check for independence of errors was used to determine that residuals were independent using Durbin-Watson between 1.5 and 2.5. Lastly, tolerance levels were investigated for multicollinearity for all independent variables to be greater than 0.10. This assumption was not met. The results section includes specifics about assumptions analysis, as well as results of the hierarchical multiple regressions for both research questions.

[16] This assumption means that the variance around the regression line is the same for all values of the predictor variables.

Chapter 4: Results

This chapter addresses the results of each research question, including analysis of assumptions.

Predictors of Overall Science Achievement

A hierarchical multiple regression was conducted to predict academic achievement on the Colorado Measures of Academic Success (CMAS) for science based on primary home language, language proficiency, and receptive and productive elements of language. Analysis began with evaluation of assumptions. Multiple regression assumptions include assessment of normality, linearity, and homoscedasticity. Normality was determined by inspecting the skewness, kurtosis, mean, median, mode, and histograms of predictor variables and the error scores. Model linearity was assessed using plots between standardized predicted values and standardized residuals, revealing no obvious divergences. Homoscedasticity was evaluated using scatterplots between standardized predicted values and standardized residuals. A scatterplot indicates homoscedasticity when the band that encloses the residuals is about the same width for all values of the predicted criterion variable (Tabachnick & Fidell, 2007). The model's scatterplot suggested general homoscedasticity. Multicollinearity was evaluated using minimum tolerance level of 0.10 (Tabachnick & Fidell, 2013) and VIF maximum tolerance level of 10 (Hair, Anderson, Tatham, & Black, 1995); the VIF recommendation of 10 corresponds to the tolerance recommendation of 0.10 (i.e., 1 / 0.10 = 10). This assumption was violated for step four in the full regression model. The violation occurred between the overall English language proficiency (ELP) score and the receptive and productive elements of languages scores. This is due to the receptive and productive elements of language being inherently within the overall English language proficiency. Therefore, to correct this violation, a three-step model was run using the original three steps in the method as outlined previously, and then a second three-step model was run using the original two steps as outlined, but then I replaced overall ELP in step three with the receptive and productive elements of language as a new step three.

Next, hierarchical multiple regression analysis was performed between overall achievement on the CMAS for science as the criterion variable and SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and overall English language proficiency in the third block as predictor variables. Then an additional hierarchical regression was performed, including steps one and two as stated above, and step three included receptive and productive elements of language as the predictor variables. Table 6 displays effect size measures (R^2), change in R^2, and adjusted R^2 for the full model, and Table 7 displays pooled unstandardized regression coefficients (*B*) and standardized regression coefficients (β). The changes in R^2 for each block suggest that for both models one and two, SES and primary home language combined accounted for only 1.0% of the variability, then by adding English language proficiency, model one accounted for 44% of the variability and model two, the receptive and productive elements of language, accounted for 48% of the variability in predicting science achievement on the CMAS for science.

Table 6

Hierarchical Multiple Regression Model Summary for Overall Achievement on the 2015 CMAS for Science

Model	Block	R	R^2	ΔR^2	ΔF
	SES	.09	.01	.01	47.49***
1	SES + HL	.09	.01	.00	3.05
	SES+HL+ ELP	.67	.44	.44	5020.19***
	SES	.09	.01	.01	47.49***
2	SES + HL	.09	.01	.00	3.05
	SES+HL+ R & P	.69	.48	.48	2843.01***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; R & P = Receptive and Productive Elements of Language; ***p<.001.

In the first block of model 1 (see Table 7), SES was a statistically significant predictor of academic achievement on the CMAS for science, $F(1, 6400) = 47.49$, $p < .001$. At step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 25.28$, $p < .001$. At step three, SES, primary home language, and English language proficiency were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 1703.47$, $p < .001$. These variables accounted for 44% of the variance in academic achievement on the CMAS for science. Block one in the second model (see Table 7), SES was a statistically significant predictor of academic achievement on the

CMAS for science, $F(1, 6400) = 47.49$, $p < .001$. At step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 25.28$, $p < .001$. At step three, SES, primary home language, and receptive and productive elements of language were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 1445.37$, $p < .001$. These variables accounted for 48% of the variance in academic achievement on the CMAS for science. Therefore, the receptive and productive elements of language increased the predictability of science achievement by an additional 4 % over English language proficiency overall. In addition, it is important to note that productive elements of language were more strongly predictive than receptive language elements. All predictor variables had statistically significant correlations with overall CMAS science achievement (see Appendix A).

Table 7

Hierarchical Multiple Regression Analysis Assessing Students' Home Language, English Language Proficiency and Receptive and Productive Elements of Language as Predictors of Overall Achievement on the 2015 CMAS for Science

Model	Variable	Block one		Block two		Block three	
		B	β	B	β	B	β
1	SES	-22.23***	-.09***	-21.25***	-.08***	-12.82***	-.05***
	HL			-6.1***	-.02***	-10.97***	-.04***
	ELP					2.76***	.66***
2	SES	-22.23***	-.09***	-21.25***	-.08***	-13.2***	-.05***
	HL			-6.1	-.02	-8.54***	-.03***
	RL					.71***	.13***
	PL					2.00***	.59***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; RL = Receptive Language; PL = Productive Language; *** p< .001.

Predictors of Physical, Life, and Earth Science Achievement

Three different hierarchical multiple regressions were calculated to predict academic achievement on the three different strands of science (physical, life, and earth) within the Colorado Measures of Academic Success (CMAS) for science based on primary home language, language proficiency, and receptive and productive elements of language. Analysis began with evaluation of assumptions and all assumptions were met except for multicollinearity for the same reasons mentioned with research question one. The procedure used within research question one to account for this violation was used for each of the regressions below.

Physical Science

Hierarchical multiple regression analysis was performed between overall achievement within the physical science strand on the CMAS for science as the criterion variables and, in model one, SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and overall English language proficiency in the third block as predictor variables. Table 8 displays effect size measures (R^2), change in R^2, and adjusted R^2 for the full model, and Table 9 displays pooled unstandardized regression coefficients (*B*) and standardized regression coefficients (β).

The changes in R^2 for each block suggest that, for both models one and two, SES and primary home language combined accounted for only 1.0% of the variability, and in model one, adding English language proficiency to the full model, accounted for 33% of the variability in predicting physical science achievement on the CMAS for science, and in model two, adding receptive and productive elements of language accounted for an 36% of the variability in predicting physical science achievement on the CMAS for science. Therefore, the receptive and productive elements of language increased the predictability of science achievement by an additional 3 % over English language proficiency overall.

Table 8

Hierarchical Multiple Regression Model Summary for Physical Science Achievement on the 2015 CMAS for Science

Model	Block	R	R^2	ΔR^2	ΔF
	SES	.09	.01	.01	47.16***
1	SES + HL	.09	.01	.00	6.55**
	SES+HL+ ELP	.57	.33	.32	3052.28***
	SES	.09	.01	.01	47.16***
2	SES + HL	.09	.01	.00	6.55**
	SES+HL+ R & P	.6	.36	.35	1741.92***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; R & P = Receptive and Productive Elements of Language; ***p<.001; **p<.01.

In the first block for model one (see Table 9), SES was a statistically significant predictor of physical science achievement on the CMAS for science, $F(1, 6400) = 47.16$, $p < .001$. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 26.88$, $p < .001$. Step three, SES, primary home language, and English language proficiency were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 1043.89$, $p < .001$. These variables accounted for 33% of the variance in academic achievement on the CMAS for science.

Block one in the second model (see Table 9), SES was a statistically significant predictor of academic achievement on the CMAS for science, $F(1, 6400) = 47.16$, $p <$

.001. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 26.88$, $p < .001$. Step three, SES, primary home language, and receptive and productive elements of language were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 891.71$, $p < .001$. These variables accounted for 36% of the variance in academic achievement on the CMAS for science. Therefore, the receptive and productive elements of language increased the predictability of the variability of achievement by an additional 3 % over English language proficiency in overall, and productive elements of language were more strongly predictive than receptive elements of language (see Appendix A).

Table 9

Hierarchical Multiple Regression Analysis Assessing Students' Home Language, English Language Proficiency and Receptive and Productive Elements of Language as Predictors of Physical Science Achievement on the 2015 CMAS for Science

Model	Variable	Block one		Block two		Block three	
		B	β	B	β	B	β
1	SES	-24.3***	-.09***	-22.73***	-.08***	-14.80***	-.05***
	HL			-9.8**	-.03**	-14.38***	-.05***
	ELP					2.59***	.58***
2	SES	-24.3***	-.09***	-22.73***	-.08***	-15.32***	-.05***
	HL			-9.8**	-.03**	-11.64***	-.04***
	RL					.43***	.07***
	PL					2.00***	.54***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; RL = Receptive Language; PL = Productive Language; *** p< .001.

Life Science

Hierarchical multiple regression analysis was performed between overall achievement within the life science strand on the CMAS for science as the criterion variables and, for model one, SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and overall English language proficiency in the third block as predictor variables. For model two, SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and receptive and productive elements of language in the third block as predictor variables. Table 10 displays effect size measures (R^2), change in R^2, and adjusted R^2 for the full models, and Table 11 displays pooled unstandardized regression coefficients (*B*), and standardized regression coefficients (β).

The changes in R^2 for each block suggest that in models one and two, SES and primary home language combined accounted for only 1.1% of the variability, and in model one, adding English language proficiency accounted for 35% of the variability, and in model two adding receptive and productive elements of language accounted for 37% of variability in predicting life science achievement on the CMAS for science.

Table 10

Hierarchical Multiple Regression Model Summary for Life Science Achievement on the 2015 CMAS for Science

Model	Block	R	R^2	ΔR^2	ΔF
1	SES	.07	.01	.01	32.78***
	SES + HL	.08	.01	.00	5.79*
	SES+HL+ ELP	.59	.35	.34	3330.43***
2	SES	.07	.01	.01	32.78***
	SES + HL	.08	.01	.00	5.79*
	SES+HL+ R & P	.61	.37	.36	1834.61***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; R & P = Receptive and Productive Elements of Language; ***p<.001; *p<.05.

In the first block for model one (see Table 11), SES was a statistically significant predictor of physical science achievement on the CMAS for science, $F(1, 6400) = 32.78$, $p < .001$. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 19.30$, $p < .001$. Step three, SES, primary home language, and English language proficiency were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 1129.70$, $p < .001$. These variables accounted for 35% of the variance in academic achievement on the CMAS for science.

Block one in the second model (see Table 11), SES was a statistically significant predictor of academic achievement on the CMAS for science, $F(1, 6400) = 32.78$, $p < .001$. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, $F(2, 6399) = 19.30$, $p < .001$. Step three, SES, primary home language, and receptive and productive elements of language were statistically significant predictors of academic achievement on the CMAS for science, $F(3, 6398) = 932.48$, $p < .001$. These variables accounted for 37% of the variance in academic achievement on the CMAS for science. Therefore, the receptive and productive elements of language increased the predictability of the variability of achievement by an additional 2 % over English language proficiency in overall, with the productive elements of language again being more strongly predictive than the receptive elements of language (see Appendix A).

Table 11

Hierarchical Multiple Regression Analysis Assessing Students' Home Language, English Language Proficiency and Receptive and Productive Elements of Language as Predictors of Life Science Achievement on the 2015 CMAS for Science

Model	Variable	Block one		Block two		Block three	
		B	β	B	β	B	β
1	SES	-20.67***	-.07***	-19.17***	-.07***	-10.84***	-.04***
	HL			-9.41*	-.03*	-14.21***	-.05***
	ELP					2.72***	.58***
2	SES	-20.67***	-.07***	-19.17***	-.07***	-11.07***	-.04***
	HL			-9.41*	-.03*	-12.2***	-.04***
	RL					.89***	.15***
	PL					1.88	.50

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; RL = Receptive Language; PL = Productive Language; *** p< .001; *p<.05.

Earth Science

Hierarchical multiple regression analysis was performed between overall achievement within the earth science strand on the CMAS for science as the criterion variables and, in model one, SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and overall English language proficiency in the third block as predictor variables. In model two, SES (dummy coded to FRL and Non-FRL) in the first block, SES and primary home language (dummy coded to Spanish and Other) in the second block, SES, primary home language, and receptive and productive elements of language in the third block were the predictor variables. Table 12 displays effect size measures (R^2), change in R^2, and adjusted R^2 for the full models, and Table 13 displays pooled unstandardized regression coefficients (*B*), and standardized regression coefficients (β).

The changes in R^2 for each block suggest that, in both models one and two, SES and primary home language combined accounted for only 1.0% of the variability, in model one, adding English language proficiency accounted for 33% of the variability, and model two adding receptive and productive elements of language accounted for 35% of the variability in predicting earth science achievement on the CMAS for science.

Table 12

Hierarchical Multiple Regression Model Summary for Earth Science Achievement on the 2015 CMAS for Science

Model	Block	R	R^2	ΔR^2	ΔF
1	SES	.07	.01	.01	33.02***
	SES + HL	.07	.01	.00	.22***
	SES+HL+ ELP	.57	.33	.32	3049.05***
2	SES	.07	.01	.01	33.02***
	SES + HL	.07	.01	.00	.22***
	SES+HL+ R & P	.59	.35	.35	1714.88***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; R & P = Receptive and Productive Elements of Language; ***p<.001.

In the first block for model one (see Table 13), SES was a statistically significant predictor of physical science achievement on the CMAS for science, [$F(1, 6400) = 33.02$, $p < .001$]. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, [$F(2, 6399) = 16.62$, $p < .001$]. Step three, SES, primary home language, and English language proficiency were statistically significant predictors of academic achievement on the CMAS for science, [$F(3, 6398) = 1032.71$, $p < .001$]. These variables accounted for 33% of the variance in academic achievement on the CMAS for science.

Block one in the second model (see Table 13), SES was a statistically significant predictor of academic achievement on the CMAS for science, [$F(1, 6400) = 33.02$, $p < .001$]. Step two, SES and primary home language were statistically significant predictors of academic achievement on the CMAS for science, [$F(2, 6399) = 16.62$, $p < .001$]. Step three, SES, primary home language, and receptive and productive elements of language were statistically significant predictors of academic achievement on the CMAS for science, [$F(3, 6398) = 870.20$, $p < .001$]. These variables accounted for 35% of the variance in academic achievement on the CMAS for science. Therefore, the receptive and productive elements of language increased the predictability of the variability of achievement by an additional 2 % over English language proficiency in overall, with the productive elements of language being more strongly predictive than the receptive elements of language. To view specific correlations between the dependent variables and independent variables mentioned in the previous sections, (see Appendix A).

Table 13

Hierarchical Multiple Regression Analysis Assessing Students' Home Language, English Language Proficiency and Receptive and Productive Elements of Language as Predictors of Earth Science Achievement on the 2015 CMAS for Science

Model	Variable	Block one		Block two		Block three	
		B	β	B	β	B	β
1	SES	-20.94***	-.07***	-20.64***	-.07***	-12.48***	-.04***
	HL			-1.85	-.01	-6.56*	-.02*
	ELP					2.67***	.57***
2	SES	-20.94***	-.07***	-20.64***	-.07***	-12.94***	-.04***
	HL			-1.85	-.01	-3.94	-.01
	RL					.55***	.09***
	PL					2.00***	.53***

Note. SES = Socioeconomic Status; HL = Home Language; ELP = English Language Proficiency; RL = Receptive Language; PL = Productive Language; *** p< .001; *p<.05.

Chapter 5: Discussion and Implications

There is a need to better understand how language impacts achievement, specifically concerning EBs, and since language is the means through which content knowledge is assessed, it was important to analyze this aspect of learning. A review of literature identified the need to inform the creation of more reliable and valid content assessments for EBs (Abedi, 2008b), to study the impact of English proficiency on EBs performance on standardized assessments (Solorzano, 2008; Wolf et al., 2008), and to further research the role language plays in content assessments and the impact they have on EBs performance (Abedi, 2008b; Abedi et al., 2004; Abedi & Lord, 2001).

This study investigated the effects of primary home language, English language proficiency, and receptive and productive elements of language on student academic achievement in science, overall and by strand, as measured by the Colorado Measures of Academic Success (CMAS) for science. There were four key findings with regard to the effects of language on science achievement overall and by strand. First, primary home language of Spanish accounts for a large part of variability in science achievement and therefore is a good predictor of achievement. This result was surprising, not only because it is not supported in the limited research around home language (Hernandez, 2007), but also because Spanish and English share many cognates due to Latin roots (Escamilla, Hopewell, Butvilofsky, Sparrow, Soltero-Gonzalez, Ruiz-Figueroa, & Escamilla, 2014;

Genesee, Lindholm-Leary, Saunders, & Christian, 2005), and science is largely Latin rooted. Therefore, one could consider that Spanish-speaking EBs would be able to recognize cognates in science much better than EBs whose primary home language is one that is non-alphabetic. Second, English language proficiency is predictive of science achievement. This is consistent with prior studies looking at whether English language proficiency was a predictor of achievement for elementary students (Abedi & Gándara, 2006; Lee & Fradd, 1998; Lee, 2006; Solano-Flores & Trumbull, 2003). Third, receptive and productive elements of language are predictive of science achievement and increase the predictability of science achievement over ELP alone. Fourth, with regards to strands of science, predictability of receptive and productive elements of language on science achievement is increased by one percent more in physical science than in life science or earth science. Overall, the findings indicate that receptive and productive elements of language are good predictors of a student's science achievement on a high stakes standardized assessment in science; with the productive elements being the strongest predictor. This was another surprising result because research supports context-embedded language learning, so one might think that the receptive elements may have been the stronger predictor.

The results presented some important factors that may influence science achievement for eighth grade EB students. In this study, as has been found in previous studies examining socioeconomic status and achievement (Caldas & Bankston, 2012; Selcuk, 2005), socioeconomic status alone was a statistically significant predictor of achievement. However, when combined with English language proficiency and receptive and productive elements of language, the latter seemed to predict achievement more than socioeconomic status. Also, since a student's primary home language was predictive of science achievement, the fact that Spanish-speaking EBs score lower than their peers on high-stakes science assessments suggest that they lacked the opportunity to learn science content. These findings will be further delineated in the subsequent sections.

Primary Home Language

The national rhetoric around EBs is that if they are Spanish-speaking they will have a harder time acquiring English; Spanish is perceived as a deficit or barrier to English language proficiency (Flores, Cousin, & Diaz, 1991; Flores, 2006). This could be driven by the fact that the majority of all EBs nationally are Spanish-speaking (NCELA, 2015) and are multigenerational, therefore this could be merely an issue of n-size. However, this study found that a students' home language is predictive of science achievement. If an EB is Spanish-speaking, their overall scores on CMAS for science decrease by 10.97 points holding all other variables equal in model one. In model two, overall scores decrease by 8.54 points holding all other variables equal. Similar results were found for each of the strands of science as well. These results have deeper implications than just achievement. If all EBs, regardless of primary home language, score lower than their peers on achievement tests, level of English language proficiency could lead to an explanation. However, when a specific subpopulation of those EBs, Spanish-speaking students, score even lower than their EB peers, something else may be in play and need further investigation.

English Language Proficiency

In both overall science and science by strand, English language proficiency (ELP) was a statistically significant predictor of achievement. As an EB student's ELP score increases, the science achievement score increases by 2.76 points in model one. Similar results held when examining each of the strands of science as well. English language proficiency predicted achievement more strongly than SES. As noted in the review of literature, research shows that although other factors such as SES and parent education level impact EBs achievement, language has the greatest impact (Abedi, 2002, 2008b, 2009; Abedi et al., 2004). Second language acquisition research states that BICS is generally acquired within one to three years; however research continues to demonstrate that CALP requires between four to seven years (Cook et al., 2011; Cummins, 2008). The results support this theory in that a students' level of ELP was predictive of academic achievement.

Receptive and Productive Elements of Language

Receptive and productive elements of language were statistically significant predictors of both overall science achievement and science achievement by strand above and beyond ELP. As an EB student's receptive scores increase, their science achievement scores increase by 0.71 points, and as their productive scores increase the science achievement score increases by 2.00 points. Similar results were found for the physical and earth strands of science as well. Results were different for Life science. Receptive language was statistically significant; however the productive language was not. When thinking about this particular result, consideration is around the use of cognates and how there may be more of them within the life sciences then the other strands of science. This could be something for further investigation.

Throughout the study, the productive elements of language remained the most predictive of achievement. This is an important result because practitioners generally use only the overall ELP score when making decisions either about programming or placement for students and it may not provide the most complete picture of their needs. Educators make the point that four domains of language are present and evaluated, regardless of how you combine them (WIDA, 2015). However, the results from this study demonstrate that receptive and productive elements of language do make a difference and evaluating how students receive and produce information is even more predictive than overall ELP alone and in particular, productive language had a stronger influence. It is important to consider why this result has occurred within this specific dataset. It could be due to the new assessment being administered in an on-line platform and this alone could have influenced how students produced their understanding. Extending this idea even further, it couls have been the specific item types that were part of the assessment. Maybe the technology-enhanced items and simulations made the assessment more accessible in general, that the students were able to express their understanding in a much clearer way. Since 2015 was only the second year of administration using an on-line platform and innovative item types, it may be something for further investigation. These results are significant, not only because they are filling a gap in the body of research on assessing EBs, but also because they are demonstrating that it is important to consider all four domains of language and not limit the focus to literacy. Educators should comprehensively address language domains within assessment and instruction.

While making instructional and assessment decisions at the classrooms level, educators need to be cognizant of how they are using language and how they are asking students to use language. ELP levels are important indicators of what students can do with language, but not the only indicator. Providing multiple opportunities for students to express their understanding of scientific concepts in verbal and/or written modes and receive information through various modalities is what the results of this study support.

As noted previously in the review of literature, there are many schools of second language acquisition theory, *Behaviorism, Innatism, and Interactionism,* and all three can be informative due to the sociocultural context with which they are embedded. Whether it is believed that students acquire or learn language through repetition, modeling, an innate ability, or interactions with their environment, focusing on how students integrate receptive and productive language is proving to be an important factor as well.

In addition, as Bailey (2007) pointed out with her adaptation of Cummins' theoretical model of BICS and CALP, there appears to be a blending of traditionally "social" language with "academic" language by thinking more broadly about the four domains of language. Beyond a shadow of a doubt, literacy is important. However, limiting thinking to literacy for second language learners may be short-sighted. It is the blending of the social (listening and speaking) with the academic (reading and writing) that is also important and possibly more important.

Limitations

One of the limitations of this study was that EB students were all grouped together and it is known that this population of students is very diverse; the complexity in the background characteristics of EBs was unavailable to the researcher or incalculable. In addition, one of the research questions in this study examined only one language group and it is known that Spanish-speaking EBs comprise a wide variety of demographic, experiential, and cultural differences. Researchers have mentioned the difficulty of having a well-defined sample in order to make generalizations and to find patterns in understanding. This study was limited by the fact that little is known about the cultural differences of the Spanish-speaking EBs within the study sample. Therefore the results of this study, with regards to primary home language, can only generalize to eighth grade, Spanish-speaking EBs in the state of Colorado. Another limitation regarding the sample is that the population of students examined was only eighth grade EBs. Therefore, the results are only generalizable to eighth graders in Colorado.

The final limitation of this study is that the results rely heavily on the receptive and productive elements of language and due to their nature as composite scores, they are compensatory. "Compensatory means that a high score in one language domain could inflate the composite score, compensating for a low score in another language domain; conversely, a low score in a language domain could bring down the composite" (WIDA, 2015, p. 9). This does not mean that the receptive and productive scores are uninterpretable; rather that caution needs to be used during interpretation of composite scores.

Implications

The implications of this research are four fold. First, it brings to the forefront the issue of construct validity in the high stakes assessment of EBs who do not have a strong command of the English language. If an assessment does not properly assess the intended construct, in this case the CMAS for science, validity of the construct comes into question. In that sense, it confirms a need to reexamine construct validity issues. Second, it adds to the educational literature on assessment of EBs specifically around Spanish-speaking EBs and using both the receptive and productive domains of language when reporting results. Third, it highlights the need to discuss how students receive and produce language, not only for predicting achievement outcomes, but for assessment design and using assessment data within classrooms. Lastly, it provides topics for professional development for science teachers, such as, focusing instruction on the four domains of language, using formative assessment as an instructional practice, and not waiting until students are able to express themselves fluently and with correct grammar or rather appear to be proficient with oracy.

Construct validity. Construct validity is basically the degree to which a test measures what it claims to be measuring and is an emergent issue in assessment of EBs. The entanglement of language with content is difficult to tease out and to assess specifically because they are inextricably linked. A student needs to be able to understand the language of science as well as communicate their understanding using scientific discourse. That, in and of itself, is a part of being scientifically literate. However, testing a student who does not have command of the target language, in this case English, allows for construct validity arguments to be made. It is important to pose the following questions: Is the assessment assessing science content or is it assessing if the student has the English proficiency to access the assessment? In regards to the CMAS for science, it is clear that for EBs, it is measuring ELP which is the wrong construct; thus, potentially invalidating the scores of these students. What kinds of inferences can be drawn from assessment results for NEP students? It is clear from the results that the only inference that can be made based on the CMAS for science scores for these students is that they do not have sufficient command of English to access the assessment.

In an era of standards-based education and accountability, assessment is not going away, however one needs to consider the usefulness of such assessments for EB students. Policymakers at the state and federal levels need to take into consideration a students' level of ELP when requiring assessment of EBs who are not proficient in the language of the assessment. "Some proportion of the academic achievement gap may be due not to an EB's lack of content knowledge, but to the content assessment's inability to accurately measure that knowledge when insufficient language proficiency stands in the way" (Bailey, 2007, p 278). This is a major equity issue in education today. Policy makers declare that all students need to be assessed so that educators can make inferences about their understanding of specific science concepts and skills. However, those inferences have limitations based on the interpretations of assessment scores that may be flawed. Herein lies the inequity.

This is not to say that these students do not need to be monitored for their ELP or their content knowledge, merely considerations of validity need to be taken into account. If states are to be held accountable for EBs as a whole group, districts are to be held accountable for collective growth, and individual teachers are held accountable for individual student growth, then the measurement tool needs to be valid and reliable in providing the necessary information so that others can make appropriate inferences about these students. As Bailey (2007) asserts, a construct needs to clearly define the necessary English language skills that are predictably needed for academic achievement.

That being said, the next generation standards (i.e., CCSS and NGSS) specify the teaching and assessment of the *communication* of content knowledge in addition to content knowledge itself. Therefore, language may become construct *relevant* in the age of these new standards and the onus will be on test developers to clearly define the content construct in order to avoid unnecessary linguistic complexity (Bailey &Carroll, 2015).

Assessment of EBs. Avoiding violations of construct validity is not the only challenge with assessing EBs. Reliance on the interpretation of results from large-scale assessments becomes problematic when making decisions around program placement, targeted support, continuation of support, exiting students out of program, and allocation of resources. Most of the decisions made rely on inferences around the acquisition of academic language. Title III of the NCLB act (2001), holds states and districts accountable for EBs growth through Annual Measureable Achievement Objectives. These indicators of growth need to include ELP and academic achievement targets. The scores on the content assessment impact this level of accountability, therefore inferences made of students' progress need to be as accurate as possible. It is important to remember that the most proficient students in the EB population, in Colorado's case these students are identified as FEP or FELL, are not a part of the EB subgroup when analyzing results from assessments. Therefore, by definition, those students in the EB subgroup are not yet proficient in English. The implication here is that the language demands placed on EBs may eclipse their display of academic content (Dutro, 2006) and are therefore not an accurate source of information for accountability purposes, unless combined within a body of evidence. As such, large scale assessments have a purpose in supporting the call for comparison and accountability, but may not inform instruction at the classroom level.

Assessments are blunt instruments that have limitations, based on how they were constructed around a single construct and hence should not be the only way to assess student learning. Eisner (2002) stated "Not everything that matters can be measured, and not everything that is measured matters" (p.178). In the current environment of Every Student Succeeds, states and districts are able to reimagine this system and have the opportunity to be more innovative. Our current types of assessment do not align with the next generation of standards which are more open to process, experimentation, innovation, and skills. One could argue that this new way of thinking about standards and assessment lends itself to be less language loaded, more context-embedded, and more relevant to students today. This would benefit EBs, as well as other students with low socioeconomic status, because it has the potential to build upon their strengths, instead of their deficits. Educators need to focus on what these students CAN DO, instead of always focusing on what they cannot do.

Receptive and Productive Elements of Language. When the focus remains on what students can do, the vision for what is possible in classrooms expands. It is important that educators begin to think more about how they are asking students to receive information and produce their understanding. With so much focus on literacy, educators are losing out on the two other domains of language, listening and speaking, which are fundamental to a students' development. The national literacy panel report (August, 2009) supports this idea that developing oral language is just as important as developing literacy skills and that educators need to take all four domains of language into consideration when developing instructional experiences for students. The results of this study demonstrate that receptive and productive elements are predictive of achievement, so it would be beneficial to be more intentional in our practice around these elements of language. For example, one could focus on oracy, but not wait for full development to start reading and writing. Or, making sure that lesson plans include opportunities for students to use all four domains of language within instruction and that the teacher is creating a language rich environment. This would be an area to focus professional development for educators. Intentional planning and explicit instruction linked to how students receive, experience, and produce their understanding would be a place to begin.

The emphasis on communication and collaboration in the next generation standards will have new ramifications on the interpretation of scores for EBs. One of these ramifications could be that EB students may find these new assessments even more difficult than the previous assessments. Therefore, it will be important for educators to redesign their instruction and classroom assessments around how students are receiving information and producing their understanding. Secondly, assessment of content knowledge will become multidimensional, assessing content knowledge and a communication ability component that determines how well students can convey their knowledge.

Professional Development for Science Teachers. High quality teaching is important for student success and it is vital that educators ensure students have equitable access to opportunities, support, and tools they need to succeed. Traditionally, EBs are tracked into low-level classes, which supplant academic content with English support services and limit EBs' access to core curricular content. Most schools do this as a way to support their EBs, however, this type of tracking is not equitable and does not allow students to reach their full potential. Hence, the need for professional development.

This study highlighted four areas of focus for professional development: home language, academic language, ELP in relation to achievement and assessment, and use of the four domains of language to support students. Since the majority of these focal areas operate within the realm of equity, it is important that educators identify what mental models they have established around the EB population of students, so that they can fully embrace professional development within any of the four areas mentioned above. Mental models are our values, beliefs, and assumptions about how the world works and from those, stories are created about other people or institutions that influence behavior (Aguilar, 2016).

Once educators have a grasp of their mental models, they can use them to think about how they view students with primary home languages different from their own, and if these views impact their ideas of teaching and learning. As mentioned in the literature review, Ogbu (1983) highlighted ideas around voluntary versus non-voluntary immigrants and the influence on achievement and Salazar (2008) used the idea of the *maleta* to convey the subtractive educational practices that are currently in our schools. If Spanish-speaking EBs score lower than other EBs, then educators need to not only investigate their mental models, but also think about providing equitable opportunities to learn for all students, especially their EBs. Educators should also draw upon a student's "funds of knowledge" (González, Moll, & Amanti, 2005) so that students can use the linguistic and cultural resources within their *maletas* to acquire new knowledge and skills. Schools across the nation have high numbers of second and third generation Spanish-speaking EBs (August, Shannahan, & Escamilla, 2009) who are underachieving. Is this due to chance or have these students become so disenfranchised, due to their lack of opportunity to learn, that reaching higher levels of achievement is difficult?

Opportunities to learn within the classroom environment should also include explicit attention to academic language. Content area teachers usually do not view themselves as language teachers, however inherent in conceptual and skill development within the content area is the development of disciplinary literacy. The results of this study demonstrate that all four domains of language contribute to academic language, not just the traditional view of CALP which only includes reading and writing. Therefore, professional development around building academic language needs to include all four domains in receptive and productive combinations. Educators could leverage the Colorado English Language Proficiency standards, in addition to support from WIDA to develop instruction around what students can do. The resources include "Can Do" indicators to assist educators in choosing ways to support students at varying levels of English proficiency based on what the student is able to do.

As educators become more comfortable with the construct of academic language within the context of the next generation standards for communication, they need to begin embedding multidimensional formative and summative assessments into their instruction. Again, this would include ideas around receptive and productive language. If classrooms need to become more collaborative and inclusive, for some with the focus on problem-based learning, educators need to be cognizant of including explicit instruction on the structures of language at the word, sentence, and paragraph levels so that students can effectively communicate their understanding of scientific concepts.

Future Directions

This research could lead to four other areas of study. First, an element for further study is to look at how students performed on innovative item types within the state assessment; namely technology enhanced items and simulations. Technology enhanced items are computer-delivered item type and include specialized interactions for collecting response data. These include interactions and responses beyond traditional selected-response or constructed-response (CDE, 2015e). Simulation refer to an assessment item type where the student interacts with the item by observing and performing tasks associated with that item (CDE, 2015e). Using the lens of how EB students receive and produce language, it is important to study how these students' achievement results differ based on how they are asked to engage with items and produce their responses to those items. Would assessment simulations provide more contexts and decrease the language load for all students taking the assessment? Would these item types "level the playing field" in a way that allows EBs better access to the item, therefore producing a more valid result of their science knowledge? These are just a few questions that would be important to ask in relation to innovative item types.

Second, comparing Problem-Based Learning science classrooms to traditional science classrooms to study disciplinary literacy through the lens of receptive and productive elements of language. One of the challenges of comparing innovative classrooms with traditional classrooms is that there may not be enough EBs within the innovative classrooms to make a true comparison. Colorado school districts choose the type of programming EBs will receive to effectively develop English language proficiency. In reality, this means that EB students are often placed in programs that focus on basic language skills rather than rigorous content, meaning that they are not getting access to grade level science content (Lee & Fradd, 1998). As a result, many EB students matriculate through grade levels without a strong foundation in science, and they continually score below their English-speaking peers on standardized achievement assessments. These practices are highly inequitable, especially since research has shown that experiential learning (Mollaie & Rahnama, 2012) and collaborative grouping

(Francis et al., 2006) are strategies that improve outcomes for EBs. It is important to use this knowledge to inform how we engage all students in science instruction and improve their opportunities to learn.

Third, the study of academic language, specifically looking at linguistic complexity of assessment items as a validity argument, would be important to add to the body of research around assessment of EBs. Science has very specific kinds of discourse or disciplinary literacy which are part and parcel to science achievement. One cannot remove the scientific language of items because the language is also what is being assessed. In addition, the implementation of the next generation standards brings to the forefront the idea of communicative competence and the possibility of a multidimensional construct that would include elements of language and content within assessment items. The best way to get at this multidimensionality of assessment is through performance-based assessment. This could be another area for further exploration.

The fourth area for further research is around primary home language linked to long term English learner status. Based on results from this study, Spanish-speaking EBs score lower than other EBs, so this needs to be explored further. Using Ogbu's (1983) ideology around immigrant versus non-immigrant minorities to frame the exploration of generational EBs, institutional marginalization and disenfranchisement could be explored as possible explanations. Also, disaggregating the "Spanish-speaking" population into subgroups (e.g., migrant, mobility rates, SES, assessments in Spanish, not allow assessment in Spanish) and comparing results to English speakers.

In addition, it may be important to explore the issue of equitable access to science courses in general for EBs and then specifically for Spanish-speaking EBs. Secondary EBs tend to be placed in English-language support classes to build up their oracy, however the rigor of the courses tends to not be the same as others (August et al., 2009). This could be due to the misconception that EBs cannot access higher level content if they do not have basic communication skills.

In my experience as a science educator and, now, currently working with science educators and administrators, I understand the inclination to focus on oral language development prior to engaging in literacy or content instruction. However, we are doing a disservice to our EB students, and society as a whole, when we unintentionally undermine their inherent desire to learn and achieve by not providing them with equitable opportunities to learn. As educators, we always act with best intentions and we want the best for our kids. In Colorado, with this ever-widening achievement and opportunity gap (CDE, 2014b), we need to take a hard look at our instructional and assessment practices, and the hidden curriculum (Uhrmacher, 1997) that schools operate under surrounding students whose first language is not English and ask ourselves, what are we afraid of? Being biliterate/multiliterate is an asset in our global society, one that other countries value. So why is it that we continue to operate under subtractive, not additive educational policies? Teachers need to employ instructional methods that are additive rather than subtractive and value students' "funds of knowledge" (González et al., 2005 Salazar, 2008).

Leadership is important and is observable at many levels. However, teacher leadership is vital. Teachers are the ones who can choose to be additive, and to integrate the four domains of language throughout their planning, instruction, and assessment. They are the ones who can provide equitable opportunities to learn within their classrooms, and view EBs through the lens of their assets. Knowing that the implementation of systemic change is a complex issue, it will also take policy leadership from state agencies and the cooperation of educator preparation institutions to understand the needs of the next generation science teacher. Science is a noun and a verb. Unfortunately, our current educational system focuses more on the noun than the verb. When educators and students understand science as a way of thinking and problem solving, then maybe the integration of the receptive and productive elements of language will not appear as difficult to implement within classrooms as some might think.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education,* 14, 219-234.

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment, 8*(3), 231-257.

Abedi, J., & Dietel, R. (2004). Challenges in the No Child Left Behind Act for English language learners. Retrieved from http://www.cse.ucla.edu/products/policy/cresst_policy7.pdf

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

Abedi, J. (2006). Language issues in item development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 377-398). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Abedi, J., & Gandara, P. (2006). Performance of English language learners as a sub-group in large scale assessment: Interaction of research and Policy. *Educational Measurement: Issues and Practice, 25*(4), 36-46.

Abedi, J. (2008a). *Linguistic modification: Part I –Language factors in the assessment of English language learners: the theory and principles underlying the linguistic modification approach.* Washington, DC: LEP Partnership.

Abedi, J. (2008b). Measuring students' level of English proficiency: Educational significance and assessment requirements. *Educational Assessment*, *13*(2/3), 193-214.

Abedi, J. (2009). Computer testing as a form of accommodation for English language learners. *Educational Assessment*, 14, 195-211.

Abella, R., Urrutia, J. & Shneyderman, A. (2005). An examination of the validity of English-language achievement test scores in an English language learner population. *Bilingual Research Journal, 29*(1), 127-144.

Aguilar, E. (2016). *ShiftingMental Models in Educators.* Edutopia: George Lucas Educational Foundation. Retrieved from http://www.edutopia.org/blog/shifting-mental-models-educators-elena-aguilar

Auckerman, M. (2007). A culpable CALP: Rethinking the conversational/academic language proficiency distinction in early literacy instruction. *Reading Teacher, 60*(7), 626-635.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners*. New York: Routledge.

August, D., Shanahan, T., & Escamilla, K. (2009). English Language Learners: Developing Literacy in Second-Language-Learners Report of the National Literacy Panel on Language-Minority Children and Youth. *Journal of Literacy Research, 41*(4), 432-452.

Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing, 19*, 453-476.

Bachman, L. & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

Bailey, A. (2007). *The Language Demands of School: Putting academic English to the test.* New Haven, CT: Yale University Press.

Bailey, A. (2008). What is Academic Language? In M. Gottlieb and G. Ernst-Slavit, *Academic Language in Diverse Classrooms: Definitions and Contexts* (p. 6-10). Thousand Oaks, CA: Corwin.

Bailey, A. & Carroll, P. (2015). Assessment of English Language Learners in the Era of New Academic Content Standards. *Review of Research in Education, 39*, 253-294.

Bailey, A. & Huang, B. (2011). Do current English language development/proficiency standards reflect the English needed for success in school? *Language Testing, 28* (3), 343-365.

Bailey, A. & Wolf, M. (2012). *The challenge of assessing language proficiency aligned to the Common Core State Standards and some possible solutions.* Paper presented at the Understanding Language Conference, Stanford University, CA.

Baker, C. (2006). *Foundations of Bilingual Education and Bilingualism.* Towanda, NY: Multilingual Matters Ltd.

Batalova, J., Fix, M. & Murray, J. (2007). *Measures of Change: The Demography and Literacy of Adolescent English Learners.* National Center on Immigrant Integration Policy, Migration Policy Institute.

Blankstein, A. & Noguera, P. (2015). *Excellence Through Equity: Five Principles of Courageous Leadership to Guide Achievement for Every Student.* Alexandria, VA: ASCD.

Bottia, M., Stearns, E., Mickelson, R., Moller, S., & Parker, A. (2015). The Relationships among High School STEM Learning Experiences and Students' Intent to Declare and Declaration of a STEM Major in College. *Teachers College Record, 117* (4), 1-46.

Boyson, B. A., & Short, D. J. (2012). *Helping newcomer students succeed in secondary schools and beyond.* Washington, DC: Center for Applied Linguistics.

Brasel, Y. (2008). *The Relationship Between Parental Involvement and Reading Comprehension Performance of English Language Learners.* Dissertation, College of Education, Walden University.

Brown, H. D. (2000). *Principles of language teaching* (4th ed.). White Plains, NY: Addison Wesley Longman, Inc.

Caldas, S. & Bankston, C. (2012). Effect of School Population Socioeconomic Status on Individual Academic Achievement. *Journal of Educational Research, 90* (5), 269-277.

Callahan, R. & Gándara, P. (2004). Nobody's Agenda: English Learners and Post-secondary Education, in M. Sadowski, (Ed.) *Immigrant and English-Language Learners: Strategies for Success*. Cambridge: Harvard Education Press.

California Association of Science Teachers (2008). *Universal design for learning guidelines version 1.0*. Wakefield, MA: Author.

Cazden, C. (2001). *Classroom Discourse: The language of teaching and learning* (2nd Ed.). Portsmouth, NH: Heinemann.

Chamot, A. U., & O'Malley, J. M. (1994). *The Calla handbook: Implementing the cognitive academic language learning approach.* White Plains, NY: Addison Wesley Longman.

Chomsky, N. (1986). *Knowledge of language: Its origin, nature and use.* Westport: Greenwood

Collier, V.P. (1995). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Research Journal, 16*, 187-212.

Collier, V.P. (1987a). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21*, 614-671.

Collier, V. (1987b). *The effect of age on acquisition of a second language for school.* Wheaton, MD: National Clearinghouse for Bilingual Education.

Colorado Department of Education. (2014a). *School and District Performance Framework Overview.* Retrieved from http://www.cde.state.co.us/accountability/spfdpf_technicalwriteup_072814

Colorado Department of Education. (2014b). *Culturally and Linguistically Diverse Learners in Colorado: State of the State2014.* Retrieved from http://www.cde.state.co.us/cde_english/stateofstate2014

Colorado Department of Education. (2015a). *CMAS Results for District and School.* Retrieved from http://www.cde.state.co.us/assessment/cmas-sciencesocial-dataandresults

Colorado Department of Education. (2015b). *Guidebook on Designing, Delivering and Evaluating Services for English Learners (EL).* Retrieved from http://www.cde.state.co.us/cde_english/elguidebook

Colorado Department of Education. (2015c). *Standards and Instructional Support.* Retrieved from http://www.cde.state.co.us/standardsandinstruction

Colorado Department of Education. (2015d). *CMAS Accommodations manual for science.* Retrieved from http://www.cde.state.co.us/assessment

Colorado Department of Education. (2015e). *CMAS Science Assessment Technical Manual.* Retrieved from http://www.cde.state.co.us/assessment/newassess-sum

Colorado Education Initiative. (2014). *The Colorado STEM Education Roadmap.* Retrieved from http://www.coloradoedinitiative.org/wp-content/uploads/2014/09/CO-STEM-Roadmap-w_Appendices.pdf

Conchas, G. & Perez, C. (2003). *Surfing the "model minority" wave of success: How the school context shapes distinct experiences among Vietnamese youth*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14750268

Cook, H. G., Boals, T., & Lundberg, T. (2011). Academic achievement for English language learners: What can we reasonably expect? *Phi Delta Kappan, 93*(3), 66-69.

Cosentino de Cohen, C., Deteerding, N., & Clewell, B. C. (2005). *Who's left behind? Immigrant children in high and low LEP schools*. Washington, DC: Urban Institute.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. Retrieved from http://www.wisd.us/campus/whs/social_studies/edd/Fall09/8344/Articles/CumminsBICSCALPSpringer2007.pdf

Cummins, J. (2003). BICS and CALP: Origins and rationale for the distinction. In C. Bratt Paulson & G. R. Tucker (Eds.), *Sociolinguistics the essential readings* (p.322-328). Malden, MA: Blackwell Publishing.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, *19*, 121-129.

Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge, England: Cambridge University Press.

Dutro, S. (2006). *Providing language instruction. Aiming high/Aspirando a lo Mejor resource.* Santa Rosa, CA: Sonoma County Office of Education.

Echevarria, J., Short, D., & Powers, K. (2006). School reform and standards-based education: A model for English-language learners. *Journal of Educational Research*, *99*(4), 195–210.

Eisner, E. (2002). *The educational imagination: On the design and evaluation of school programs* (3rd ed.) Upper Saddle River, NJ: Merrill Prentice Hall.

Escamilla, K. & Hopewell, S. (2010). Transitions to biliteracy: Creating positive academic trajectories for emerging bilinguals in the United States. In Petrovic, J., *International Perspectives on Bilingual Education: Policy, practice, and controversy, (*p.69-93). Charlotte, NC: Information Age Publishing.

Escamilla, K., Hopewell, S., Butvilofsky, S., Sparrow, W., Soltero-Gonzalez, L., Ruiz-Figueroa, O., & Escamilla, M. (2014). *Biliteracy from the Start: Literacy Squared in Action.* Philadelphia, PA: Caslon.

Fisher, D., Rothenberg, C. & Frey, N. (2007). *Language Learners in the English Classroom.* Urbana, IL: National Council of Teachers of English.

Francis, D. J., Rivera, M., Lesaux, N., Keiffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for the instruction and academic interventions.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Freeman, Y. & Freeman, D. (2009). *Academic language for English language learners and struggling readers: How to help students succeed across content areas.* Portsmouth NH: Heinemann.

Flores, A. (2007). Examining Disparities in Mathematics Education: Achievement Gap or Opportunity Gap? *The High School Journal, 19* (1), 29-42.

Flores, B. (2006). The Intellectual Presence of the Deficit View of Spanish-Speaking Children in the Educational Literature during the 20th Century. In *Latino Education: An Agenda for Community Action Research.* Mahwah, New Jersey: Lawrence Earlbaum Associates, Publishing.

Flores, B., Cousin, P., & Diaz, E. (1991). Transforming deficit myths about learning, language, and culture. *Language Arts, 68*(5), 369-379.

Gallagher, A., Bridgeman, B., & Cahalan, C. (2000). *The effect of computer-based tests on racial/ethnic, gender, and language groups.* Princeton, NJ: Educational Testing Service (ETS).

García, O., Kleifgen, J., & Falchi, L. (2008). From English language learners to emergent bilinguals. In *Equity Matters: Research Review* No. 1. New York: A Research Initiative of the Campaign for Educational Equity.

Gee, J. (1990). *Social linguistics and literacies: Ideology in discourses.* London, England: Falmer Press.

Gee, J. (1999). *An introduction to discourse analysis theory and method.* London, England: Routledge.

Genesee, F., Lindholm-Leary, K., Saunders, W. & Christian, D. (2005). English Language Learners in U.S. Schools: An Overview of Research Findings. *Journal of Education for Students Placed at Risk, 10* (4),

Goldenberg, C. (2008). Teaching English Language Learners: What Research Does and Does Not Say. *American Educator,* Summer.

González, N., Moll, L. & Amanti, C. (2005). *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Gottlieb, M. & Ernst-Slavit, G. (2009). *Academic Language in Diverse Classrooms: Definitions and Contexts.* Thousand Oaks, CA: Corwin.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). *Multivariate Data Analysis* (3rd ed.). New York: Macmillan.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara: University of California Linguistic Minority Research Institute.

Halliday, M. & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective.* Oxford: Oxford University Press.

Hernandez, C. (2007). Home Language Use and Hispanic Academic Achievement: Evidence from Texas High Schools. *Penn McNair Research Journal,1* (1), 1-30.

Howard, E., Páez, M., August, D., Barr, C., Kenyon, D. & Malabonga, V. (2014). The Importance of SES, Home and School Language and Literacy Practices, and Oral Vocabulary in Bilingual Children's English Reading Development. *Bilingual Research Journal, 37*(2), 120-141.

Krashen, S. (1987). *Principles and Practice in Second Language Acquisition.* Prentice-Hall International.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation.* New York, NY: Cambridge University Press.

Lawrence, J. F., White, C., & Snow, C. E. (2010). The words students need. *Educational Leadership, 68*(2), 23-26.

Lee, O. (2005). Science Education with English Language Learners: Synthesis and Research Agenda. *Review of Educational Research, 75*(4), 491-530.

Lee, O. & Fradd, S. (1998). Science for All, Including Students from Non-English-Language Backgrounds. *Educational Researcher, May (*27), 12-21.

Lemke, J. (1990). *Talking science: language, learning, and values.* Brooklyn, NY: Ablex Publishing Corporation.

Lightbown, P. M., & Spada, N. (2004). *How languages are learned* (2nd ed.). Oxford, England: Oxford University Press.

Lindholm-Leary, K. & Hernández, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: native English speakers, fluent

English/previous ELL, and current ELLs. *Journal of Multilingual and Multicultural Development*, *32*(6), 531-545.

Long, M. H. (2007). *Problems in SLA.* New York, NY: Lawrence Erlbaum Associates.

Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record, 109*(4), 897-926.

Martinello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review, 78*(2), 333-368.

Matrundola, D. L. T., Chang, S., & Herman, J. (2012). *Evaluation of calipers II: Using simulations to assess complex learning site visit findings* (CRESST Report 821). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mollaei, F & Rahnama, H. (2012). Experiential Education Contributing to Language Learning. *International Journal of Humanities and Social Science, 2(21), 268-279.*

National Academies Press (2012). *A Framework for K-12 Science Education: Practices, Cross Cutting Concepts, and Core Ideas.* National Research Council: Washington, DC.

Nasir, N. & Hand, V. (2006). Exploring Sociocultural Perspectives on Race, Culture, and Learning. *Review of Educational Research, 76* (4), 449-475.

National Center for English Language Acquisition (NCELA) (2015). County-by-County Maps of the EL Student Population. Retrieved: http://www.ncela.us/content/28_maps08_09

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Noble, T., Suarez, C., Rosebery, A., O'Connor, M., Warren, B., & Hudicourt-Barnes, J. (2012). "I never thought of it as freezing": How students answer questions on large-scale science tests and what they know about science. *Journal of Research in Science Teaching, 49*, 779-803.

Norris, S. P., & Phillips, L. M. (2009). Scientific literacy. In D. Olson & N. Torrance (Eds.). *The Cambridge handbook of literacy* (pp. 271-285). New York, NY: Cambridge University Press.

Norton, B. and Toohey, K. (2001). Changing Perspectives on Good Language Learners. *TESOL Quarterly, 35 (2)*, 307-322.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Ogbu, J. (1983). Minority Status and Schooling in Plural Societies. *Comparative Education Review 27*(2), 168-190.

Pellegrino, J. (1999). *The Evolution of Educational Assessment: Considering the Past and Imagining the Future.* A William H. Angoff Memorial Lecture presented at Educational Testing Service, Princeton, New Jersey.

Quellmalz, E.S., Timms, M.J., Silberglitt, M.D., & Buckley, B.C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching, 49*(3), 363-393.

Reese, L., Goldenberg, C., and Saunders, W. (2006). Variations in Reading Achievement among Spanish-Speaking Children in Different Language Programs: Explanations and Confounds. *The Elementary School Journal, 106,* (4), p 363-385.

Rogoff, B., Mosier, C., Mistry, J., and Göncü, A. (1993). Guided Participation in Cultural Activity by Toddlers and Caregivers. In *Monographs of the Society for Research in Child Development, 58*, (8), pp. 1-174.

Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J.V. Wertsch, P. del Rio, & A. Alvarez (Eds.), *Sociocultural studies of mind* (139-164). Cambridge, UK: Cambridge University Press. Reprinted (2008) in K. Hall & P. Murphy (Eds.), *Pedagogy and practice: Culture and identities*. London: Sage.

Rymes, B. (2010). Classroom discourse analysis: A focus on communicative repertoires. In N. Hornberger & S. McKay (Eds.), *Sociolinguistics and language education* (pp.528.548). Tonawanda, NY: Multilingual Matters.

Salazar, M. (2008). English or nothing: The impact of rigid language policies on the inclusion of humanizing practices in a high school ESL program. *Journal of Educational Equity and Excellence 41*(3):341–56.

Sato, E. (2010). Language for achievement: A framework for academic English language. WestEd. Retrieved from http://www.cde.ca.gov/sp/el/er/documents/achievementlang.pdf

Scarcella, R. (2003). *Academic English: A Conceptual Framework*. The University of California Linguistic Minority Research Institute. Technical Report 2003-1.

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education, 12*(4), 431-459.

Schleppelgrell, M. J. (2004). *The language of schooling.* New York, NY: Routledge.

Seluck, S. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research, 75*(3), 417-453.

Short, D. & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners* – A report to Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (p.112-133). New York, NY: Cambridge University Press.

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328*(5977), 450-452.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English language learners. *Educational Researcher, 32*(2), 3-13.

Solorzano, R.E. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*(2), 260-329.

Spolsky, B. (1989). *Conditions for second language learning: Introduction to a general theory.* Oxford: Oxford University Press.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Allyn and Bacon.

Uhrmacher, P.B. (1997). The Curriculum Shadow. *Curriculum Inquiry, 27*(3), 317-329.

Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal, 36*(2), 299-316.

Vygotsky, L. (1978a). Interaction between learning and development. In Guavain & Cole (Eds.) *Readings on the Development of Children* (p. 34-40). New York: Scientific American Books,.

Vygotsky, L. (1978b). *Mind in society*. Cambridge, MA: Harvard University Press.

World Class Instructional Design and Assessment (WIDA) (2015). *Annual Technical Report for ACCESS for ELLs.* Retrieved from http://www.cde.state.co.us/assessment/ela-additionalresources

Wolf, M., & Leon, S. (2009). An investigation of the Language Demands in Content Assessments for English Language Learners. *Educational Assessment, 14*(3/4), 139-159.

Young, J.W., Cho, Y., Ling, G., Cline, F., Steinberg, J., & Stone, E. (2008). Validity and fairness of state standards-based assessments for English language learners. *Educational Assessment 13*(2/3), 170-192.

Appendix A

Correlation Table for Overall CMAS Science Achievement and Overall ELP, RL, PL, and HL

	CMAS Overall	Overall ELP	Receptive Language	Productive Language	Home Language
CMAS Overall	1	.66**	.68**	.51**	-.35**
Overall ELP		1	.87**	.93**	.02
Receptive Language			1	.63**	.04*
Productive Language				1	-.01
Home Language					1

Note. * p < .05; ** p < .01

Correlation Table for CMAS Physical Science Achievement and Overall ELP, RL, PL, and HL

	CMAS Physical Science	Overall ELP	Receptive Language	Productive Language	Home Language
CMAS Physical Science	1	.57**	.42**	.59**	-.34**
Overall ELP		1	.87**	.93**	.02
Receptive Language			1	.63**	.04*
Productive Language				1	-.01
Home Language					1

Note. * p < .05; ** p < .01

Correlation Table for CMAS Life Science Achievement and Overall ELP, RL, PL, and HL

	CMAS Life Science	Overall ELP	Receptive Language	Productive Language	Home Language
CMAS Life Science	1	.59**	.46**	.59**	-.32**
Overall ELP		1	.87**	.93**	.02
Receptive Language			1	.63**	.04*
Productive Language				1	-.01
Home Language					1

Note. * p < .05; ** p < .01

Correlation Table for CMAS Earth Science Achievement and ELP, HL, and RL and PL

	CMAS Earth Science	Overall ELP	Receptive Language	Productive Language	Home Language
CMAS Earth Science	1	.57**	.43**	.59**	-.32**
Overall ELP		1	.87**	.93**	.02
Receptive Language			1	.63**	.04*
Productive Language				1	-.01
Home Language					1

Note. * p < .05; ** p < .01