USING THE ACCESS TEST AS A PREDICTIVE MEASURE OF ENGLISH LEARNER SUCCESS ON THE BIOLOGY END-OF-COURSE-TEST IN GEORGIA

by

SUSANNE WAKEMAN

JOHN PETROVIC, COMMITTEE CHAIR
DAVID DAGLEY
BOB JOHNSON
ROBERT SUMMERS
SARA TOMEK

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Education in the Department of Educational Leadership, Policy, and Technology Studies in the Graduate School of The University of Alabama

TUSCALOOSA, ALABAMA

2013
ABSTRACT

In order to be successful in an educational setting, a student must acquire a particular register of language specific to academia, or an academic register. It cannot be acquired through ordinary social intercourse. It is specific to school. Acquiring the academic register is the greatest challenge English Learners (EL) face during their academic careers. The federal government has recognized that challenge and has placed the burden, where it rightly belongs, on educators.

With the passage of the No Child Left Behind Act of 2001, schools, districts, and states are accountable for the annual progress of ELs’ language proficiency with an emphasis on the acquisition of the academic register. The goal is for all students to attain a high school diploma. In order for ELs to do this, they must not only learn vocabulary related content and concepts, but the language that crosses content areas such as ‘analyze,’ ‘infer,’ and ‘summarize.’ They must learn multiple meanings of seemingly simple words, such as ‘like,’ ‘left,’ and ‘right,’ which cause confusion for those who only know the more common definition in their social register.

English Learners are required to take federal and state mandated assessments of their content knowledge and their English language proficiency. In an era of data-informed instruction, these sets of data are available to the educator in order to best inform decisions related to the individual needs of each student. This study sought to identify a correlation between academic language proficiency as measured by the Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS) test, 2007
edition, and the Biology End-of-Course-Tests (EOCT) for 164 ninth graders at one high school in Georgia between the years 2007 and 2012.

A strong, positive correlation was found between the overall subscale score on the ACCESS and the biology ECOT test, meaning that if an EL scored high on the ACCESS, he or she would also score high on the biology EOCT. While all four of the language domains of listening, speaking, reading and writing had a positive correlation with the biology EOCT scores, writing had the greatest predictive value of the four. Analyses also included the three ACCESS composite subscale scores of oral, comprehension, and literacy. Literacy, which is made up equally of the reading and writing subscale scores, proved to have the greatest value of all of the ACCESS subscales for predicting an EL’s success on the biology EOCT.

Gender and length of time (LOT) in English to Speakers of Other Languages (ESOL) program were factored into a step-wise regression analysis and found to play useful roles in predicting EL success on the biology EOCT. In regard to gender, three out of four ELs who were successful on the biology EOCT were male. After the third year in an ESOL program, LOT had a negative effect on EL achievement. While these findings are consistent with recent research, an exact explanation for this phenomenon is not available in the literature. Additional data is necessary to determine the influence of school and program models, sociocultural factors, and individual differences (ID) on student success.
### LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS</td>
<td>Assessing Comprehension and Communication in English State-to-State for English Language Learners</td>
</tr>
<tr>
<td>AIR</td>
<td>American Institute for Research</td>
</tr>
<tr>
<td>AM</td>
<td>acquisition model</td>
</tr>
<tr>
<td>AMOs</td>
<td>Annual Measurable Objectives</td>
</tr>
<tr>
<td>AoA</td>
<td>age of acquisition</td>
</tr>
<tr>
<td>API</td>
<td>Academic Performance Index</td>
</tr>
<tr>
<td>AYP</td>
<td>Adequate Yearly Progress</td>
</tr>
<tr>
<td>B</td>
<td>parameter estimate</td>
</tr>
<tr>
<td>b</td>
<td>standardized regression parameter</td>
</tr>
<tr>
<td>BEA</td>
<td>Bilingual Education Act</td>
</tr>
<tr>
<td>CCRPI</td>
<td>College and Career Readiness Performance Index</td>
</tr>
<tr>
<td>CCSS</td>
<td>Common Core States Standards</td>
</tr>
<tr>
<td>CELLA</td>
<td>Comprehensive English Language Learner Assessment</td>
</tr>
<tr>
<td>CEP</td>
<td>Center for Educational Policy</td>
</tr>
<tr>
<td>CRCT</td>
<td>Criterion Referenced Competency Test</td>
</tr>
<tr>
<td>CRT</td>
<td>Critical Race Theory</td>
</tr>
<tr>
<td>DBP</td>
<td>developmental bilingual programs</td>
</tr>
<tr>
<td>EAA</td>
<td>English Acquisition Act</td>
</tr>
<tr>
<td>EEOA</td>
<td>Equal Opportunities Act</td>
</tr>
<tr>
<td>EL</td>
<td>English learner</td>
</tr>
<tr>
<td>ELA</td>
<td>English Language Arts</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ELD</td>
<td>English Language Development</td>
</tr>
<tr>
<td>ELDA</td>
<td>English Language Development Assessment</td>
</tr>
<tr>
<td>ELL</td>
<td>English language learner</td>
</tr>
<tr>
<td>ELPS</td>
<td>English language proficiency standards</td>
</tr>
<tr>
<td>EOCT</td>
<td>End of Course Test</td>
</tr>
<tr>
<td>ESEA</td>
<td>Elementary and Secondary Education Act</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second Language</td>
</tr>
<tr>
<td>ESOL</td>
<td>English to Speakers of Other Languages</td>
</tr>
<tr>
<td>FFEL</td>
<td>Flexibility and Focus in Education Law</td>
</tr>
<tr>
<td>GADOE</td>
<td>Georgia Department of Education</td>
</tr>
<tr>
<td>GHSGT</td>
<td>Georgia High School Graduation Test</td>
</tr>
<tr>
<td>ID</td>
<td>individual differences</td>
</tr>
<tr>
<td>IRB</td>
<td>Institution Review Board</td>
</tr>
<tr>
<td>ITBS</td>
<td>Iowa Test of Basic Skills</td>
</tr>
<tr>
<td>L1</td>
<td>first language</td>
</tr>
<tr>
<td>L2</td>
<td>second language</td>
</tr>
<tr>
<td>LAD</td>
<td>Language Acquisition Device</td>
</tr>
<tr>
<td>LEP</td>
<td>Limited English Proficient</td>
</tr>
<tr>
<td>LOT</td>
<td>Length of Time in an ESOL program</td>
</tr>
<tr>
<td>M</td>
<td>mean</td>
</tr>
<tr>
<td>MALDEF</td>
<td>Mexican American Legal Defense and Education Fund</td>
</tr>
<tr>
<td>MEG</td>
<td>magneto-encephalography</td>
</tr>
<tr>
<td>MPI</td>
<td>model performance indicator</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>NAEP</td>
<td>National Assessment of Educational Progress</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>NCLB</td>
<td>No Child Left Behind Act</td>
</tr>
<tr>
<td>NECAP</td>
<td>New England Common Assessment Program</td>
</tr>
<tr>
<td>NUSC</td>
<td>Nogales Unified School District</td>
</tr>
<tr>
<td>NWLC</td>
<td>National Women’s Law Center</td>
</tr>
<tr>
<td>OCR</td>
<td>Office of Civil Rights</td>
</tr>
<tr>
<td>OELA</td>
<td>Office of English Language Acquisition</td>
</tr>
<tr>
<td>OLS</td>
<td>ordinary least squares</td>
</tr>
<tr>
<td>P</td>
<td>probability value</td>
</tr>
<tr>
<td>PET</td>
<td>positron emission tomography</td>
</tr>
<tr>
<td>PK-K</td>
<td>pre-kindergarten – kindergarten</td>
</tr>
<tr>
<td>r</td>
<td>Pearson correlation coefficient</td>
</tr>
<tr>
<td>r²</td>
<td>r-squared statistic</td>
</tr>
<tr>
<td>REL</td>
<td>Regional Education Laboratory at Educational Development Center</td>
</tr>
<tr>
<td>SCASS</td>
<td>State Collaborative on Assessment &amp; Student Standards</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>SEI</td>
<td>Structured English Immersion</td>
</tr>
<tr>
<td>SES</td>
<td>socio-economic status</td>
</tr>
<tr>
<td>SIFE</td>
<td>students with interrupted formal schooling</td>
</tr>
<tr>
<td>SIOP</td>
<td>Sheltered English Observation Protocol</td>
</tr>
<tr>
<td>SLA</td>
<td>second language acquisition</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>t</td>
<td>statistical examination of two population means</td>
</tr>
<tr>
<td>TCAP</td>
<td>Tennessee Assessment Program Achievement Test</td>
</tr>
<tr>
<td>TVAAS</td>
<td>Tennessee Value Added Assessment System</td>
</tr>
<tr>
<td>UG</td>
<td>Universal Grammar</td>
</tr>
<tr>
<td>W-APT</td>
<td>WIDA-ACCESS Placement Test</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>WIDA</td>
<td>World-class Instructional Design and Assessment</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to thank my husband, Eric Wakeman, for his support and encouragement. I would not be in education if it were not for him. When we taught in Japan together, he noticed that I loved teaching English Learners. He suggested that I pursue a master’s degree and has pushed and prodded me all the way to this terminal degree. A good marriage makes you a better person than if you were alone; this he has made of me.

I would like to thank Dr. John Petrovic for chairing my dissertation committee. He has been an invaluable source of information. He has done what good teachers do; he has led me to deeper levels of understanding and scholarship while keeping this dissertation within my realm of possibility. Along with the dissertation committee, I thank him for his criticism and critique of my work along with many words of encouragement. Dr. David Dagley taught the very first course I took at The University of Alabama; I am pleased that he is present as I finish this adventure in learning. I thank the staff at the Gadsden Center, particularly Roger Woodard, for their help and support over the years.

Finally, I would like to thank my children, my parents, family, and friends who celebrate this accomplishment with me. My daughter, a future Ph.D. in geology, told me that she was going to get her doctorate before she had children so she could spend weekends with them. I thank her and her brother with all my heart for generously allowing me to spend so many weekends as a graduate student. I thank my parents for inculcating the importance of a good education in my life. I thank the rest of my wonderful family and friends for their constant words of support and encouragement which sustained me throughout this process. Hebrew 13:2.
REFERENCES ...........................................................................................................................................82

APPENDICES:

A  ACCESS TEACHER REPORT 2011 ........................................................................................................93

B  PERMISSION LETTER FROM THE SCHOOL DISTRICT ................................................................95
# LIST OF TABLES

1. Percentage of ELs Failing Ninth Grade EOCTs ................................................................. 5
2. Demographic Frequencies of ELs ....................................................................................... 55
3. Distribution of ELs by Length of Time in an ESOL Program (LOT) .............................. 56
4. Distribution of Biology EOCT Scaled Scores by LOT .................................................. 57
5. Linear Regression Analysis of EOCT and ACCESS Overall Subscale .................... 58
6. Linear Regression Analysis of EOCT and ACCESS Overall Subscale, Gender, and LOT ... 59
7. Linear Regression Analysis of EOCT and ACCESS Subscales for Language Domains ... 60
8. Linear Regression Analysis of EOCT and Combined ACCESS Subscales .................. 61
9. Correlation of Biology EOCT with Subscales Reported for the ACCESS Test ............ 62
10. Linear Regression Analysis of EOCT and ACCESS Overall Subscale, Gender, and LOT ... 63
LIST OF FIGURES

1  Number of ELs and overall subscale scores from the springs of 2007-2012 .........................46

2  The WIDA English language proficiency continuum ..........................................................48
CHAPTER 1
INTRODUCTION

English Learners (ELs) face linguistic challenges that native English speakers do not. They must acquire not only social, but also academic language, or “second language instructional competence” (MacSwan & Rolstad, 2003), in order to successfully manage school-related tasks such as participation in class discussions, reading texts, and writing informational reports (Diaz-Rico & Weed, 2002; Dutro & Moran, 2003; Hakuta, Goto, & Witt, 2000; Scarcella & Rumberger, 2000; Short & Fitzsimmons, 2007). Unless a child has some sort of pathology, the majority of children are proficient in their first language by the time they enter school. However, ELs “may exhibit linguistic errors [in the L2, English] of a sort that typically developing school-age children do not exhibit in their native language” (MacSwan & Rolstad, 2010, p. 175). Maturation is expected in L2 competence over time with the acquisition of both social and academic registers through the sustained experience of schooling. The challenge for English Speakers of Other Languages (ESOL) teachers is to help students acquire an academic register so that they can fully participate in their education in English.1

The No Child Left Behind Act of 2001 (NCLB) (U.S. Department of Education, 2004) requires states to assess the progress of EL language acquisition. In an effort to meet these mandates, the Georgia Department of Education (GADOE) turned to the World-class Instructional Design and Assessment (WIDA) consortium to provide English language

---

1 Research shows that it takes between 4 and 7 years for ELs to acquire an academic register depending on the student’s academic background in his first language and regardless of instructional method (Hakuta, 2002; Hakuta et al., 2000; Thomas & Collier, 2002).
assessment services for EL students. WIDA designed the Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS) test to measure both EL students’ social and academic language competency (Gottlieb, Cranley, & Cammilleri, 2007). The ACCESS tests EL language competence in the domains of speaking, listening, reading, and writing. The combined subscale score for speaking and listening is called the “oral proficiency score” (WIDA, 2012).² Reading and writing together are referred to as “language literacy.” Reading and listening make up the “comprehension” subscale score. The overall subscale score is a combination of all four domains. The ACCESS test also assesses students’ English language competence across the five content standards of social and instructional language, the language of language arts, mathematics, science, and social studies. ACCESS scores are used to determine whether an EL continues to receive ESOL services or exits the program (U.S. Department of Education, 2004; GADOE, 2012a; WIDA, 2011).

If ELs have passed a valid and reliable English language assessment, then it may be assumed that they are able to function at grade level in the regular classroom in English. However, there is provision within the GADOE (2012a) ESOL program guidelines that requires the academic success of exited ELs to be monitored for 2 years after exiting the program. This is to ensure that if exited ELs are not successful in the regular classroom without ESOL support, they can be brought back into the ESOL program and again receive language support (GADOE

² Gottlieb (2003) defines proficiency is being able “to use language accurately and appropriately in its oral and written forms in a variety of settings” (p. 244) WIDA describes a continuum of linguistic abilities, dividing them into what they call “proficiency levels” and rating those abilities with “proficiency scores” (2012). In doing so, WIDA imbues the term proficiency with an arbitrary sense of rightness, which is a social construction favoring one register over another. However, the word proficiency simply refers to how well, or the ease to which, one can do something; it cannot refer to first language acquisition which all humans are innately predisposed to accomplish (MacSwan & Rolstad, 2009; Petrovic, 2012). The second definition of proficiency is the meaning used in this paper. The term competence will be employed by the author to describe an EL’s ability to use the academic register accurately and appropriately.
By requiring the monitoring of exited ELs, there is an implicit understanding that English language assessments may not be reliable for all ELs. The WIDA (2011) consortium echoes this belief by encouraging states to adopt multiple criteria for exiting ELs from ESOL programs. The GADOE (2012) has chosen to supplement ACCESS scores with the Criteria Referenced Competency Test (CRCT) scores in reading for Grades 1 through 8 and with End-of-Course Test (EOCT) scores in Language Arts for Grades 9 through 12. If an EL scores within Level 4, Expanding, or Level 5, Bridging, on the ACCESS test, he/she must also have a passing score on the CRCT in reading or on the language arts EOCTs in order to exit ESOL services. The implication is that CRCTs and EOCTs scores outweigh the ACCESS in predicting EL success in the regular classroom.

Statement of the Problem

Secondary educators are searching for ways to predict students’ success on standards-based tests. Researchers have studied the predictive value of standards-based assessments in earlier grades to determine if there is a relationship between those scores and high school graduation tests. A positive relationship has been found in California, Tennessee, Texas, and Georgia between standards-based test scores and high school graduation test scores (Demps & Onwuegbuzie, 2001; Greeson, 2009; Heatley, 2002; Hobbs, 2005; McNamara, 2004; Wood, 1999). However, the data were not disaggregated for ELs. In New England and Georgia, researchers found the ACCESS test is a significant predictor of EL performance on standards-based assessment in the content areas for elementary and middle grades students (Baker, 2011; Parker, Louie, & O’Dwyer, 2009).
In Georgia, a study by the Department of Education has shown a moderately positive relationship between ACCESS scores and the EOCT in ninth grade literature and composition (Walker, 2008). Of the research studies included in the literature review, the study conducted by the GADOE (Walker, 2008) was the only one that applied the knowledge gained in some useful way for ESOL teachers. Correlating the two test scores, in order to determine what ACCESS score is necessary for an EL student to be successful on an EOCT, is helpful to educators in both administration and in the classroom. The GADOE compared ACCESS scores to ninth grade Literature and Composition and American Literature EOCTs. This is significant, because the GADOE requires ELs to not only have an ACCESS score of 4.0 or higher in order to exit ESOL services, but they must also pass the standards-based reading test for their grade level. For ninth graders, this could be one of the aforementioned EOCTs.

The GADOE’s study (Walker, 2008) shows that an EL will most likely pass the ninth grade Literature and Composition EOCT with a 4.3 in overall proficiency on the ACCESS and the American Literature EOCT with a 4.8 in overall proficiency on the ACCESS. The study also found that the reading score was the strongest predictor of success on the Literature and Composition EOCT. From this information we could conclude that while the required ACCESS score for exiting ESOL services is somewhat low for ninth graders, requiring multiple criteria exhibits perspicacity on the part of the GADOE. However, far more ELs pass ninth grade literature EOCTs than Biology EOCTs. As seen in Table 1, ninth grade Limited English Proficient (LEP) students in Georgia consistently failed the Biology EOCT in greater numbers than any other subject from the 2007-2008 to 2010-2011 school years (GADOE, 2012b). This might suggest that a higher level of academic English competence may be necessary in order to pass the Biology EOCT than the English Literature EOCTs in ninth grade. If so, the current
Exiting criteria for the ESOL program may not be predictive of success for ninth grade ELs in the regular classroom in all content areas.

Table 1

Percentage of ELs Failing Ninth Grade EOCTs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature</td>
<td>54%</td>
<td>55%</td>
<td>57%</td>
<td>67%</td>
</tr>
<tr>
<td>Math I</td>
<td>61%</td>
<td>60%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Biology</td>
<td>63%</td>
<td>64%</td>
<td>68%</td>
<td>69%</td>
</tr>
<tr>
<td>Economics</td>
<td>55%</td>
<td>53%</td>
<td>57%</td>
<td>61%</td>
</tr>
</tbody>
</table>

In order to explore the relationship between the ACCESS test and the Biology EOCT, this study will perform analyses similar to those done by the GADOE (Walker, 2008). It will attempt to advance the search for ways to predict students’ success on standards-based tests by examining the relationship between the ACCESS scores of ninth grade ELs and their scores on the high school EOCTs in Biology.

Significance of the Problem

A study of the relationship between ACCESS and Biology EOCT scores is significant for two reasons. The first reason is to discover if there is a predictive relationship between the two sets of data. If exiting scores on the ACCESS test predict success on the Biology EOCT, then educators can use the information to identify students for intervention programs in critical content areas rather than waiting for EOCT results to determine which students need remediation. This will be helpful to schools concerned with achieving Adequate Yearly Progress for the EL subgroup as required by NCLB (U.S. Department of Education, 2004). The second
reason this study is significant is to provide insight into the relationship between the two data sets and whether the Biology ECOT score should be included in determining whether an EL is ready to exit an ESOL program. Presently, the decision to exit an EL from ESOL services is made using the ACCESS overall subscale score and the ninth grade Literature and Composition EOCT score only.

Research Questions

The following research questions were developed for this study.

1. How does performance on the ACCESS test of English language competence predict EL performance on the Biology EOCT in ninth grade after controlling for student characteristics?

2. Which language domain is the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing?

Assumptions

The ACCESS test determines whether an ELL student has acquired academic language competency at grade level. Therefore, an ELs prior educational experience, motivational level, number of years the student has lived in the United States, gender, ethnicity, socioeconomic status, disability status, years in ESOL programs, and age were assumed neutral to this study. Because it is administered to every EL in the state of Georgia, the size of the student population, racial make-up, socioeconomic status, and EL density of the student body of a school was assumed to be neutral to this study also.
The ACCESS tests the English Language Proficiency Standards (ELPS) designed by WIDA. Federal guidance requires that at least one content standard per grade-level cluster be linked to the ELPS (Chi, Garcia, Surber, & Trautman, 2011; U.S. Department of Education, 2003, 2004). “Linking assures that linguistic discourse elements associated with the language of mathematics are included in language proficiency standards” (Chi et al., 2011, p. 10). WIDA’s assumption is that if a student is proficient in the general academic language of a content area, then he should be able to access the content and be successful in the regular class without ESOL services.

It is further assumed that ACCESS testing procedures were consistent in all schools in Georgia due to the annual training that WIDA (2011) requires ESOL teachers to complete shortly before the test is administered. The training culminates in a test that ESOL teachers must pass in order to administer the ACCESS. If the teacher passes the test at 80% or above, the teacher is considered qualified to administer the test (GADOE, 2012a). It is further assumed that EOCT procedures were consistent due to training that proctors of the test are required to attend before each administration. Therefore, the testing environments and procedures were similar for every student taking both tests and were assumed to be neutral to this study.

Limitations

The purposive, random sampling procedure decreased the number of students participating in the study. This limited the generalizability of the study to other populations, particularly to schools in states that do not use the ACCESS test. The study was limited to five cohorts because Georgia began using the 2007 edition of the test during the 2007-2008 school years (GADOE, 2012a).
The ACCESS test reports results in four domains of language: listening, speaking, reading, and writing, as well as the combined domains of oral, literacy, comprehension, and overall proficiency (WIDA, 2011). It includes items from five ELPS: the language of social and instructional language, language arts, mathematics, science, and social studies. While an EOCT assesses student knowledge of the course content, the ACCESS purports to test general academic language and not be content specific (Johnson, 2009). This is a crucial distinction for WIDA, because the ACCESS test is divided into grade level clusters.

The ACCESS has five grade-level clusters: Pre-kindergarten-Kindergarten (PK-K), 1-2, 3-5, 6-8, and 9-12. For this reason, ACCESS scores are scaled depending on the EL’s grade level and according to language domains and not academic subjects (GADOE, 2012a). The limited number of items on the ACCESS makes it difficult to compare content area scores to EOCT scores (Parker, et al., 2009). There are between 6 and 20 questions per content area taken from the content standards of between two to four grade levels. In addition, the questions are drawn from content matter common to 27 separate state content standards. Because the results of the study are predictive and not causal, educators may not use the data to evaluate ESOL instructional programs used at the schools.

Delimitations

The data used in this study were drawn from the results of the 2007 edition of the ACCESS test, which sought to link content questions to content common to 27 separate state standards. The conclusions drawn from this study may not reflect the changes in the 2012 edition of the ACCESS, which is based on the Common Core States Standards Initiative (CCSS) (2012) of which Georgia is a part. The district and school were not randomly selected.
CHAPTER 2

LITERATURE REVIEW

It is important to set the use of assessments into a historical context in order to understand how the purpose has changed depending on the needs of society and the spirit of the times. Today, standards-based assessment has become linked to high-stakes testing. This has led to implementation of data-informed instruction in order to close achievement gaps between high and low performing students so that no child is left behind in public schools. Educators are searching for ways to use standards-based assessments to predict student success in secondary schools in order to provide the learning necessary for all students to achieve at high levels. This review of literature provides an overview of standards-based assessment, its value as a predictor of student success on tests, the history of legislation affecting ELs, English language assessment, the challenges inherent in second language acquisition, and the ESOL program in Georgia.

Standards-Based Assessment

Assessment, in order to determine achievement and critical thinking, has long been a part of education in the west. In the 14th century, Joan Celc introduced examinations into the curricular organization of his Latin school. Because of the large student population, he divided the curriculum into eight forms and employed assessments to determine which students to promote to the next level of learning. This criterion referenced view of assessment prevailed until the 19th century when it became associated with norm-referenced assessment through psychological testing as a means of intellectual measurement. Now, as in the past, assessment
takes the form of questions and answers. Success entails answering as many questions correctly as possible (Wilbrink, 1997). Recently, in the United States, the greatest change in student assessment concerns who determines what is to be learned. There has been a shift from the educator to the state. Standards-based assessment provides the accountability state and federal governments require to ensure that all students master common standards (Wang, Beckett, & Brown, 2006).

Today, assessment has been linked to high-stakes testing through NCLB with the threat of high-stakes consequences for schools that do not meet state determined goals for Adequate Yearly Progress (AYP) (Goertz & Duffy, 2003; Wang et al., 2006). NCLB contends that “accountability is a crucial step in addressing the achievement gaps that plague our nation” (U.S. Department of Education, 2004, p. 18). Included in the list of groups of students who are most at risk are ELs. In this way, NCLB has become a driving force for school reform by placing emphasis on each sub-group’s success rather than on the student population as a whole.

In 1999, California anticipated NCLB by implementing the Academic Performance Index (API). The purpose of API in California is similar to that of AYP for the nation as a whole. Both measure the academic attainment and yearly progress of schools, disaggregating the data in order to report the achievement of sub-groups of the student population (California Department of Education, 2009; U.S. Department of Education, 2004). Prior to the implementation of API, Deerfield High School in California was considered a high achieving school. However, its success was not based on the achievement of all of its students. Only 10% of the student body was responsible for the academic success of the entire school. With the implementation of API, no longer could a school be considered high achieving when 90% of its students were failing. API forced educators at Deerfield to focus on the achievement of each sub-group of the student
Standards-based assessment has not only been used as a summative instrument, but also for formative purposes as well. When assessment is used as a predictor of future achievement, it becomes a useful tool for educators. Gewertz (2007) tells of a school in Massachusetts that requires students who have failed the exit exam to substitute remediation courses for electives. However, The Center for Educational Policy (CEP) states that “while remediation for students who have failed an exit exam is essential, interventions that prepare students for exit exams are also crucial” (2007, p. 41). The CEP (2007) reviewed the data from five states to see what types of interventions they have in place to help students before they take the exit exam. One of the interventions included was using assessment data to inform instruction in order to focus on content that will be tested.

Other studies report the use of assessment data to predict student success. Heatley (2002) describes how Deerfield High School in California successfully used API data and other data to target students for remediation before taking the state exit exams. The high school was not selected at random. Instead, it was chosen because the district had a data system design linked to state mandates and the student population reflected the ethnic mix of the state at large. Using qualitative research methods, Heatley drew from interviews, observations, surveys, and document reviews to analyze a school district engaged in the ongoing use of data to inform instruction. Through the use of district-enhanced, state content standards; quarterly benchmark testing aligned with those standards; quarterly staff training in the use of disaggregated data; and curriculum teams, which included special education and EL staff charged with designing instruction aligned with the standards; all stakeholders were aware of the district-wide
expectation to use data to inform instruction. Students performing below grade level on quarterly benchmark tests were immediately identified for additional instruction in that area. The result of these synchronized efforts was an improvement in student performance that brought about Deerfield High School being identified as a National Blue Ribbon School with 98% of high school graduates gaining admission to college.

Another example of the use of data to inform instruction was provided by Morrison, Morrison, and Bedford (2007). Their study described a process for using assessment data to identify students’ individual needs and provide them with the necessary instruction to succeed on standards-based tests. It was called the Strategic Monitoring/Intervention Process. The key is to focus on subgroups that are in danger of failing AYP. First, administrators identify the AMO, or the percentage, of the subgroup that must pass in order for the subgroup to make AYP. The percentage is multiplied by the number of students in the group. This is the number of students in the subgroup that must pass the test in order for the school to make AYP. Next, a spreadsheet is created of the students in each subgroup arranged from those who are most likely to pass the test to those most likely to fail. A student’s chances of success or failure are based on previous test scores and other data the school feels is relevant. Then, a line is drawn on the spreadsheet showing how many of those students must pass in order to make AYP. Interventions are then matched with the needs of those students likely to fail the test. Progress is monitored for all students on the list and the order and interventions may change depending on student progress. By focusing on the individual student and providing strategic help in areas of academic weakness, Argyle Middle School, and other schools in Maryland, made AYP in 2006.

NCLB requires states to be accountable for the academic progress of all students as proven by standards-based assessment (U.S. Department of Education, 2004; Wang et al., 2006).
Davis (2004) found that there was a significant relationship between grades in content courses such as Biology I and whether a student passed the Mississippi Subject Area Test in Biology I, when the content taught in class matched the standards that were tested on the graduation exam. Thus, standards-based assessment has led states and school districts to implement data-informed instruction (CEP, 2007; Heatley, 2002; Peterson, 2007).

Wood (1999) researched the relationship between the English Language Arts (ELA) portion of the GHSGT and students’ eighth grade reading scores on the Iowa Test of Basic Skills (ITBS). Wood’s sample included 177 students who graduated in 1997, 1998, and 1999. In addition, only the score from the first time the students took the GHSGT their junior year was included in the study. Wood disaggregated the data for female, male, Caucasian, and Black students. No other ethnicity was enrolled in these classes in the high school at the time, nor was there mention made of special education or EL students being included in the sample. There was no variation in the results of the study in regard to gender, ethnicity, or year of graduation. The data consistently showed a significant relationship between the two tests leading the researcher to conclude that scores on the ITBS could be used as a predictor of future success on the GHSGT.

Greeson (2009) found a predictive relationship between eighth grade CRCT scores and the GHSGT. He discovered that if a student passed a content area assessment on the CRCT, then it was highly probable that the student would also pass that section of the GHSGT in 11th grade. In addition, only the science CRCT was able to predict success on all sections of the GHSGT. The language arts CRCT only predicted success on the language arts section of the GHSGT. The data were disaggregated for ethnicity and gender, but the results were insignificant. Greeson’s recommendation was for districts to use eighth grade CRCT scores to predict student
failure on the GHSGT and target those students for intervention programs in order to better prepare them for success on the GHSGT.\(^3\)

Demps and Onwuegbuzie (2001) also explored the relationship between students’ scores on all five subtests of the GHSGT and the Iowa Test of Basic Skills (ITBS) in the eighth grade. The Demps and Onwuegbuzie sample consisted of 102 students in the class of 1999 from a rural high school in Georgia. Of the students, 78 were African-American, 22 were Caucasian, 1 was Asian American, and 1 was Hispanic. No mention was made of special education or EL students being included in the sample. A significant correlation was found between the eighth grade ITBS reading test scores and those on all subtests--writing, ELA, math, social studies, and science--of the GHSGT. There was a stronger correlation between the ITBS reading test and the GHSGT in ELA and social studies for all students regardless of ethnicity. Females were found to have stronger correlations between the ITBS reading test and the GHSGT in writing, math, and social studies. Students who scored in the bottom quartile on the ITBS reading test failed some portion of the GHSGT. The authors recommended targeting these students for interventions early in their high school career.

Mando (2007), in an unpublished dissertation, reported a strong correlation between ACCESS scores and the eighth grade Math CRCT of 187 ELs in one school district in Georgia. This was the first year that the ACCESS test was used in Georgia. ELs with high English proficiency scores also had high CRCT scores. While the GADOE grants testing waivers to ELs for 1 year after arriving in the United States--the ACCESS test substituting for the reading and

\(^3\) Greeson’s district is situated near the school district included in this study. Drawing on internal research similar to Greeson’s, the district has begun identifying students by their CRCT scores for inclusion in intervention programs including an invitation to enroll in a new high school whose mission is to help students who do not do well in the traditional educational setting (Murkerson, 2009; Rojas, 2012). ELs were initially excluded from this opportunity due to programmatic issues, which have since been resolved (Rojas, 2012).
ELA CRCT--they do have to take the standardized test in mathematics. Mando (2007) observed that standardized tests in mathematics have become less computational and more problem solving. She recommended granting ELs a waiver for the math test until they reach a certain level of language proficiency; however, she did not specify what that level should be.

McNamara (2004) researched the relationship between the Tennessee Assessment Program Achievement Test (TCAP) and high school graduation exam scores. McNamara sampled 200 students from one county high school in middle to upper socioeconomic neighborhoods and 35 students from two inner-city high schools in lower socioeconomic neighborhoods during the 1998-1999 school years. The data were not disaggregated for ethnicity. McNamara found a significant correlation between eighth grade TCAP scores on the Mathematics General Subtest, the Total TCAP score, and first semester grades in Algebra 1. The strongest predictor of success on the Gateway Algebra 1 was the numeric course grade. The total TCAP score was more significant than the Mathematics General Subtest. The students identified as most likely to succeed on the Gateway Algebra 1 test were male students from the county high schools.

Hobbs (2005) also investigated the relationship between the TCAP in Grades 3, 5, and 8 and the Gateway high school exiting subtests in English, math, and science. Hobbs’ sample included 314 students from one school district in Tennessee from 2001 to 2003. The findings indicate that the eighth grade TCAP scores in science and language are the strongest predictors of success on the each of the Gateway subtests. The Tennessee Value Added Assessment System (TVAAS) was also evaluated as a predictor of success on the Gateway. Each subtest of the TVAAS correlated with the same subtest on the Gateway. Ethnicity was not considered a factor in this study, because 99% of the student population was White. Students with higher SES
performed better than those with lower SES; however, the predictive value of the TCAP was the same for both groups. Students without disabilities performed better than those with disabilities. In addition, there were fewer predictors for disabled students.

The Regional Education Laboratory at the Educational Development Center (REL) published a study (Parker et al., 2009) in which ACCESS data from Rhode Island, New Hampshire, and Vermont and data from the New England Common Assessment Program (NECAP) were used to determine whether an EL’s performance in the four language domains could predict success on the standards-based content assessments in fifth and eighth grades. The results showed a positive correlation between ACCESS scores in the domains of language literacy, reading and writing, and outcomes on reading, writing, and math content assessments. Parker et al. (2009) also found that each additional year spent in an ESOL program after the fifth year, resulted in significantly lower scores on the NECAP in reading, writing, and math.

Baker (2011) studied 1,059 Hispanic ELs in Grades 1 through 5 in a district in northwest Georgia. Participants had to have taken both the ACCESS and the reading and mathematics CRCT tests in 2010. A moderately positive relationship was found between ELs’ ACCESS and CRCT scores. Variation was due more to disability status, years in ESOL, and grade level, than an EL’s proficiency level or gender. ELs with disabilities scored significantly lower than ELs without disabilities. There was no significant difference in the scores of ELs who had been in an ESOL program for 3 or more years; however, there were far fewer participants in Grades 3 through 5 (351) than in first and second grades (608). ELs who scored 800 or above on the reading CRCT also had higher overall scores on the ACCESS.

Flores, Batalova, and Fix (2012) researched the educational trajectories of ELs in Texas. Their sample included two groups: one consisted of all students; the other was made up of
students who entered kindergarten in 1995 and graduated on-time in 2006 with no gaps in their enrollment. Similar to the high school in this study, a high percentage of the EL population in 6th through 12th grades were born in the United States. Their study found that former EL students belonging to the on-time cohort were in general more successful than any other group of ELs. Flores et al. (2012) identified other factors correlating with student success. For example, ELs who worked while in high school were more likely to attend college. While poverty negatively influenced student trajectories, access to college-credit courses in high school was a positive factor in students continuing their education after high school graduation. Long-term ELs, spending 5 or more years in a support program, were the least successful due to poor academic literacy skills in both English and their first language caused by gaps in their education.

The Georgia Department of Education (Walker, 2008), in an effort to give direction to educators deciding whether an EL should be placed in ninth grade Literature and Composition or tenth grade American Literature and Composition courses, undertook a study that sought to determine the ACCESS score necessary for an EL to be successful in the regular class. Comparing the scores of 962 ELs, the GADOE found that a moderately strong, positive relationship existed between the scaled scores for the EOCTs and the overall composite score on the ACCESS. Because the overall composite score on the ACCESS test is weighted by domain (35% reading, 35% writing, 15% listening, and 15% speaking), the state was interested in which domain had the strongest correlation when weighted equally. The results show that the ACCESS reading subscale score significantly outweighed the other domains in predicting success in the ninth grade Literature and Composition and tenth grade American Literature and Composition courses. It was also more predictive than the overall composited score.
The GADOE study (Walker, 2008) showed that the greater the EL student’s proficiency, the higher the scaled score on the EOCT. However, no EL scored in the upper range of the scaled scores in either of the Literature and Composition courses. In a range from 200-600, a 400 is necessary to pass the course. While some ELs scored 6.0 on the ACCESS, allowing them to exit ESOL services with a perfect score indicating fluent English proficiency, the highest score that an EL achieved was 484 on the ninth grade Literature and Composition EOCT and 474 on the American Literature and Composition EOCT. The study recommended that ELs should have 4.8 on the ACCESS overall composite score in order to be placed in a Literature and Composition course. Furthermore, an EL should score 4.3 on the ACCESS overall composite score in order to be placed in American Literature and Composition course.

History of Bilingual and ESOL Legislation since the Civil Rights Act

Over the decades, the federal government has gradually defined what it means to be an EL in public schools. Title VI of the Civil Rights Act of 1964 forbids agencies receiving federal funds from discrimination based on ethnicity, color, or national origin. However, requiring equal access to a public education did not guarantee that states would comply. Under Title I and Title III of the Elementary and Secondary Education Act (ESEA) of 1965, 18 ESOL and bilingual programs in the Southwest were receiving funding for Spanish-speaking students (Crawford, 1999). Unfortunately, this did not provide equal access to every EL in the Southwest, let alone the whole United States. A new title was necessary to address the needs of all children whose first language was not English.

In 1968, President Lyndon Johnson, a former elementary school teacher from Texas who understood the plight of ELs, signed the Bilingual Education Act (BEA) into law. It was also
known as Title VII of the Elementary and Secondary Education Act (ESEA). For the first time, the federal government committed monies specifically for EL educational programs, teacher training, resource development, and parental involvement. At the time, BEA was viewed as compensatory and limited to students of low socioeconomic status (SES). Even so, the federal government did not fund BEA the first year. The second year, BEA received $7.5 million, which doubled the number of ELs receiving support services to 27,000 (Crawford, 1999). BEA also required the inclusion of some sort of bilingual education in a proposal in order to qualify for grants. However, the fundamental, political issue that would plague BEA was whether the federal government was appropriating monies for the maintenance of an EL’s first language or whether the goal of public education was English acquisition.

While federal and state governments were legislating education for limited English proficient (LEP) students, parents of ELs began suing local school districts using the premise outlined in Crystal City, Texas, during the student boycotts of 1970. They claimed that equal treatment under the law did not equate to equal opportunities for success for their children. This was evident by low test scores and high drop-out rates for ELs. The Office of Civil Rights (OCR) took matters into its own hands, issuing a memorandum to school districts with a student population of more than 5% LEP requiring them to provide support services for these students. Case by case, lower federal courts began to define what it meant to be and how to instruct EL students.

In 1974, parents of children of Chinese descent brought a class action suit against the San Francisco Unified School District for failing to educate their children. When the lower federal courts ruled in favor of the school board, the parents ultimately appealed the decision to the
Supreme Court of the United States. In *Lau v. Nichols*, the Supreme Court made the following ruling:

The failure of the San Francisco school system to provide English language instruction to approximately 1,800 students of Chinese ancestry who do not speak English, or to provide them with other adequate instructional procedures, denies them a meaningful opportunity to participate in the public educational program and thus violates § 601 of the Civil Rights Act of 1964, which bans discrimination based ‘on the ground of race, color, or national origin,’ in ‘any program or activity receiving Federal financial assistance,’ and the implementing regulations of the Department of Health, Education, and Welfare. (pp. 565-569)

While the Supreme Court refused to dictate pedagogy, ELs were now protected by case law. The subsequent reauthorization of ESEA in 1974 dropped the low SES requirement for participation. The number of ELs who now qualified for services ballooned, but funding was limited to only 6% of the eligible population. The question concerning the purpose of BEA--first language maintenance or English acquisition--still remained unanswered.

Using the *Lau v. Nichols* decision as a platform, the U.S. Commissioner of Education, Terrel Bell, developed guidelines for public schools serving ELs. Known as the Lau Remedies, these guidelines went beyond the Supreme Court’s decision. The Lau Remedies required school districts to identify ELs and evaluate their English proficiency in order to receive services and to exit them from support programs. It established student/teacher ratios, required specialized teacher training, and provided a timeline for implementation. The OCR began enforcing the Lau Remedies with the power to deny federal funding to districts that did not comply. For districts that violated EL rights, bilingual education was prescribed for elementary schools, but allowed ESOL in the secondary grades.

Despite the passage of the Equal Education Opportunities Act (EEOA) of 1974 prohibiting discrimination against students, ELs continued to experience segregation in schools. EEOA bars racial segregation of students and compels districts to dismantle barriers to equal
participation in educational programs. Despite the passage of EEOA, the American Institute of Research (AIR) revealed that a majority of ELs continued being segregated from regular instructional programs after achieving English language fluency.

In 1978, AIR caused a furor when it published a report stating that 86% of ELs stayed in bilingual programs after they were redesignated fluent English speakers. The inference was that BEA had inadvertently created a segregated population. Finally, a decision had to be made. Seeing the two goals as mutually exclusive, the debate was narrowly framed as to whether the goal of BEA was first language maintenance or English language acquisition. In the 1978 reauthorization of ESEA, Congress amended BEA stating that a student’s L2 could be used “only to the extent necessary to allow a child to achieve competence in the English language” (p. 128). The focus of federally funded EL instruction would be on students acquiring academic English with the goal of participating successfully in the mainstream classroom.

In 1981, EEOA played a pivotal role in deciding Castañeda v. Pickford (Crawford, 1999). The plaintiff sued saying his daughters were being segregated from other children because of their first language and ethnicity. He also charged that the children were not learning because Raymondville Independent School District was not providing bilingual education programs at his daughters’ grade levels. Castañeda further argued that there was no mechanism for evaluating the EL support programs implemented by the school district.

The Fifth Circuit Court of Appeals ruled in favor of Castañeda. In doing so, they went beyond Lau v. Nichols and invoked EEOA in its reasoning. The Court stated that whether or not the district intended to discriminate, it had a “specific obligation” to provide appropriate instruction for ELs. The Court went on to define what “appropriate action” meant, using a three-part test as to the efficacy of instruction used to support ELs. First, the district must use “a
program informed by an educational theory recognized as sound by some experts in the field” (Castañeda v. Pickard, 1981). Second, there must be sufficient resources to implement the program. Third, the program’s effectiveness must be evaluated after a trial period.

The impact of Ronald Reagan’s presidency was not in the changes to Title VI in the 1984 reauthorization of ESEA. While additional grants became available for family literacy programs and special EL populations such as gifted and special needs students, the major impact of Reagan’s presidency was felt in his view of Civil Rights. The OCR had “to prove ‘discriminatory intent,’ rather than merely documenting ‘discriminatory effects,’ to find that a district had violated Title VI of the Civil Rights Act” (Crawford, 1999, p. 56). Enforcement of the Lau Remedies dropped precipitously. However, a 2005 memorandum from OCR reiterates their usage of the three-part Castañeda standard in determining whether a school is in compliance with Title VI and EEOA.

The OCR’s memorandum (2005) cited several cases in which the three-part Castañeda standard was used to prove whether a school was in compliance with federal mandates or not. In Keys v. School District No. 1, Denver (1983), the courts found that the district failed the second prong of the Castañeda standard, because EL teachers did not have the necessary skills to provide comprehensible instruction for students. In addition, Hispanic students were not on parity with their English-speaking peers, in terms of the drop-out rate as well as in the use of comparable curriculum materials. Conversely, in Teresa P. v. Berkeley Unified School District (1989), the court found that ELs were performing as well as or better than their English-speaking peers and thus met the third prong of the Castañeda standard.

Two other cases cited by the OCR (2005) in its memorandum hold both state and local educational agencies responsible for the monitoring of EL programs through EEOA. In Idaho
Migrant Council v. Board of Education (1981), the 9th Circuit Court defined the term “state agency” as including both state and local school boards. This set the scene for Gomez v. Illinois Board of Education (1987). At the time, Illinois only provided services for ELs when there were 20 or more ELs in a school. The plaintiffs charged that the state had abrogated its duty in not providing adequate guidelines in identifying ELs such that many schools were able to report less than 20 EL enrolled and thus did not have to provide services for those students. After establishing that the state is charged with supervising school districts and enforcing regulations, the 7th Circuit Court ruled in favor of the plaintiff basing the ruling on the second prong of the Castañeda standard. It remanded the case to the lower court noting that the third prong of the Castañeda standard was not addressed because there were no services at those schools to monitor.

In an ongoing case beginning in 1992, parents of ELs reenrolled in the Nogales Unified School District (NUSD) filed suit alleging that the district was in violation of EEOA. Flores v. State of Arizona (1992) claimed that the district was not providing for their children’s educational needs, particularly stating that the support program was underfunded and had insufficient resources. In 2002, the 9th Circuit Court ruled in favor of the parents, requiring the state of Arizona to fund EL programs. This narrow focus on funding brought the suit to the attention of the Supreme Court. In 2009, the Supreme Court reversed and remanded the lower court decision because of subsequent state and federal legislative action that remediated the EEOA violations (Horne v. Flores, 2009).

Litigation in Flores v. Arizona (1992) lasted almost two decades and covered three presidencies, all eager to fulfill their campaign promises to reform public education. Added to this were changing governors, judges, lawyers, school board members, and legislators both at the
state and federal levels. The Supreme Court observed that the conditions that prompted the original law suit in 1992 (*Flores v. Arizona*) did not exist in 2009 (*Horne v. Flores*).

The Supreme Court noted in its decision (*Horne v. Flores*, 2009) that NUSD had made strides to provide better instruction for ELs by improving teacher quality, which brought the district into compliance with federal guidelines. One of the reasons for this was the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983). *A Nation at Risk* pointed out that in general, the nation’s teachers lacked the training to effectively teach ELs. In 1984 with the reauthorization of ESEA, the federal government responded to *A Nation at Risk* with “a needs assessment of qualified teachers and provided for pre-service and in-service training to make sure there were enough qualified teachers” (Bunch, 2011, p. 326) for ELs.

In the interim, the State of Arizona passed legislation resolving the funding discrepancy cited in the suit. The state also adopted an instructional delivery method based on sound educational theory in accordance with the *Castañeda v. Pickard* (1981) decision rendering that issue moot (*Horne v. Flores*, 2009). In addition, the NUSD also lowered the student/teacher ratios and EL class size in response to the assertions made in the *Flores v. Arizona* suit. Finally, the programmatic changes required by NCLB (U.S. Department of Education, 2004), such as providing “evidence of the progress and achievement of Nogales’ ELL students through its assessment and reporting requirements” (*Horne v. Flores*, 2009), served to settle the original claims in *Flores v. Arizona* (1992). The Supreme Court’s opinion was that EEOA was concerned with instructional quality, not fiscal equity, and that the lower court should have focused on program outcomes rather than the amount of funding it was receiving.

While *Flores v. Arizona* was being litigated, the Clinton Administration was leaving its mark on educational reforms. In Title I of the Improving America’s Schools Act of 1994, ELs
were identified with other at-risk groups as needing to meet the same content standards as fluent English speaking students in order to succeed academically. In order to measure student achievement, standards-based assessments were necessary for which the act provided funding.

The next major legislation to affect ELs came with the passage of the emotively named NCLB. As in past reauthorizations of the ESEA, NCLB (U.S. Department of Education, 2004) added new levels of accountability to the education of ELs. In order to highlight the purpose of EL instruction, BEA was renamed the English Acquisition Act (EAA) or Title III of NCLB. The Office of English Language Acquisition (OELA) was established to administer the new federal mandates for ELs (Bunch, 2011). Besides requiring states to teach ELs academic language, they must also assess ELs in the four language domains of reading, writing, speaking, and listening. States must show that a certain percentage of ELs progressed in academic language acquisition annually. The impetus for such mandates comes from the continuing gap in achievement between fluent English speakers and ELs.4

Because ELs must not only acquire English, but also learn content, there is consensus among educators that bilingual programs best prepare ELs to be successful in school (MacSwan & Rolstad, 2009; Ramirez, Yuen, & Ramey, 1991; Thomas & Collier, 2002). The implication is that the achievement gap reflects the accumulated effects of not understanding academic content while learning English. However, the federal government may not prescribe bilingual education

---

4 Not all ELs fail to achieve academic success. Research affirms that children of any language background who come from middle or upper class homes do better in school (MacSwan & Rolstad, 2010). Socio-economic status (SES) has proven to be a consistent predictor of success in school (Coleman, et al., 1966; Park, et al. 1995). Since the majority of ELs come from homes of lower SES, the statistics for the majority of ELs are grim. The 2000 census reported that 19% of school age children have immigrant parents. Of those children, 14% were born in the United States and 5% were foreign born. Only 6% of these children were described as LEP in that they spoke English “less than very well” (Capps, et al., 2005). The Urban Institute (2006) reported that 51% of children born to immigrant parents came from poverty, 32% of their parents did not have high school degrees, and 58% had at least one parent who was also LEP.
as a remedy to the achievement gap. The Blackmun Caveat to the *Lau v. Nichols* (1974) decision stated that if bilingual education is too costly for a district, English as a Second Language (ESL) services could alternatively be provided. In addition, the 1978 reauthorization of ESEA limited the scope of bilingual instruction to the support of English language acquisition, which, somewhat paradoxically, has been used to limit bilingual education programming.

First and Second Language Acquisition

The study of language acquisition takes the form of explaining what happens when a person acquires a language. It is agreed that humans are innately predisposed to acquire a first language (L1) (Chomsky, 2006). Chomsky (1968, 2006) proposed that language acquisition happens within the brain through a device he first presented as the Acquisition Model (Lenneberg, 1967), later referred to as the Language Acquisition Device (LAD) (Ellis, 2006). His hypotheses were developed in opposition to behaviorism, which stated that all learning was the result of some sort of stimulus-response. Skinner (1957) championed the behaviorist approach to language acquisition. He said that language was acquired through constant linguistic input and positive reinforcement. Chomsky (1959) disagreed with Skinner’s position pointing out that independent of intelligence or instruction, a child’s grammatical competence outperformed input. Instead, Chomsky put forth a theory of syntax acquisition in which language is acquired unconsciously by the brain. He called it Universal Grammar (UG).

An early focus in the study of linguistic universals was on typological features. An example would be that all languages have nouns or that there is a preference for preverbal negation (Dahl, 1979) among languages. Chomsky (1968, 2006) postulated that L1 language acquisition was governed by UG, which functions as abstract, universal linguistic principals that
constrain the creation of grammatical rules within all languages (Celce-Murcia & Larsen-Freeman, 1983; Ellis, 2008; Lenneberg, 1967; White, 1989, 2003).

For example, UG allows for determiners in languages, but does not proscribe them, as in Japanese, which has no articles. The nature of the linguistic input determines which rules are selected. An English speaker has only two determiners used to signify whether the speaker has a specific referent or not. The indefinite determiner also indicates that the noun is singular. English determiners do not indicate possession nor are they declined depending on their placement in a sentence. German determiners, on the other hand, communicate gender, case, number, and possession. Spanish allows for only masculine and feminine articles, which also indicate whether the noun is singular or plural. The question for second language acquisition (SLA) researchers was how long the LAD was available to L2 learners, assuming there was a Critical Period (Lenneberg, 1967) after which it was impossible for an L2 learner to achieve native speaker competence. At the time, it was felt this happened around puberty when the brain was thought to lateralize creating the longitudinal fissure.

Before the advent of neuroimaging, Chomsky (2006) proposed that the LAD was located in the brain. This organ was responsible for determining which rules are acquired within the constraints of UG. This seemed plausible at the time based on research into aphasics. In 1861, Broca presented his research locating the part of the brain considered to be responsible for processing vocabulary and syntax (Ellis, 2008, Sousa, 2011). People with damage to this area of the brain have difficulty creating and understanding complex grammatical sentences. In 1874, Wernicke isolated an area of the brain thought to control the meaning of language (Ellis, 2008, Sousa, 2011). People with damage to this part of the brain are able to speak, but are neither comprehensible to the listener nor can they comprehend what they hear.
Since the 1990s, neurolinguists have shed light on brain activity during L1 and L2 acquisition through the use of functional magnetic resonance imaging (MRI), positron emission tomography (PET), and magneto-encephalography (MEG). There are several conclusions drawn from the research. Since lateralization happens much earlier than puberty and may occur before birth (Ellis, 2008), it plays no part in limiting SLA.

The second set of conclusions concern UG. Language acquisition engages many different areas of the brain associated with language and general cognitive activity. For example, the hippocampus is important in the formation of new, declarative (episodic) memory and housing lexis. The basal ganglia are involved in motor control, learning, cognition, and procedural memory related to grammar. Early ELs have been shown to use additional brain activity when producing language and less when comprehending their L2 (Abutalebi, Cappa, & Perani, 2011; Binder, Westbury, McKernan, & Possing, 2005). However, proficient L2 speakers and L1 speakers are shown to access the same areas of the brain, leading Binder et al. (2005) to conclude that “when proficiency increases, L2 processing ‘converges’ on the neural representation of L1” (p. 205). The results of these studies show that there is neither one organ in the brain responsible for L1 acquisition nor are there separate systems responsible for L1 and L2 acquisition (Abutalebi et al., 2011; Binder, Desai, Graves, & Conant, 2009; Binder et al., 2005; Ellis, 2008; Perani & Abutalebi, 2005).

Generative grammarians postulate now that “UG provides a genetic blueprint, determining in advance what grammars can (and cannot) be like” (White, 2003, p. 2). So far, developmental biology has only provided evidence that the FOXP2 gene located on the 7th chromosome is involved in the speech disorder dyspraxia affecting the fine motor control of the tongue and lips, as well as linguistic and grammatical impairment involving the afore mentioned
basal ganglia and other parts of the brain involved in language processing (Lai, Gerrelli, Monaco, Fischer, & Copp, 2003). How UG is involved with SLA is still unclear. While the study of neurolinguistics and developmental biology add to our knowledge of language acquisition, it is still not definitive. Many researchers believe this is only part of the puzzle that is language acquisition and do not know to what extent other aspects such as cognitive development, socio-affective factors, and individual differences (ID) play in the acquisition of a second language (Ellis, 2008). Chomsky (1968, 2006) eloquently wrote:

The moral is not to abandon useful tools; rather it is, first, that one should maintain enough perspective to be able to detect the arrival of that inevitable day when the research that can be conducted with these tools is no longer important; and, second, that one should value ideas and insights that are to the point, though perhaps premature and vague and not productive of research at a particular stage of technique and understanding. (p. 19)

Without formal instruction, native speakers attain linguistic proficiency in their first language by the age of 3 (Sousa, 2011). A child’s L1 at age 6 is equally complex as an adult’s language; however, a child’s language is social and necessary to survival in the home and community. The difference in a child’s language and an adult’s lies primarily in the size of their vocabulary. It is generally accepted that words that are more concrete, or highly imageable, are easier to comprehend. Dual coding theory (Swaab, Baynes, & Knight, 2002) explains how there are two semantic systems located in two slightly overlapping areas of the brain: one verbal-based and the other image-based (Binder et al., 2005; Binder et al., 2009). Concrete words, such as “cat” draw upon both systems in making meaning. Abstract words such as “perseverance” use only the verbal-based code having less informational content from which to draw meaning. Therefore, concrete words are more easily acquired than abstract (Binder et al., 2005; Sousa, 2011). Because L1 and L2 speakers draw from the same linguistic system in the brain when processing language, it would stand to reason that concrete words are easier for L2 speakers to
acquire than abstract words (Ellis, 2008; Fillmore & Snow, 2002; Gersten et al., 2007; Johnson, 2009; Scarcella, 2003).

A register is a language used in a particular setting and acquired through some sustained experience or endeavor. No register is inherently better or worse than another register (MacSwan & Rolstad, 2009; Petrovic, 2012). The register associated with home, family, and friends, has many labels within the literature: basic, casual, conversational, every-day, face-to-face, informal, interpersonal, interactional, oral proficiency, and social, to name few. In this paper, it will be referred to as a social register.

A social register is initially acquired at home and normally used to communicate with family and friends about topics that are immediate, familiar, and of personal interest. Meaning is conveyed through facial expression, gestures, voice, pictures, and/or realia. The language of a social register is largely concrete, imageable, and easily comprehended. When communication is unclear, interlocutors engage in a negotiation of meaning, repairing misunderstandings and clarifying concepts. Researchers agree that a social register is more quickly acquired than an academic register, which is more abstract and relies on the language itself for clarification (Ellis, 2008; Fillmore & Snow, 2002; Gersten et al., 2007; Johnson, 2009; Scarcella, 2003).

An academic register is a social construction (Petrovic, 2012) acquired through instruction usually beginning when a child enters school. Chamot and O’Malley (1994) defined academic language as “the language used by teachers and students for the purpose of acquiring new knowledge and skills. Academic language has very specific purposes, including imparting new information, describing abstract ideas, and developing students’ conceptual understanding” (p. 40). Johnson (2009) tells us that there are two categories of language used in schools: content-specific and general academic. Content-specific language refers to vocabulary particular
to a subject. For example, democracy and geography are associated with social studies in the same way that addition and subtraction are uniquely linked to mathematics. General academic language is required across all content areas and is described in terms of cognitive actions, concepts, and text transitions, which are both complex and abstract (Gersten et al, 2007; Johnson, 2009; Marzano, 2004; Schleppegrell, 2001; Zwiers, 2008). Examples include analyzing, synthesizing, and evaluating, which are not easily visualized.

While WIDA mirrors Johnson’s (2009) distinction between content-specific and general academic language as one reason for justifying the clustering of grade levels, in its research agenda, WIDA (2012b) also states in its “guiding principles of language development,” that “students’ development of academic language and academic content knowledge are interrelated processes.” In addition, WIDA (2012b) defines an academic register as the following:

The oral and written text required to succeed in school that entails deep understanding and communication of the language of content within a classroom environment; revolves around meaningful application of specific criteria related to Linguistic Complexity at the discourse level, Language Forms and Conventions at the sentence level, and Vocabulary Usage at the word/phrase level within the particular context in which communication occurs. (p. 112)

Vocabulary Usage connotes “general, specific, and technical language” (WIDA, 2012). An overlap may be inferred between content-specific and general academic language.

Cummins’ (1981) distinguished between a social register being context-embedded and an academic register being context-reduced. While useful, the distinction between Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) has caused some misunderstandings among teachers. First, CALP is often interpreted as another degree of proficiency. Nevertheless, this distinction has little to do with proficiency in the linguistic sense (MacSwan, 2000; Petrovic, 2012). Furthermore, CALP has been defined as a “decontextualized language.” The problem here is that it is not decontextualized, as Gee (2002)
and Petrovic (2012) pointed out--Hyper-contextualized--as school language and even more so within specific content areas where different vocabulary and different vocabulary and different modes of expression are required (Gee, 2008) among teachers. Both the social and academic registers are contextualized (Gee, 2002; Petrovic, 2012). The critical difference between the two registers lays in the degree in which they employ concrete and abstract language not the domain in which the language is communicated. Furthermore, concrete language and abstract language draw upon different orthographic neighborhoods in the brain. Greater concreteness results in greater semantic connectivity because it activates both the image-based semantic system, located in the frontal lobe of the brain, and the verbal-based semantic system, located in the parietal and occipital lobes of the brain (Swaab et al., 2002). Abstract language is accessed through the verbal-based system only. This reduces the amount of informational context available, because of the number of words co-activated during target word recognition (Midgley, Holcomb, & Grainger, 2009) and the lack of imageability inherent in abstract language (De Bleser et al., 2003; Swaab et al., 2002).

Second language acquisition (SLA) is an additive process. It occurs when any language is acquired after a person’s L1. This can happen at any point in a person’s life (Ellis, 2008; Perani & Abutalebi, 2005). For children of immigrant parents in the United States, it most often happens when they enter school (Capp et al., 2008). While most English speakers begin acquiring an academic register when they start school, ELs must acquire both a social register and an academic register at the same time. Researchers concur that it takes less time to acquire a social register than an academic register, but that one does not necessarily have to precede the first. Both a social register and an academic register can be acquired simultaneously. It is

---

5 Gee (2002) and Petrovic (2013) referred to the academic register as hyper-contextualized since language is limited to one context, the verbal-based system.
generally accepted in the literature that it takes between two and five years to acquire a social
register and between four and seven years to acquire an academic register (Hakuta, 2002; Hakuta
et al., 2000; Thomas & Collier, 2002).

Thomas and Collier’s (2002) report analyzed 210,054 student records obtained over a 5-
year period (1996-2001) from across the United States in order to ascertain which type of
instructional support best serves ELs. The major findings support the use of developmental
bilingual programs (DBP) over English as a Second Language (ESL) programs. Students in
DBPs reached the 50th percentile on a nationally-normed, standardized reading test while those
in ESL programs reached a median 23rd percentile. Students whose parents waived services
fared the worst, scoring at the 12th percentile in 11th grade.6 These students were also the most
likely to drop out of school, the inference being that any language support is better than no
support at all.

Regardless of the type of L2 support services, Thomas and Collier (2002) found that ELs
needed between four and seven years of language support for the L1 and L2 achievement gap to
be closed. The greatest factor affecting a student’s success was the amount of formal schooling
the student had in his first language. ELs with four to six years of schooling in their L1 out-
performed students with only one to three years of schooling in their L1. Thomas and Collier
(2002) recommended that ELs be required to remain in some sort of support programs for four to
six years, because the highest achieving ELs in the study needed that much time to attain fluency
in the academic register enabling them to be successful throughout their school career.

While age of acquisition (AOA) is debated in the literature, much of the discussion
revolves around brain lateralization and Lenneberg’s (1967) Critical Period Hypothesis.

---

6 Flores et al. (2012) research on student trajectories in Texas also found that ELs whose parents
waived services were less successful that those participating in a language support program.
Neurolinguists have shown that brain lateralization happens before puberty and that the AOA does not affect which areas of the brain are used in L2 processing as much as the speaker’s level of proficiency (Wartenburger et al., 2003). While many assume that children are better second language learners, research shows that younger ELs lose their second language once exposure has ceased (Montrul, 2008). Older ELs are considered more efficient learners because of their access to learning strategies and previous experience learning a second language (Collier, 1988; Cummins, 1981; Ellis, 2008). The only aspect of SLA in which younger ELs consistently outperform older ELs is in the attainment of native-speaker pronunciation (Collier, 1988; Ellis, 2008). Ellis (2008) summarizes the research in this way: “There is no clear end point beyond which L2 learners will fail to achieve native-speaker proficiency. Rather there is a gradual decline in the ability to learn an L2 with age starting from early childhood” (p. 26).

“Individual differences (ID) have been found to be the most consistent predictors of L2 learning success” (Dörnyei, 2005, p. 6) in the research on SLA, perhaps because they have been most studied. IDs are the “dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree” (Dörnyei, 2005, p. 4). Personality, prior knowledge in her L1, intelligence, learning strategies, and working memory (Dörnyei, 2005; Ellis, 2008; Perani & Abutalebi, 2005) are some of the factors which may help or hinder an EL’s final L2 competency as well as how quickly that plateau is reached. Ellis (2008) identified the IDs most commonly found in the research: language aptitude, motivation, and anxiety.

Language aptitude refers to a person’s ability to learn an L2 independent of his level of current achievement. Language aptitude includes skills such as enhanced auditory ability, linguistic ability, and working memory. Dörnyei (2005) points out that intelligence, ability, and
aptitude are synonymous in the literature. However, Ellis (2008) treats language aptitude and intelligence separately, because intelligence is not specific to language aptitude and may obscure the unique propensities inherent in language learning. Many tests have been developed to identify people with this special talent or flair. However, consensus has not been reached as to an exact definition.

Motivation is defined by *The New World Encyclopedia* (2008) as the initiation, direction, intensity, and persistence of behavior. In SLA, research has shown a direct correlation between motivation and levels of success (Masgore & Gardner, 2003). Dörnyei tells us that “High motivation can make up for considerable deficiencies both in one’s language aptitude and learning conditions” (2005, p. 65). Gardener and Lambert (1972) identified two types of motivation related to SLA: instrumental and integrative. The first type of motivation suggests a need to accomplish a goal such as passing a test or getting a job; the second type refers to the EL’s desire to integrate into the L2 culture (Dörnyei 2005; Ellis, 2008).

Language anxiety is an affective factor that can influence the end-state of a learner’s SLA (Ellis, 2008). Anxiety can either have a facilitating affect or it can interfere with SLA. While low levels of anxiety may spur learners to try harder, too much anxiety is debilitating. Research tells us that language anxiety is often linked to ELs perceived lack of proficiency when compared to other learners and is most evident when an EL is speaking in a second language (Woodrow, 2006). Krashen (1982), borrowing from Dulay and Burt (1977), hypothesized that a learner’s Affective Filter responds to emotion. Like a wall, the Affective Filter rises or lowers depending on the student’s feelings. If a student feels safe, the Affective Filter would lower allowing him to receive input. Thus it is important for educators to create a safe learning environment free from adverse consequences in order to promote SLA.
While there is extensive research into the internal influence of IDs in language acquisition, language aptitude and motivation have been found to be the most influential in determining how fluent a person becomes in a second language. IDs are not static; their influence varies depending on the sociocultural and socio-environmental situation of the EL. Some of the external influences on EL success lay within their home, school, and individual educational opportunities (Baker, 2011; Flores et al., 2012; Hoyle, O’Dwyer, & Chang, 2011; Kieffer, 2008; Parker et al., 2009; Portes & Hao, 2004; Sanchez, Ehrlich, Midouhas, & O’Dwyer, 2009; Starratt, 2003; Thomas & Collier, 2002). Home level factors may include SES, poverty level, parents’ level of education, generational distance from the first immigrant, growing up in an intact family, and Length of Time (LOT) in-country. School factors may include the quality and training of the teachers, program models, opportunities to learn academic content, size, poverty level, racial composition, and location of school. Individual educational opportunities vary in terms of disability status, prior education, students’ initial level of English language proficiency when entering school, transience, interrupted formal schooling, background schemata, attendance, prior achievement in reading, length of time in a language support program, and gender. Without qualitative research to inform quantitative results, researchers fail to help stakeholders identify ways in which they can increase student success.

English Language Assessment

When President Johnson signed the first ESEA in 1964, he began the work that would shape not only how states would provide services for ELs, but also the assessments related to those services. The same year, the National Assessment of Educational Progress (NAEP), also known as the Nation’s Report Card, was formed (Bunch, 2011; Crawford, 1999). This
organization would play a pivotal role in educational assessment reform and in framing the
test language of EL testing in NCLB (Bunch, 2011).

The assessment requirements of NCLB are the culmination of reforms implemented by
NAEP since its inception in 1964 (Bunch, 2011). Section 3121 of NCLB (U.S. Department of
Education, 2004) states,

(d) EVALUATION MEASURES. – A State shall approve evaluation measures for use
under subsection (c) that are designed to assess – (1) the progress of children in attaining
English proficiency, including a child’s level of comprehension, speaking, listening,
reading, and writing skills in English; (2) student attainment of challenging State student
academic achievement standards on assessments described in section 1111(b)(3); and (3)
progress in meeting the annual measurable achievement objectives described in section
3122. (P.L. 107–110 §3121)

In addition, evaluation of an EL’s English competence must be tied to research-based,
defensible, performance standards, each with clearly defined, multiple levels of abilities (Bunch,
2011; Gottlieb, 2003).

NCLB provided grants through OELA to states and agencies to design English language
assessment tests (P.L. 1007-110 §3122). OELA is required to collect and report the data
generated by these assessments and conduct program evaluations. The purpose of EL testing is
to provide data for the biennial evaluation required by Title III. It also provides data to states
for reporting Annual Measurable Objectives (AMOs) necessary for demonstrating AYP. AMOs
state the percentage of students in each subgroup required to meet the specific goals of AYP.
AMO percentages incrementally increase annually with the goal of 100% of students passing
state standards-based assessments by 2014 (GADOE, 2010). The reasoning for such high

---

7 Title I requires states to assess student content knowledge (U.S. Department of Education,
2004). ELs are required to master the same content as non-ELs. Title II requires states to assess
ELs English language proficiency (U.S. Department of Education, 2004). While these two tests
are positively related, a confounding factor may be that the ACCESS test is not tied to single
grade level content, but to a cluster of grade levels (Marion, 200WIDA, 2013).
expectations in such a short timeframe was that as overall student performance improved, it would be easier for schools to meet AMOs (P.L. 1007-110 §3122).

According to NCLB, by 2014, 100% of ELs are expected to score at the proficient level in reading and mathematics: the same goal as fluent English speakers. NAEP interprets proficiency in terms of the validity of a state’s assessment (Stoneberg, 2007).

The State’s academic achievement standards fully reflect its academic content standards for each required grade and describe what content-based expectations each achievement level represents. The ‘proficient’ achievement level represents attainment of grade-level expectations for that academic content area. (U.S. Department of Education, 2004)

Georgia defines proficiency in terms of scores on state standard assessments. For the CRCT for Grades 1-8, a proficient score is between 800 and 849. A score between 850 and 950 exceeds expectations. High school students must pass the Georgia High School Graduation Test in ELA with a score of 200 or better. EOCTs have a list of specific performance level indicators. Students performing at the proficient level demonstrate competent, clear, or adequate understanding and knowledge of a list of skills associated with each content area. The highest score possible on an EOCT is 600 with a passing score of 400 (GADOE, 2012b).

As the CEP (2010) points out, ELs who attain that level of competence are exited from English language support programs, while non-English speakers are continually entering the program. For an EL subgroup, 100% proficiency is an unrealistic target. AMOs for the adequate yearly progress of English acquisition is a more attainable goal.

In 2012, the Obama administration waived some of the NCLB requirements for Georgia and nine other states with the Bringing Flexibility and Focus to Education Law (FFEL) (2012). All states may apply for waivers that will be in effect until next ESEA reauthorization. At that time, the conditions of the waiver may be superseded by the new law.
What is not changing under FFEL is the spotlight on subgroup performance. Data will still be reported separately for the ELs and other groups at risk. Priority Schools, the bottom 5% of lowest-performing schools, and Focus Schools, those in the next 10% identified by low graduation rates and subgroup performance, will be required to implement research-based instruction tailored to that school’s needs in closing the large achievement gap. FFEL is cancelling the 2014 deadline for 100% proficiency in reading and mathematics for all students. This is replaced by flexible goals determined by individual states, but requiring progress to be shown toward college and career-ready content standards. Accountability is expanded from the single criterion-referenced test to multiple forms of assessment such as grades and portfolios.

FFEL does not change Section 3121 of NCBL (U.S. Department of Education, 2004). ELs are still required to take an annual assessment of their language ability in all four domains of language as proscribed by Title III. In order to design and develop such tests, OELA made a request for proposals in 2002 (Bunch, 2011). Nine proposals were submitted and four consortia were chosen to receive Enhanced Assessment Instrument grants through Title VI, Section 6112 of ESEA (1965): Mountain West, State Collaborative on Assessment & Student Standards (SCASS), Accountability Works, and WIDA. Originally named for the consortium states Wisconsin, Delaware, and Arkansas, WIDA now stands for World-Class Instructional Design and Assessment and is a consortium of 31 states including Georgia. Georgia was initially a

---

8 WIDA’s lead state is Wisconsin and the test developed is called the Assessing for Comprehension and Communication State to State for English Language Learners (ACCESS). The Pennsylvania Enhanced Assessment Group developed the Comprehensive English Language Learning Assessment (CELLA), which is currently used only in Florida. The State Collaborative on Assessment & Student Standards (SCASS) Consortium designed the English Language Development Assessment (ELDA). Finally, the Mountain West Assessment Consortium’s (MWAC) assessment is no longer being used.
member of the SCASS consortium. Membership in the consortia is fluid with the Mountain West consortia having dissolved in 2006 (Northwest Regional Education Laboratory, 2006).

According to Bunch (2011), the tests developed by the four consortia had many characteristics in common. All consortia provide for the annual assessment and collection of data for state and federal reporting purposes. Development of assessments was linked to the content standards of member states and English Language Proficiency Standards. All include assessment in the language domains of reading, writing, listening, and speaking. Test items include multiple-choice, short-answer, extended response, and oral responses scored by standardized rubrics. All tests are characterized by technical rigor relating to item development. The consortia engage in pilot testing, and provide multiple forms of the assessment. The reliability scores for the four assessments ranged from Alpha .73 - .97 across all levels and test forms indicating a high inter-correlation between test items.

One criticism of standardized testing of ELs raises the concern that for all limited English proficient students, every test is an English language proficiency test regardless of the purpose or content being assessed (Abedi, 2002, 2008; Pitoniak et al., 2009). While a positive correlation between an English proficiency test and content test may lend some validity to their use, educators must include other factors, such as disability status, gender, and LOT in an ESOL program when making high-stakes decisions that may impact an EL’s educational outcomes (Pitoniak et al., 2009).

---

9 Linkage requires that at least one content standard from one of the grade levels in the cluster be represented in the test questions (U.S. Department of Education, 2003, 2004).
10 WIDA’s initial Alpha: .82-.97, the highest score among four consortia.
ESOL Program in Georgia

The Title III Biennial Report to Congress (2008) showed that most EL services focused on English literacy. Georgia is an English-only state, meaning that the language of instruction in public schools is English. Bilingual or dual language programs are not approved program delivery models. Early-exit, transitional programs, in which bilingual instruction is approved for 1 year, are allowed, even though research has shown that late-exit programs are more effective (Ramirez et al., 1991; Rolstad, Mahoney, & Glass, 2005; Thomas & Collier, 2002).

Most ELs in Georgia receive English Language Development (ELD) through a choice of state-approved ESOL service delivery models. In a survey of program types, the GADOE (2012a) found that most ELs are served through pull-out or push-in delivery models. In the pull-out delivery model, students are pulled out of their regular class and brought to the ESOL teacher’s room for instruction. ESOL teachers that push-in to classrooms are neither co-teaching nor are they teaching the same lesson as the regular teacher in another part of the room; they exclusively help ELs in small groups or through individual conferencing during regular instruction to meet English Language Proficiency Standards (ELPS) and state content standards. In addition to ELD classes, middle schools and high schools may provide EL services through Structured English Immersion (SEI) content classes. SEI classes are similar to sheltered instruction in that all students in the class are ELs (GADOE, 2012a).

Whether ELs are being pulled out of classes, ESOL teachers are pushing into classes, or ELs are attending SEI classes, the purpose of ESOL services in Georgia is to prepare students to be successful in the regular classroom in English (GADOE, 2012a). The school district included in the sample provides training in the Shelter Instruction Observation Protocol (SIOP), a method
of content delivery that incorporates best practices in the field of ESOL instruction, to both
ESOL and general education teachers at all grade levels (Echevarria, Vogt, & Short, 2008).

Summary

A review of the literature indicates that standards-based assessments in earlier grades can
be used as a predictor of student success on high school exiting exams. Two of the studies
discussed in this chapter disaggregated data for ethnicity. Greeson (2009) found that it was not a
significant variable. Davis (2004) found that there was a slight predictive value for ethnicity and
final grades in a course, but which ethnicity was not revealed. Neither study disaggregated data
for ELs who had exited an ESOL program. No study used an English proficiency exam to
predict success on high school graduations tests. However, several studies indicated that
ACCESS scores could be a predictor for standards-based assessment scores in math, reading, and
English language arts for fifth and eighth graders (Parker et al., 2009), first through fifth graders
(Baker, 2011), and for enrollment in ninth grade Literature and Composition and in American
Literature courses in Georgia (Walker, 2008). The relationship between ACCESS scores and
Biology EOCTs has not been studied. However, Greeson (2009) in Georgia and Hobbs (2005)
in Tennessee found that the science subtest in eighth grade was able to predict success on every
content area of the high school graduation test.

As required by BEA and NCLB, the focus for EL instruction has become the acquisition
of an academic register. This has become manifest in ELPSs and in systematic, annual
assessment of ELs’ proficiency. The ACCESS test has become the most widely used English
language assessment in the United States. While the ACCESS was developed to meet the testing
requirements promulgated by NCLB, OLEA, and NAEP, the question still remains for some
educators as to whether the ACCESS accurately predicts student success in the regular education classes. In Walker’s (2008) study comparing ACCESS scores to ninth grade Literature and Composition EOCT scores, ELs who were deemed proficient by the ACCESS test in Georgia, scored an average of 420 on both ninth grade Literature and Composition and American Literature EOCTs. When the range of pass scores for EOCTs is from 400-600, ELs are barely making the grade. EL scores on Biology EOCTs are even more grim; from the 2007-2008 to the 2010-2011 school years, an average of 64% of LEP students in the Georgia failed the Biology EOCT (GADOE, 2012c).11

In 2005, ELs comprised 10% of the total national student population (NETC, 2005). Traditionally, ELs score lower than native English speakers on academic assessments of content knowledge in English (Abedi & Dietel, 2004). This is a significant group of students with a particular standards-based assessment that evaluates their academic language proficiency in English. Comparing academic language proficiency scores to standards-based content assessment scores may give insight into the particular abilities and challenges facing ELs so that Georgia educators might help them be successful on ECOTs.

11 An average of 64% of ELs failed the Biology EOCT in the district included in this study (GADOE, 2012c).
CHAPTER 3

METHODOLOGY

This study examined the relationship between the ACCESS test and the ninth grade Biology EOCT in Georgia. The purpose of this study was to determine whether ACCESS scores can be used to predict EL success on the ninth grade Biology EOCT. Furthermore, this study sought to establish a relationship between the two tests in order to predict what overall subscale score an EL might have on the ACCESS while also passing the Biology EOCT. This chapter presents the research questions and study design. It describes the setting, student sample, the assessments being measured, the procedures for data analysis, and ethical considerations in undertaking this research.

Research Questions

This study was guided by two research questions:

1. How does performance on the ACCESS test of English language competence predict EL performance on the Biology EOCT in ninth grade after controlling for student characteristics?

2. Which language domain is the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing?
Research Design

This study was originally designed to be similar to the 2008 GADOE study (Walker, 2008). Following Walker’s model, linear regression analyses were employed to determine the relationship between an EL’s scores on the ACCESS test and Biology EOCT. The primary analysis focused on the relationship between the overall subscale score and the Biology EOCT score. Secondary analyses explored the relationship between the EOCT score and the ACCESS scores in the four language domains of speaking, listening, reading, and writing. It differs in the inclusion of a step-wise regression analysis to determine the influence of the covariates of gender and LOT on the results. The regression equation is also particular to the data provided by the analyses.

Research Setting and Subjects

In north Georgia lays the heart of the world’s carpet industry. Of the world’s carpets, 70% are produced here by such companies as Shaw, Mohawk, and Beaulieu of America (Georgia Humanities Council and the University of Georgia Press, 2011). In the 1990s, the production of carpet grew so rapidly that major manufacturers picked up busloads of Mexican workers at the border between Mexico and the United States and brought them directly to this area (Valencia & Tripp, 2012). The influx of immigrant workers took the community by surprise. In 1990, the U.S. Census reported a total population for the county of 15,934 (U.S. Census Bureau, 1995). In 2010, the population had grown to 102,599; 32,471 of whom are Latin@ (U.S. Census Bureau, 2012). By 2010, the number of Latinas alone had grown to twice as many as the total citizenry in 1990.
To illustrate the impact of the population growth on the local school systems, when the district included in the study began its ESOL program in 1990, there were only two teachers serving 55 students in Grades 6 through 12 (Pressley, 2012). In spring 2011, there were 25 teachers serving 1,120 students in Grades Kindergarten through 12. ELs make up 16.7% of the district’s student population (GADOE, 2012c). This is a significant portion of the school population identified by NCLB as having specific learning needs (U.S. Department of Education, 2004).

The subjects in this study are or were ninth grade ELs who attended high school during the school years 2006-2007 through 2011-2012. A purposive, random sample of student scores was selected from approximately 500 students using two criteria: 1) the scores must be from ninth grade, and 2) the scores must be available for both the ACCESS test and the Biology EOCTs for school years 2006-2007 through 2011-2012.

![Figure 1. Number of ELs and overall subscale scores from the springs of 2007-2012.](image-url)
A purposive sampling is also called a judgment sampling (Gay et al., 2006). The sample in this study is purposive, because the researcher has chosen a school district with a population that has a high number of EL students in ninth grade. The sample is random, because it contains a certain number of members chosen at random from the total accessible population as “the best single way to obtain a representative sample” (Gay et al., 2006, p. 101).

Assessments

This study compared two standardized assessments that EL students in Georgia must take in ninth grade. The first was the ACCESS test and the second was the EOCT in Biology. The ACCESS test is used by the GADOE to annually measure the progress of an EL’s language acquisition. The overall subscale score is also used to determine whether an EL should continue in ESOL services or exit the program. The EOCT in Biology is used by the GADOE to determine mastery of content material. Beginning in the 2012-2013 school year, EOCTs replaced the Georgia high school graduation test as a means of determining whether a student may graduate from high school or not.

ACCESS for ELLs

The ACCESS test was developed by the WIDA consortium through a grant from the OLEA for the development of English competency tests that meet the NCLB (U.S. Department of Education, 2004) requirements. WIDA provides a screening test called the WIDA-ACCESS Placement Test (W-APT) (WIDA, 2011). Once students have been designated as EL, they must submit annually to the ACCESS test of English competence. The score on the ACCESS is used by educators to determine whether an EL will exit ESOL services or remain in the ESOL program (GADOE, 2012a). It is also used to measure student progress from one year to the next.
The ACCESS test is divided into grade level clusters (WIDA, 2012). Kindergarten, Grades 1-2, Grades 3-5, Grades 6-8, and Grades 9-12. There are six proficiency levels identified by WIDA: Level 1 = Entering, Level 2 = Beginning, Level 3 = Developing, Level 4 = Expanding, Level 5 = Bridging, and Level 6 = Reaching. Scores range from 1.0 to 6.0 and correlate to the proficiency levels: scores ranging from 1.0 to 1.9 equal Level 1 -- Entering, scores ranging from 2.0 to 2.9 equal Level 2 -- Beginning, etc. Scores may not be rounded up or down.

![Figure 2. The WIDA English language proficiency continuum. Reprinted from World-Class Instructional Design & Assessment (WIDA) Consortium, Madison, WI. Copyright 2011 by the Board of Regents of the University of Wisconsin System.](image)

There are three ability tiers that may be administered to ELs: Tier A, Tier B, and Tier C. According to directives from the GADOE (2012a), Tier A is designed for ELs who have been in the American school system for 1 year or less. The proficiency levels tested are 1 through 3. Tier 2 is reserved for second year students who are deemed by their ESOL teacher not to have the ability to exit the ESOL program at that time. Tier 3 is used for all other ELs. Only scores
obtained on the Tier C test are used for exiting purposes, because it is the only level that encompasses Level 6, which is considered the equivalent to a fluent English speaker.

Per NCLB requirements, the ACCESS test is concerned with an EL’s competency in the academic register. Therefore, the test is based on the five ELPS developed by WIDA (2012) for social and instructional languages, mathematics, reading, science, and social studies. The ESLP are worded in generalities:

- *English Language Proficiency Standard I*: English language learners communicate for **Social** and **Instructional Purposes** within the school setting.¹²
- *English Language Proficiency Standard II*: English language learners communicate information ideas, and concepts necessary for academic success in the content area of **Language Arts**.
- *English Language Proficiency Standard III*: English language learners communicate information ideas, and concepts necessary for academic success in the content area of **Mathematics**.
- *English Language Proficiency Standard IV*: English language learners communicate information ideas, and concepts necessary for academic success in the content area of **Science**.
- *English Language Proficiency Standard V*: English language learners communicate information ideas, and concepts necessary for academic success in the content area of **Social Studies**. (WIDA, 2012a)

ELPS are further defined by model performance indicators (MPI) based on whether the assessment is formative or summative, the language domain, and level of proficiency. Each MPI has three parts: 1) a language function, 2) a topic, and 3) examples of what ELs in WIDA’s proficiency levels 1 through 5 can do. An example of a summative MPI from WIDA’s (2012) website for the Grades 9-12 cluster contains the following information: “Example topic: Food chains--Life cycles; Domain: Speaking; Level 1 = Identify components of chains or cycles from diagrams or graphic organizers; Level 2 = Give examples of components or functions of chains or cycles from diagrams or graphic organizers; Level 3 = Describe sequence within chains or cycles from diagrams or graphic organizers; Level 4 = Explain the importance or impact of the

---

¹² Words emphasized by author.
iterative nature of food chains or life cycles; Level 5 = Discuss how food chains or life cycles within ecosystems are interdependent.” Language functions are listed by WIDA’s Madison Academic Language Working Group as one of the five components of academic language. The other components are vocabulary, grammar, text structure, and genre (WIDA, 2012). The language functions included in this example are to identify, give examples, describe, explain, and discuss.

WIDA (2012) releases test items so that ELs may practice using the testing format. Examples for the ELPS of Science cluster for Grades 9-12 do not contain items from the reading and writing domains. The ACCESS listening and speaking items released by WIDA concern the solar system, which is not included in the ninth grade Biology content standards (GADOE, 2012). However, interpreting data from tables, graphs, and charts is listed as an ongoing skill Biology students should possess.

The ACCESS tests all four language domains, but does not weigh them equally when determining an overall score. Reading and writing are elements of the academic register and thus account for 35% each of the overall subscale score on the ACCESS. Listening and speaking account for 15% each. Scores are adjusted for grade level by establishing different thresholds for each grade level in a grade-level cluster taking the same test. Scores are reported in two ways: 1) on a scale of 1 to 6, which correlates to the WIDA proficiency levels, and 2) as scaled scores adjusted for grade levels. The former is illustrated with a bar graph and distributed to parents. The latter is distributed to teachers. Raw scores are included on the Teacher Report by content area and language domain; however, the number of items linked to content areas is so few that it
would not be viable to explore a correlation between content areas on the ACCESS and content areas on state assessments such as EOCTs.\textsuperscript{13} (Parker, et al., 2009).

\textit{End-of-Course Tests}

Georgia no longer requires high school students to take a graduation test (GADOE, 2012c).\textsuperscript{14} From school year 2011-2012, EOCTs replaced the GHSGT as a component of its College and Career Readiness Performance Index (CCRPI). EOCTs will serve four other functions as well. First, it can be used as a diagnostic tool for students. Second, the tests can “provide data to evaluate the effectiveness of classroom instruction at the school, system, and state levels” (GADOE, 2012c). Third, they will provide 20% of a student’s final grade in the course that is being assessed. Finally, they will serve as the state’s high school accountability assessment per NCLB requirements (U.S. Department of Education, 2004).

There are 10 EOCTs that students must take during their high school careers in order to graduate (GADOE; 2012c). They are Literature and Composition, American Literature and Composition, Algebra, Geometry, Mathematics I, Mathematics II, Biology, United States History, Physical Science, and Economics/Business/Free Enterprise. EOCTs required for ninth graders include Literature and Composition, Mathematics I, Biology, and Economics/

\textsuperscript{13} The Teacher Report for 2011, Grade Level Cluster 9-12, Tier C, shows that the WIDA English Language Proficiency Standard for the Language of Science had a total of 9 test items within the listening and reading domains. Science and Mathematics English Language Proficiency Standards were combined for scoring in the Speaking task with a total of 5 items. The number of items was not reported for the writing tasks. Instead, a raw score and total possible points were reported for the combined standards of Mathematics and Science in the area of Linguistic Complexity, Vocabulary Usage, and Language Control.

\textsuperscript{14} According to Khote (2012), until the school year 2011-2012, the graduation test took precedence over EOCTs. Up until then, teachers did not prepare their students well for the EOCT. Therefore, he believes the EOCT tests prior to 2011-2012 school year are not an accurate indicator of students’ abilities and knowledge.
Business/Free Enterprise. Specific Performance Level Descriptors are provided by the GADOE for each EOCT, indicating whether a student exceeds the standard, meets the standard, or does not meet the standard. These are outlined according to the topics covered in each course. For example, the Biology ECOT encompasses the five topics of cells, organisms, genetics, ecology, and evolution. Lists of required content and accompanying skills are included. Scores are reported in three ways: performance levels, scaled scores, and a conversion score. Scaled scores range from 200 to 600, with 400 equaling the pass or “meets the standard” criteria. Georgia has chosen 70%\textsuperscript{15} to be the grade below which students fail. Therefore, a scaled score below 400 equals conversion score of 70% (GADOE, 2011). The conversion score is used in determining a student’s final grade in the course (GADOE, 2012d). One retake opportunity is allowed for students who do not meet the standard. Failing this, they must reenroll and complete the course again before attempting the EOCT a third time.

Procedures for Data Analysis

Beginning in the fall of 2012, student data, including ACCESS scores, Biology EOCT, gender, home language, and years in country, were collected on CD and hardcopy from the school district. The data were deidentified so that anonymity was preserved. Using the Statistical Package for the Social Sciences (SPSS), the data were uploaded and analyses occurred as outlined below.

\textsuperscript{15} Per GA state board rule 160-4-2-13 statewide passing score (GADOE, 2011). No explanation was provided as to why 70% was selected as the passing score. However, rubrics have been established, based on content standards, indicating what criteria are necessary to pass EOCTs. This limits the influence of social and political issues which are imbued with the conflicting values and goals of the individuals making the decisions (Kane, 1994)
First, histograms were created for both the scaled scores (SS) on the Biology EOCT and for the overall proficiency levels on the ACCESS. A bell curve was evident, so a Pearson’s Product Moment Correlation (r-test) was conducted. Because a high positive score was found on the r-test, several secondary analyses followed.

Second, a series of linear regression analyses were performed in order to determine which subscales of the ACCESS test were better predictors of success on the Biology EOCT. Covarients initially included in a step-wise regression analysis were gender, and length of time in an ESOL program.

The variables in this analysis were coded for SPSS in this way. In terms of gender, females were coded as 1, males were coded as 2. For ethnicity, there were two groups represented in the sample. Asians were coded as 1 and Latin@s were coded as 2. ELs without disabilities were coded as 1 and ELs with a disability status were coded as two. Length of time in and ESOL program (LOT) was coded by the number of years an EL had been in an ESOL program at the time of testing. ELs with less than one year in an ESOL program were coded as zero. ELs with between one and two years in an ESOL program were coded as one. Because of the small number of ELs in the sample having been in the program for six or more years, these students were combined into one group and coded as 6.

Ethical Considerations

Several ethical considerations were taken concerning this study. There was no direct student involvement in this study due to the use of data collected as a normal part of the educational process. In order to maintain confidentiality, neither names nor institutional identification numbers were used to identify any student individually. This was a study to
determine whether one set of test scores can predict success on another test; therefore, cause and
effect inferences may not be drawn from the results of the analyses. In addition, the data were
not disaggregated by school year curtailing any conclusions that might have been drawn about
the ESOL programs used by the high school or its EL population at any specific time.
CHAPTER 4

RESULTS

Group Demographics

Table 2 shows the EL group demographics in terms of disability status, gender, and ethnicity. The overwhelming majority of ELs in this sample (99%, n = 163) did not have a disability. ELs with disabilities (1%, n = 1), while included in the total sample, did not constitute a large enough sample of students to warrant separate analyses. Latinas (94.5%, n = 154) accounted for a larger percentage of the sample than Asians (5.5%, n = 9). Thus ethnicity was not included as a separate covariate. However, gender was fairly evenly distributed between females (43.9%, n = 72) and males (56.1%, n = 92) and was included as a covariate in the following analyses.

Table 2

Demographic Frequencies of ELs

<table>
<thead>
<tr>
<th>Student-level characteristics</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1</td>
<td>1.0%</td>
</tr>
<tr>
<td>No</td>
<td>163</td>
<td>99.0%</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin@</td>
<td>155</td>
<td>94.5%</td>
</tr>
<tr>
<td>Asian</td>
<td>9</td>
<td>5.5%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>92</td>
<td>56.1%</td>
</tr>
<tr>
<td>Female</td>
<td>72</td>
<td>43.9%</td>
</tr>
</tbody>
</table>

The majority of ELs (76%, n = 124) failed the Biology EOCT. There were gender differences in the pass/fail rate, $\chi^2(1) = 7.675, p = .006$. A higher proportion of males passed the
Biology EOCT (75%, n = 40) as compared to females (25%, n = 40) who passed the Biology EOCT.

All of the ELs in the sample (N = 164) were in ninth grade. Table 3 shows the length of time (LOT) the students had been in an ESOL program at the time of testing. The largest group of ELs (32.9%, n = 54,) were in their second year in an ESOL program. The average number of years spent in an ESOL program was 2.42 (M = 2.42, SD = 1.66) years. The majority of students in the study (79.9%, n = 131) had been in an ESOL program for 4 years or less at the time of testing. One student had been in an ESOL program for 10 years; this student was not the EL designated as having a disability.

Table 3

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of students</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>13</td>
<td>7.9%</td>
</tr>
<tr>
<td>1</td>
<td>35</td>
<td>21.3%</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>32.9%</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>17.7%</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>9.1%</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>6.1%</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>2.4%</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1.2%</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.6%</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.6%</td>
</tr>
<tr>
<td>Total</td>
<td>164</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 3 lists the mean test scores for ELs depending on the number of years spent in an ESOL program. Table 4 collapses years 5 through 10 into one group. Mean group differences were tested using an ANOVA, $F (5, 158) = 4.955, p < .001.$ The differences were significant. Students who were in the ESOL program longer, scored lower than those who were in the
program fewer years. ELs that were in the program for less than 1 year, scored the highest on
the Biology EOCT. ELs that were in the program between 1 and 2 years had the second best
scores. Those who were in the ESOL program between 2 and 5 years scored the lowest on the
Biology EOCT.

Table 4

Distribution of Biology EOCT Scaled Scores by LOT

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of students</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>13</td>
<td>395.92</td>
<td>26.091</td>
</tr>
<tr>
<td>1</td>
<td>35</td>
<td>389.09</td>
<td>21.909</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>383.41</td>
<td>21.520</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>383.34</td>
<td>19.993</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>383.80</td>
<td>19.972</td>
</tr>
<tr>
<td>5+</td>
<td>18</td>
<td>385.89</td>
<td>18.598</td>
</tr>
<tr>
<td>Total</td>
<td>164</td>
<td>385.91</td>
<td>21.299</td>
</tr>
</tbody>
</table>

Data Analyses

Research Question 1: How does performance on the ACCESS test of English language
competence predict EL performance on the Biology EOCT in ninth grade after controlling for
student characteristics?

A significant correlation was found between the ACCESS scores and the Biology EOCT
scores, \( r = .414, n = 164, p = .001 \). This implies that there is a moderately strong relationship
between the test scores. They have a positive relationship, meaning that if ACCESS scores are
higher, then the Biology EOCT scores will also be higher.
Additionally, in a linear regression analysis, the ACCESS overall subscale had a moderately positive relationship with the dependent variable, the Biology EOCT, $B = 9.707, t(163) = 5.787, p = .001$. Therefore, the ACCESS overall subscale is a significant predictor of success on the Biology EOCT. For every unit increase in the ACCESS test, the Biology EOCT score increased 9.707 points.

A stepwise multiple regression analysis was conducted to evaluate whether the ACCESS overall subscale, gender, and LOT were useful in predicting Biology EOCT scaled scores. At step 1 of the analysis, the ACCESS subscale score was entered into the regression equation. At step 2, gender was entered into the regression equation. Finally, at step 3, LOT was entered into the regression equation. The multiple correlation coefficient was $R^2 = .248$, indicating approximately 24.8% of the variance of the Biology EOCT scaled score could be accounted for by the ACCESS overall subscale score, gender, and LOT. Thus the regression equation for predicting a passing score on the Biology EOCT is as follows:

$$EOCT = \text{Constant} + \text{ACCESS Subscale} + \text{Gender} + \text{LOT}$$

The regression equation for a male EL who scores 400 on the Biology EOCT looks like this:

$$EOCT = 336.295 + 10.273 + 8.681 + (-2.602 \times \text{LOT})$$

The Biology EOCT constant is 336.295, or average score attained by ELs on the EOCT when the covariates of gender and LOT are zero. For every unit increase in the ACCESS test, the Biology EOCT score increased $B = 10.273$ points. Males, $b = 8.681, t(160) = 2.931, p <$
have significantly higher Biology EOCT scores than females, adding 8.681 points to the constant.\textsuperscript{16} Biology EOCT scores decrease the longer an EL spends in an ESOL program, $B = -2.602$, $t(160) = -2.892$, $p = .004$. For each year spent in an ESOL program, the student’s score decreases by -2.602 points.

Table 6

\textit{Linear Regression Analysis of EOCT and ACCESS Overall, Gender, and LOT}

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>$b$</th>
<th>$t(163)$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOCT</td>
<td>336.852</td>
<td>7.830</td>
<td>43.018</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>ACCESS</td>
<td>10.273</td>
<td>1.638</td>
<td>.441</td>
<td>6.273</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>8.681</td>
<td>2.962</td>
<td>.203</td>
<td>2.931</td>
<td>.004</td>
</tr>
<tr>
<td>Length of time</td>
<td>-2.602</td>
<td>.899</td>
<td>-.203</td>
<td>-2.892</td>
<td>.004</td>
</tr>
</tbody>
</table>

A Latino’s best chance for passing the Biology EOCT falls within his first 2 years in an ESOL program. Few EL students ($n = 40$) of either gender passed the Biology EOCT with an ACCESS Overall proficiency level less than 4.5 on a scale of 1.0 to 6.0. The mean scale score attained by the ELs ($n = 40$) who passed the Biology EOCT was 416. The passing score on the Biology EOCT is 400; the highest score possible is 600.

Research Question 2: Which language domain is the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing?

In order to better predict success on the Biology EOCT, a step-wise linear regression analysis was performed on the sample, $N = 164$. The dependent variable was the Biology EOCT scaled scores. Included stepwise in the analysis were the ACCESS subscale, gender, and LOT.

\textsuperscript{16} Being female adds no points to the regression equations. The regression equation for females achieving a score of 400 on the Biology EOCT test is:

$$EOCT = 336.295 + 10.273 + (-2.602 \times LOT)$$
The linear regression analysis shows that each domain has significance in predicting success on the Biology EOCT. Increases in all the subscales lead to an increase in the Biology EOCT test. Every unit increase in writing, $b = 9.782, t (163) = 4.557, p < .001$, increases the EOCT score the most by 9.782 points. Every unit increase in reading, $b = 6.930, t (163) = 5.803, p < .001$, increases the Biology ECOT test by 6.930 points. Every unit increase in listening, $b = 6.456, t (163) = 3.948, p < .001$, increases the ECOT score by 6.456 points. Every unit increase in speaking, $b = 2.589, t (163) = 2.555, p = .012$, increases the ECOT score the least by 2.589 points.

<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>B</th>
<th>SE</th>
<th>$b$</th>
<th>$t(163)$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td>349.105</td>
<td>6.456</td>
<td>1.636</td>
<td>.292</td>
<td>3.948</td>
<td>.000</td>
</tr>
<tr>
<td>Speaking</td>
<td>363.702</td>
<td>2.589</td>
<td>1.013</td>
<td>.197</td>
<td>2.555</td>
<td>.12</td>
</tr>
<tr>
<td>Reading</td>
<td>356.293</td>
<td>6.930</td>
<td>1.194</td>
<td>.407</td>
<td>5.803</td>
<td>.000</td>
</tr>
<tr>
<td>Writing</td>
<td>337.890</td>
<td>9.782</td>
<td>2.147</td>
<td>.335</td>
<td>4.557</td>
<td>.000</td>
</tr>
</tbody>
</table>

The ACCESS reports scores for the language domains of listening, speaking, reading, and writing as well the combined subscales of Comprehension (listening and reading), oral (speaking and listening), literacy (reading and writing), and overall, which is a combination of all four language domains. Several analyses were conducted to see if the combined subscales were in fact better predictors of success on the Biology EOCT than the individual domains or the overall subscale. Table 8 shows the results of a linear regression analysis of these subscales and the Biology EOCT. All of the combined subscales were significant; however, the literacy subscale, $b = 10.448, t (163) = 6.305, p < .001$, was the most predictive of the three, as changes
in the literacy score yield greater increases in the Biology EOCT scores. However, as all of the subscales increase, the Biology ECOT scores also increase. This indicates that combined subscale scores are also useful in predicting EL success on the Biology EOCT. However, for every unit increase in literacy, \( b = 10.448, t(163) = 6.305, p < .001 \), the Biology EOCT increases by 10.448 points, the greatest increase of all the subscale scores. Therefore, the literacy subscale may be a better predictor of EL success than the overall subscale \( (B = 9.707, t(163) = 5.787, p < .001) \) or the writing subscale \( (b = 9.782, t(163) = 4.557, p < .001) \).

Table 8

*Linear Regression Analysis of EOCT and Combined ACCESS Subscales*

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Constant</th>
<th>B</th>
<th>SE</th>
<th>( b )</th>
<th>( t(163) )</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension</td>
<td>342.474</td>
<td>8.747</td>
<td>1.438</td>
<td>.425</td>
<td>6.081</td>
<td>.000</td>
</tr>
<tr>
<td>Oral</td>
<td>353.418</td>
<td>5.370</td>
<td>1.381</td>
<td>.295</td>
<td>3.889</td>
<td>.000</td>
</tr>
<tr>
<td>Literacy</td>
<td>335.914</td>
<td>10.448</td>
<td>1.657</td>
<td>.437</td>
<td>6.305</td>
<td>.000</td>
</tr>
</tbody>
</table>

As reported in Table 9, the Biology EOCT correlates positively with all of the subscales of the ACCESS test. Higher scores on each of these subscales result in higher scores on the Biology ECOT. However, in a correlation analysis of the ACCESS subscales, literacy and the overall subscales had the greatest positive correlation, \( r = .925, p < .001 \). The high \( r \) value \( (r = .927, p < .01) \) indicates that either score could be used to predict student success, because they provide the researcher similar information.
Table 9

*Correlation of Biology EOCT with Subscales Reported for the ACCESS Test*

<table>
<thead>
<tr>
<th></th>
<th>EOCT</th>
<th>Listening</th>
<th>Speaking</th>
<th>Reading</th>
<th>Writing</th>
<th>Comp</th>
<th>Oral</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td>.168*</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>.156**</td>
<td>.255*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>.304*</td>
<td>.486*</td>
<td>.313*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>.232*</td>
<td>.381*</td>
<td>.457*</td>
<td>.531*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>.294*</td>
<td>.751*</td>
<td>.328*</td>
<td>.916*</td>
<td>.560*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral</td>
<td>.203*</td>
<td>.603*</td>
<td>.892*</td>
<td>.466*</td>
<td>.534*</td>
<td>.591*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy</td>
<td>.295*</td>
<td>.502*</td>
<td>.433*</td>
<td>.840*</td>
<td>.835*</td>
<td>.837*</td>
<td>.557*</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>.209*</td>
<td>.614*</td>
<td>.638*</td>
<td>.776*</td>
<td>.830*</td>
<td>.834*</td>
<td>.817*</td>
<td>.925*</td>
</tr>
</tbody>
</table>

*Note.* *p < .01, **p < .001, N = 164

As seen in Table 10, a stepwise multiple regression analysis was conducted to evaluate whether the other ACCESS subscale scores, together with gender, and LOT were better predictors of Biology EOCT scores than the ACCESS overall score. At step 1 of the analysis, the ACCESS subscale score was entered into the regression equation. Then at step 2, gender was entered into the regression equation. Finally, at step 3, LOT was entered into the regression equation. The regression equation for predicting a passing score on the Biology EOCT was

\[
\text{EOCT} = \text{Constant} + \text{ACCESS Subscale} + \text{Gender} + \text{LOT}
\]
Table 10

Linear Regression Analysis of EOCT and ACCESS Overall Subscale, Gender, and LOT

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>b</th>
<th>t(163)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td>349.105</td>
<td>8.198</td>
<td>42.584</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td><strong>Listening</strong></td>
<td>6.397</td>
<td>1.628</td>
<td>0.290</td>
<td>3.929</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>9.539</td>
<td>3.141</td>
<td>0.223</td>
<td>3.037</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>LOT</strong></td>
<td>-2.580</td>
<td>1.122</td>
<td>-0.170</td>
<td>-2.299</td>
<td>0.023</td>
</tr>
</tbody>
</table>

| **Constant**           | 363.702 | 6.712 | 54.184 | 0.000  |
| **Speaking**           | 2.589   | 1.015 | 0.197 | 2.251  | 0.012 |
| **Gender**             | 9.680   | 3.238 | 0.226 | 2.989  | 0.003 |
| **LOT**                | -2.061  | 0.990 | -0.161 | -2.081 | 0.039 |

| **Constant**           | 356.293 | 5.144 | 69.260 | 0.000  |
| **Reading**            | 6.897   | 1.202 | 0.405 | 5.738  | 0.000 |
| **Gender**             | 7.468   | 3.018 | 0.175 | 2.474  | 0.014 |
| **LOT**                |         |       |       |        |      |

| **Constant**           | 337.890 | 9.360 | 36.100 | 0.000  |
| **Writing**            | 9.802   | 2.129 | 0.338 | 4.604  | 0.000 |
| **Gender**             | 9.783   | 3.097 | 0.229 | 3.159  | 0.002 |
| **LOT**                | -2.369  | 0.944 | -0.185 | -2.508 | 0.013 |

| **Constant**           | 342.474 | 7.173 | 47.748 | 0.000  |
| **Comp**               | 8.769   | 1.435 | 0.427 | 6.109  | 0.000 |
| **Gender**             | 7.902   | 2.992 | 0.185 | 2.641  | 0.009 |
| **LOT**                | -1.987  | 0.890 | -0.155 | -2.232 | 0.027 |

| **Constant**           | 353.418 | 7.431 | 47.557 | 0.000  |
| **Oral**               | 5.392   | 1.380 | 0.296 | 3.907  | 0.000 |
| **Gender**             | 8.942   | 3.166 | 0.209 | 2.824  | 0.005 |
| **LOT**                | -2.365  | 0.967 | -0.85 | -2.446 | 0.016 |

| **Constant**           | 335.914 | 7.812 | 43.002 | 0.000  |
| **Literacy**           | 10.404  | 1.653 | 0.436 | 6.292  | 0.000 |
| **Gender**             | 9.133   | 2.954 | 0.213 | 3.091  | 0.002 |
| **LOT**                | -2.260  | 0.890 | -0.177 | -2.540 | 0.012 |

*Note.* *excluded from regression analysis, r = -.089, N = 164, p = .129

When each language domain is analyzed individually using a regression analysis with the covariates of gender and LOT introduced step-wise, the overall relationship remains positive. However, the number of points added to the Biology ECOT score based on domain varies from the results of the simple regression analysis reported in Table 7. By adding the covariates of gender and LOT, the amount of points added to the Biology EOCT decreased for listening and
reading, speaking remained the same, and writing increased. In all cases, male gender increases the Biology EOCT score. LOT has a negative effect on all domains except for reading, from which it was excluded having no significance in the variation in the Biology EOCT score.

Summary of Findings

Using linear regression analyses and controlling for the student and school characteristics of gender, ethnicity, and LOT, statistical analyses were performed on extant test score data of 164 ninth grade ELs in order to determine if there was a predictive relationship between the two test scores. Two research questions guided the analyses. First, how does performance on the ACCESS test of English language competence predict EL performance on the Biology EOCT in ninth grade after controlling for student characteristics? A moderately strong, positive predictive relationship was shown between Biology EOCT scaled scores and the ACCESS overall subscale after controlling for student characteristics. Second, which language subscale is the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing? No one language subscale on the ACCESS test was a stronger predictor of success on the Biology EOCT than another after controlling for student characteristics. However, combined language subscales were a stronger predictor of success than when they were analyzed independently. The literacy subscale, an equal combination of the reading and writing subscale scores, and the overall subscale scores accounted for the greatest variance in the Biology EOCT scores. Because the predictive capacity of the two subscales was similar, the ACCESS overall subscale was deemed sufficient for predicting EL success on the Biology EOCT.

Using regression analysis, student characteristics were taken into account in order to create a fuller picture of the EL who successfully passes the Biology EOCT. It was found that,
in general, males perform better on the Biology EOCT than females. Also, ELs with fewer years in an ESOL program had higher scores than ELs who had been in the program longer. However, the students who passed the Biology EOCT accounted for 24% of the total student sample indicating that three out of four ELs fail the Biology EOCT regardless of gender or years in an ESOL program. The one out of four ELs who passed the Biology EOCT did so with an average score of 416, well below 600, the highest score possible.

In Chapter 5, findings will be discussed in relation to the literature review. A discussion of the significance of this study to educators in Georgia and its limitations will be presented. Finally, suggestions for future research will be considered.
CHAPTER 5

CONCLUSIONS, DISCUSSION, AND IMPLICATIONS

Introduction

In accordance with federal mandates (U.S. Department of Education, 2004), English Learners’ (ELs) proficiency is assessed annually. Georgia has chosen the ACCESS test developed by WIDA as the assessment used to measure EL language proficiency in the four domains of language: listening, speaking, reading, and writing. The subsequent data are used not only for federal and state accountability purposes, but also by teachers and administrators as formative and summative evaluations. Decisions made using this data could have a profound impact on the opportunities afforded ELs as they plan their career pathways in high school. Being designated an EL limits classes available for electives, or alternatively affords students much need language support in order to be successful in their core content classes.

This study sought to determine if there is a predictive relationship between the ACCESS test and the Biology EOCT at one high school in Georgia. First, the ACCESS overall subscale scores and the Biology EOCT scores were analyzed to see if there exists a predictive relationship between the two tests. Second, the ACCESS subscale scores were analyzed to see which of the four language domains reported by the ACCESS test had the greatest predictive relationship to the Biology EOCT. The ACCESS provides combined domain subscale scores in order to differentiate between social and academic language proficiency. These too were analyzed in order to judge their usefulness in predicting EL success on the Biology EOCT. Student variables
such as gender and length of time in an ESOL program (LOT) were introduced into the analyses to learn what significance they might have on the predictive value of the data.

Discussion

Sample

The data employed in this study were drawn from one high school in North Georgia between 2007 and 2012. A purposive, random sample was taken of ELs in the ninth grade who also took the Biology EOCT the same school year. The student population \((N = 164)\) was predominately Latin@, constituting 94.5\% \((n = 154)\) of the sample. While the results may give insight into the needs of Latin@ ELs, they may not be applied to the Asian ELs who were part of the sample, because the sample size 5.5\% \((n = 9)\) was too small to be significant. The same may be said for ELs with disabilities, because only one student in the sample had a designated disability at the time of testing. Those included in this study were of normal cognitive ability. However, some understanding may be found pertaining to gender, because gender was fairly evenly distributed between males \((56.1\%, n = 92)\) and females \((43.9\%, n = 72)\).

Research Questions

Two questions framed the research:

1. How does performance on the ACCESS test of English language competence predict EL performance on the Biology EOCT in ninth grade after controlling for student characteristics?

2. Which language domain is the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing?
In order to answer the first question as to whether there is a predictive relationship between the ACCESS overall subscale scores and the Biology EOCT, a linear regression analysis was performed on the data. The results showed a significant, positive correlation between the ACCESS overall subscale scores and the Biology EOCT scores, \( r = .414, n = 164, p = .001 \). This means that if an EL student’s ACCESS overall subscale score is high, then the Biology EOCT score will also be high. If an EL student’s ACCESS overall subscale score is low, then the Biology EOCT score will be low, too.

A linear regression analysis was conducted to answer the second research question. Analyses were performed to determine which language domain subscale score was the strongest predictor of success on the Biology EOCT: listening, speaking, reading, or writing. All of the four domain scores on the ACCESS subscales had a significant, positive relationship with the Biology EOCT. The ACCESS writing subscale score had the greatest predictive value adding the greatest amount of points to the Biology EOCT score. For every unit increase in the ACCESS writing subscale score, 9.782 points were added to the Biology EOCT test. The reading domain contributed 6.930 points and the listening increased the Biology EOCT score 6.456 points. While significant, the speaking domain contributed the least to the Biology EOCT score, only adding 2.589 points per unit increase.

Other Analyses

In addition to individual language domain subscale scores, the ACCESS test reports combinations of subscale scores. A linear regression analysis was done on the oral, comprehension, and literacy subscale scores. The oral subscale score is made up of the listening and speaking subscale scores. The comprehension subscale score is a combination of the
listening and reading subscale scores. Finally, the literacy subscale score is comprised of the reading and writing subscale scores. Again, analyses showed a significant, positive relationship between all three composite subscale scores and the Biology EOCT scores.

Of the three composite subscale scores, the literacy subscale score exhibited the strongest predictive relationship to the Biology EOCT score. The literacy subscale augmented the Biology EOCT score by 10.448 points for every unit increase. This is the greatest increase reported by any subscale scores including the overall subscale score (9.707 points) and the writing subscale score (9.782 points). Of all of the subscale scores, the literacy subscale is composed of the two domains most closely identified with the academic register: reading and writing. Proficiency in an academic register is imperative for success in school. A high correlation between the literacy subscale score and the biology EOCT scores lends further support to previous research emphasizing the necessity for ELs to be proficient in the academic register in order to be successful in school (Boscardin & Aguirre-Munoz, 2006; Johnson, 2009).

Finally, a step-wise, linear regression analysis was performed in order to find out what role gender and LOT in an ESOL program played in EL success on the Biology EOCT. Both covariates were found to play a useful role in predicting EL success on the Biology EOCT. In terms of gender, males significantly outperformed females on the Biology EOCT three to one. LOT had a negative impact on the Biology EOCT score. What follows is a more in-depth discussion of these two results.

**Gender**

Gender played an important role in this study. EL males scored much higher than females on both the ACCESS test and the Biology EOCT. Three out of four ELs who were successful on the Biology EOCT were male (75%, n = 40). In addition, being male added 8.681
points to Biology EOCT score, while being female added no points for every unit increase in the ACCESS overall subscale score. These results show a clear gender gap between males and females in this sample.

Recently researchers have observed a trend toward a closing gender gap. The National Assessment of Educational Progress (NAEP) (Center on Educational Policy, 2010) reported that, in general, females outperformed males in reading in the fourth, eighth and 12th grades from 1992 until 2009. In the same report, the CEP concluded that the gender gap in math, in which males outperformed females, had closed. The results showed that male and female students were on parity. This may be explained in part at the secondary level by the increased enrollment of females in higher level math courses (Niederle & Vesterlund, 2010).

Chang and Choi (2011) have one possible explanation for the so-called gender gap. In a survey of 8,912 eighth grade students, they found that while females had higher grades in math classes on tests given in the classroom than males. Alternately, males had higher math scores on standardized tests. They suggested that “female students tend to be more socially oriented in their learning strategies and behaviors, whereas male students are more independent and autonomous in completing learning tasks” (p. 17). The Latinas in this study had a higher mean on the listening, $M = 4.092, n = 72$, and speaking $M = 4.475, n = 72$, than they did for reading, $M = 3.879, n = 72$, and writing $M = 3.874, n = 72$. This lends support to Choi and Chang’s supposition that females perform better in a more collaborative setting.

---

17 NAEP did not differentiate results by ethnicity.
18 Ellison and Swanson (2007) found that students scoring over 800 on the SAT were two to one male. In 2003, Xie and Shauman (as cited in Niederle and Swanson, 2010) that the top 5% of high school math students were male outperforming females two to one.
19 Chang and Choi did not differentiate results by ethnicity.
Niederle and Vesterlund (2010) studied the influence of competition on gender success. They found when competition is not a factor in a task, females and males performed similarly. However, males responded strongly to the addition of competition and compensation into an activity, significantly outperforming females. In fact, the authors observed females actively avoiding competition. They cautioned researchers not to rely solely on the results of standardized test to evaluate a person’s interest or ability, because it may not paint a complete picture.

At the high school used in this study, Latino ELs consistently outperformed Latina ELs on the Biology EOCT. In an interview with Beth Jordan (2013), the lead teacher in the ESOL department at DHS, she postulated scenarios based on conversations with students that may explain the findings in this study. Speaking specifically about the ELs at her high school, Jordan said,

The Latino males we have typically come to us on grade level in their native language. In contrast, the females usually have significant gaps in their education, many never having attended school (in their home country) until they reach (this high school) at age 14 to 17 years old. Many times this is because a learning disability has been identified and they cannot afford a special school. Other times they say they have had to care for siblings or work. Also, some of the families are migrant and they are not in one place long enough to be in school or be promoted from one grade to the next.

The literature supports the advantage for ELs coming to our schools at grade level (Krashen, 1996; Thomas & Collier, 2002). If Latinos are indeed coming to this high school on grade level, then they have what Krashen (1996) called “de facto bilingualism.” This means that the content knowledge, literacy skills, and learning strategies that ELs bring with them create similar conditions to that of a successful bilingual program. It provides ELs with comprehensible input (Krashen, 1982) so that they may more easily acquire English. If Latinas are not coming to this school with de facto bilingualism, then they are immediately at a
disadvantage as compared to some Latinos in this study. Along with interrupted formal
schooling (SIFE) and transiency, the type of language support programs at the high school may
explain why so many ELs in this study failed the Biology EOCT (DeCapua, Smathers & Tang.,
2007; Fletcher & Navarrete, 2011; Morse, 2005; NWLC & MALDEF, 2009). While some ELs
may have come to this high school with undiagnosed learning disabilities, normal ELD
milestones may appear to be manifestations of a learning disability (Baca & Cervantes, 2004;
Fletcher & Navarrete, 2011; Gersten & Baker, 2000; Klinger, Hoover, & Baca, 2008), especially
if their AoA was in the middle grades or in high school.

According to Fletcher and Navarrete (2011), “a very typical profile of a student acquiring
a second language reveals a high nonverbal score with poor performance in the language-based
areas of reading, writing, and speaking or listening . . . indicative of a discrepancy between
scholastic aptitude and achievement” (pp. 11-12). Fletcher and Navarrete (2011) tell us that SLA
is “a transient problem amenable to educational intervention” (p. 11). They warn that an EL’s
“real learning potential may be masked by their inability to learn due to language differences” (p.
12). It is imperative that we provide these students with the support they need to be successful in
the regular classroom.

Bilingual education20 may meet the needs of the three-fourths of students in this sample
who failed the Biology EOCT (Baca & Cervantes, 2004). Bilingual education is intended to
support the acquisition of English while allowing ELs to learn content in their L1. For this
sample, it would give Latinas the same opportunity to learn content as their male counterparts
who came to this school with de facto bilingualism because they were on grade level in their L1.

---

20 The GADOE allows for transitional bilingual education for one year. While research has
shown that late-exit dual language programs are the most effective in preparing ELs for success
in the regular classroom (Thomas & Collier, 2002), any language support in an EL’s L1 will
improve their opportunity to learn.
Because the culture of the majority of the ELs included in the sample is Latino, it is reasonable to look into the possible role of culture on the results of these analyses. In research conducted by the National Women’s Law Center (NWLC) and the Mexican American Legal Defense and Educational Fund (MALDEF) (2009), many issues are discussed that contribute to the 41% of Latinas dropping out of high school nationwide. Besides Latinas having the highest pregnancy and teen birth rate of all ethnicities, Latinas surveyed said they are burdened by the gender and ethnic stereotypes of being “submissive underachievers and caregivers” (NWLC & MALDEF, 2009, p. 2).

Given the high dropout rate for Latinas across the nation, gender plays a significant role in the success of Latinas in American schools. In Georgia, the total Latin@ drop-out rate for the 2010-2011 school year was 4.6% (GADOE, 2013). The same year at the high school in the study, the total Latin@ drop-out rate was 0.3%, well below the state and national averages. The drop-out rate for the EL subgroup was 5.5% at the state level and 1.5% at the high school level. This was the largest drop-out rate for any subgroup reported for AYP at either the state or high school level. It would behoove the agencies involved to look more closely at this data in order to identify ways in which to help ELs graduate from high school.

*Length of Time in an ESOL Program*

An unexpected outcome of this research was the negative effect of additional time spent in an ESOL program. Baker (2011) and Flores et al. (2012) found that being in a language support program after three years was detrimental to EL academic success. These results may be explained by the fact that each EL is unique and depending on a plethora of factors related to home, school, and individual educational opportunities, will acquire English at differing rates.
Hence, these results do not necessarily run counter to previous research suggesting that ELs need four to seven years of language support in order to be academically successful in school (Hakuta, 2002; Hakuta et al., 2000; Thomas & Collier, 2002). If some Latinos are indeed coming to this high school on or near grade-level, their de facto bilingualism (Krashen, 1996) may, in part, explain their success on the Biology EOCT and in English language acquisition in general.

The focus of the discussion thus far has been on the one-fourth of ELs who have been successful on the Biology EOCT. The overarching piece of data that should frame conclusions drawn from this study is that the overwhelming majority of ELs (76%, \( n = 124 \)) failed the Biology EOCT. Three out of four students did not have the knowledge or academic language in English to be successful on the Biology EOCT test.

Unlike this study, Baker (2011) was able to follow students across multiple years and found that while their English language proficiency as measured by the ACCESS test continued to increase over time, academic achievement scores plateaued after the third year. She posited several explanations for this phenomenon which echo Oakes’ (1985) observation about differences in educational inputs resulting in differing outputs. One of these include EL tracking into courses that separate them from fluent English speaking peers with whom to collaborate and interact using an academic register (Boscardin & Aguirre-Munoz, 2006; Ellis, 2008). Another explanation suggested that because of an EL’s fluency in a social register, generally acquired sooner than an academic register (Cummins, 1999, Hakuta, 2002; Hakuta, et al., 2000; Thomas & Collier, 2002), teachers may not be aware that explicit instruction in the academic register is needed (Boscardin & Aguirre-Munoz, 2006) and fail to provide comprehensible input so that students have an opportunity to learn content material.
The purpose of this study was not to evaluate the programs used at this high school. Consequently, information was not gathered as to the learning environment, or programs used at this high school to teach biology to ELs through the years for each cohort of students. However, the models currently used are push-in and sheltered classes (Jordan, 2013) with English as the language of instruction. An early-exit, dual-language academy is an option for non-English speaking Latinas within their first year in-country (GADOE, 2013) for English Language Arts. Bilingual education in content areas other than ELA is not an option at this high school.

Conclusions

In regard to the first research question, a significant, positive, predictive relationship was found to exist between the ACCESS test and the Biology EOCT. Therefore, the ACCESS can be used by teachers and administrators to predict EL success on the Biology EOCT. Using the regression equation, ELs can be identified for academic intervention in order to prepare them to be more successful on the Biology EOCT. Because ELs have consistently scored lower on the Biology EOCT than in other content areas, the ACCESS scores might be a useful tool in identifying ELs who need additional support in other content areas as well. Alternatively, ELs who perform successfully on the Biology EOCT test may be considered to have acquired sufficient English proficiency to not warrant inclusion in academic English support courses. Their electives can be chosen to fulfill their career pathway requirements.

In terms of the second research question, all four language domain subscale scores on the ACCESS test equate to a positive increase per unit in an EL’s Biology EOCT score. The writing subscale score was the strongest predictor of success followed by reading. Because these two domains are so closely linked to academic learning, it is not surprising that their combined score
in the literacy subscale should be the best predictor of EL success on the Biology EOCT. While the literacy subscale score provides a stricter criterion than the overall subscale score because it is specifically linked to the academic register, the overall subscale score meets the requirements of NCLB (U.S. Department of Education, 2004) in that it reports a composite score comprised of all four language domains. Since the overall subscale has a strong, positive correlation to the Biology EOCT, it may also be used in predicting EL success.

Gender played a surprising role in the outcomes of this research. Latinos were more successful on the Biology EOCT at the high school in this study than Latinas three to one. Latino culture may play some role in differing expectation for males and females in terms of educational outcomes (DeCapua, et al., 2007). Anecdotal evidence indicates that some Latina ELs who entered this high school come with limited or interrupted schooling in their home countries (Jordan, 2012). These would be plausible explanations for why Latina ELs underperformed on the Biology EOCT at this high school.

Implications and Recommendations

There are important implications to be drawn from this research. Beginning in 2013, the Georgia DOE disseminated new exiting criteria for ELs (GADOE, 2013). As in previous years, the first ACCESS score considered in exiting an EL from language support services is the overall subscale score. An overall subscale score of 5.0 remains the cut-off score for a clear exit from the ESOL program. If the ACCESS overall score falls between 4.0 and 4.9, then a second criterion must be considered in determining if a Language Assessment Conference (LAC) should be convened among teachers to decide whether or not the EL should be exited from ESOL services. In the past, a high school EL’s score on the Literature EOCT was used in addition to
the ACCESS overall subscale score. In Grades 1 through 8, an EL’s CRCT score in reading was used. With the adoption of the CCSS, the CRCT no longer reflects the standards taught in Georgia schools. While new CRCTs and EOCTs that conform to the CCSS are being developed, the state has chosen to supplement the exiting criteria for the ESOL program with the literacy subscale score.

Besides the questionable policy of using scores derived from the same test (The American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, 1999), the anecdotal information from the high school in the study is that fewer ELs have been exited through a LAC this year than in previous years. This may be due to the fact that the literacy subscale score was found to be a stronger predictor of success and thus has a lower mean score than the overall subscale score. The result is that it is unlikely for an EL to score higher on the literacy subscale than the overall subscale. For example, the ELs included in the sample \(N = 164\) scored an average of 3.96 on the literacy and 4.21 on the overall subscale. Consequently, this may have an impact on 2013 AMOAs at both local and state levels. It may be in both the students’ and the state’s best interest to choose an alternative criterion for determining whether an EL should be exited from ESOL services.

An implication that can be drawn from the data and research involves gender. While nationally schools are failing Latinas, Latinas in Georgia are dropping out of school at a much lower rate when compared to national figures. Moreover, all Latinas at the high school in the study have a lower drop-out rate than the state. In addition, graduation rates for Latinas have been increasing over the last 6 years at both the state and school levels. Given this data, it may be inferred that the high school in the study is working to meet the needs of its Latino students in

\[21\] In the spring of 2012, the school district exited eight ninth grade ELs using the LAC. In the spring of 2013, the school district exited one ninth grade EL using the LAC (Jones, 2013).
terms of school completion. What is of concern is the retention rate at the high school. From 2006-2011, Latinas made up 72.3% of the students who were retained (GADOE, 2013). This percentage refers to all Latinas in Grades 9 through 12. However, demographics from the 2010-2011 school years show that 43% of all high school students were in ninth grade (GADOE, 2013). It may be assumed that most student retentions occur in ninth grade. The ninth grade curriculum offers a great challenge for all students and particularly Latinas. Studies such as this one may help identify students for intervention programs so that they can be successful the first time they attempt challenging academic courses. In addition to the language support programs in place, bilingual education in all content areas should be offered for ELs who do not come to school with content knowledge at grade level in their L1. If an EL is required to learn both language and new content at the same time, this limits the amount of working memory available for content processing (Sousa, 2011). The result is limited opportunities to learn either language or content. As an EL’s L2 proficiency increases, working memory is freed from language processing and can be utilized for content processing. Bilingual education provides ELs with content knowledge while they are learning English (Baca & Cervantes, 2004).

In 2009, the school district established a second, alternative high school for underperforming students for the purpose of intervening before students fail. The second, smaller high school offers flexible scheduling and on-line courses for students who do not do well in the traditional high school setting. Free tutoring is offered 4 days a week after school. Currently, the second high school provides sheltered content courses in ninth grade and push-in ESOL services in the upper grades to support the linguistic needs of ELs. Recalling that only three out of four ELs were successful on the Biology EOCT, these practices may improve the success of
all ELs at this high school. Again, ELs would benefit from bilingual education at this school as well.

It is beyond the scope of this study to determine whether the ACCESS is a valid and reliable test of EL English proficiency. However, the findings of this study raise concern as to whether the ACCESS test is exiting ELs before their academic register is developed enough to meet the linguistic demands of the Biology EOCT. To wit, using the regression equation, we find that, in general, a female EL must score a 6.0 on the ACCESS overall subscale, in order to be successful on the Biology EOCT. A 6.0 is a perfect score on the ACCESS indicating that the EL has reached native speaker proficiency. If the results of this study could be generalized, then states that exit female ELs at the 5.0 as recommended by WIDA (2013) are terminating support services that may still be necessary to ensure success for these students in regular content courses.

From previous research, we expect to see minority students, in this case ELs, underachieving on standardized tests (Abedi, 2002; National Center for Educational Statistics, 2010). In the results of this study and others (Baker, 2010; Flores et al., 2012; Garcia, Lawton, & Diniz de Figueiredo, 2010; Mahoney, Haladyna, & MacSwan; 2007; Parker et al., 2009) we see ELs making progress on an English language proficiency test and yet not being successful on content tests. If the ACCESS does indeed test the language proficiency of the academic register (WIDA, 2013), why then does this progress in language acquisition not translate into academic success in the classroom? Viewing this conundrum through the lens of Critical Race Theory (CRT), we see race, racism, and power colliding (Choudhury, 2008; Delgado & Stefancic, 2001). The Federal pressure on states and local educational agencies is to report an ever-increasing percentage of ELs exiting language support programs each year in order to make AMOAs. The
politics of high stakes testing create an environment that may contribute to states selecting a test that does not meet the academic needs of students, but rather supports the need to meet federal mandates. Those responsible for preparing ELs to meet the demands of a 21st century workforce should carefully consider first what is best for the student.

With the advent of the CCSS, educators have a de facto national curriculum. Since most members of the WIDA consortium are using the CCSS, it would better serve ELs and other stakeholders to assess EL proficiency at grade level rather than in grade level clusters. The ACCESS test would be a more valid measure if it linked test items to the language functions and vocabulary ELs have learned at grade level. This would be more reliable than linking the test items to language ELs have yet to be exposed and arbitrarily scaling scores for students in the lower grades of the cluster. It is highly recommended that grade level tests should be developed to measure EL proficiency for both formative and summative purposes.

Further Research

Several studies presented themselves for further research based on the results of this study. The data used in this research were from the 2007 version of the ACCESS test. In the spring of 2013, the 2012 version of the ACCESS was introduced. It would be interesting to see if analyses performed on the results of the new test version produce similar conclusions.

While this research compared the ACCESS subscale scores to the Biology EOCT scores, it would be useful to determine if the Biology EOCT is indeed a better predictor of overall EL success, as compared to other EOCTs such as the ninth grade Literature and Composition EOCT or the Geometry EOCT.
Beginning in 2013, high school students no longer have to pass the GHSGT in order to graduate from high school. EOCTs in certain content areas will replace this graduation requirement. In addition, EOCTs now compose 20% of a student’s course grade. With this new emphasis on EOCTs, teachers must prepare their student specifically for this test using the CCSSs to guide their instruction. It would be interesting to compare EOCT scores before 2013 to those after to see if the new value placed on EOCT scores at the state level translates to higher scores at the school level.

In the past, undocumented students found American high school to be the end of their academic pursuits in the United States. In 2012, the Georgia state legislature passed Senate Bill 458 making it virtually impossible for undocumented students to pursue higher education in the state of Georgia. Knowing that education was barred for these students would likely result in less motivation for these students to do well or even graduate from high school. With the passage of Deferred Action for Childhood Arrivals (DACA) (US Citizenship and Immigration Services, 2013), undocumented youth who meet certain requirements may apply for a 2-year employment authorization. Whether DACA will be interpreted by the Georgia Board of Regents as allowing these students in-state tuition in the Georgia university system is still contested. It will be interesting to see if this new opportunity will motivate undocumented students, some of whom may be ELs, to finish high school thus increasing the high school completion rate and lowering the high school dropout rate.
REFERENCES


Flores v. Arizona, 516 F.3d 1140 (9th Cir. 2008), cert. granted, 129 S. Ct. 893 (2009).


Georgia Department of Education. (2007). What Georgia educators need to know about Georgia’s testing program. Atlanta, GA: Author.


Mahoney, K., Haladyna, T., & MacSwan, J. (2007). A validity study of the Stanford English Language Proficiency Test (SELP) as used for classifying English Language Learners. Annual meeting of the University of California Linguistic Minority Research Institute (UCLMRI). Arizona State University, Tempe, May 2-5.


Mulkerson, L. (2009). Schools succeed with teamwork and the HSTW practices. In *High schools that work: Building the capacity of school leaders and teachers to implement reform at all levels*. Atlanta, GA: SHREB.


## ACCESS for ELLs® English Language Proficiency Test

### Teacher Report – 2011

#### District: [Blank]

#### School: [Blank]

#### Grade: 9

#### Tier: C

#### Grade Level Cluster: 8-12

#### State ID: [Blank]

#### District ID: [Blank]

#### Birth Date: [Blank]

**Report Purpose:** This report provides information regarding the levels of social and academic English language proficiency the student has attained. Social language is used to communicate for everyday purposes. Academic language is used to communicate the content of language arts, mathematics, science, and social studies. This report can be used to monitor progress from year to year and to help determine instructional strategies by content areas and standards. Please refer to the ACCESS for ELLs® Interpretive Summary for more information on the meaning and use of these scores. You may also refer to the complete Interpretive Guide for Score Reports at www.wida.us for more detailed information.

## Student’s level of English proficiency by language domains

<table>
<thead>
<tr>
<th>Language Domain</th>
<th>Scale Score</th>
<th>Confidence Band</th>
<th>Proficiency Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Possible</td>
<td>See Interpretive Summary for definitions</td>
<td>Possible</td>
</tr>
<tr>
<td></td>
<td>100 - 600</td>
<td></td>
<td>1.0 - 6.0</td>
</tr>
<tr>
<td>Listening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral Language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Score (Composite)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- A: Oral Language = 50% Listening + 50% Speaking
- B: Literacy = 50% Reading + 50% Writing
- C: Comprehension = 70% Reading + 30% Listening
- D: Overall Score = 25% Reading + 25% Writing + 15% Listening + 15% Speaking

Overall Scores are computed when all 4 domains have been completed.

### Student’s performance by WIDA English Language Proficiency Standards

Due to varying numbers of items and their levels of difficulty, raw scores should be used with caution. See the Interpretive Guide for Score Reports for details.

#### COMPREHENSION (Listening and Reading)

<table>
<thead>
<tr>
<th>English Language Proficiency Standards</th>
<th># of Items Correct</th>
<th>Total # of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social &amp; Instructional Language</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Language of Language Arts</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Language of Mathematics</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Language of Science</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Language of Social Studies</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

#### SPEAKING TASKS

<table>
<thead>
<tr>
<th>English Language Proficiency Standards</th>
<th>Raw Score</th>
<th>Total Possible Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social &amp; Instructional</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Language Arts/Social Studies</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Mathematics/Science</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

#### WRITING TASKS

<table>
<thead>
<tr>
<th>English Language Proficiency Standards</th>
<th>Linguistic Complexity</th>
<th>Vocabulary Usage</th>
<th>Language Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social &amp; Instructional</td>
<td>Raw Score</td>
<td>Total Possible Points</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mathematics &amp; Science</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Language Arts &amp; Social Studies</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

**Description of Proficiency Levels**

- 1 novice – Knows and uses minimal social language and minimal academic language with visual and graphic support
- 2 Beginning – Knows and uses some social and general academic language with visual and graphic support
- 3 Developing – Knows and uses social and general academic language with visual and graphic support
- 4 Developing – Knows and uses social and general academic language with visual and graphic support
- 5 Developing – Knows and uses social and general academic language with visual and graphic support
- 6 Beginning – Knows and uses social and general academic language at the highest level measured by this test

April 28, 2011
APPENDIX B

PERMISSION LETTER FROM THE SCHOOL DISTRICT
November 13, 2012

To Whom It May Concern:

I have reviewed Susanne Wakeman’s research request. Permission has been granted for Susanne to collect the data as submitted in her proposal. The only request that I make is upon completion of the study that the findings are shared with the district. If I can be of further assistance in this process, please let me know.

Sincerely,

Rhonda Hayes, Ph.D.
Director of School Support Services
Dalton Public Schools