

ABSTRACT

Title of Dissertation: **HOW LONG DOES IT TAKE FOR ENGLISH
LEARNERS TO BE RECLASSIFIED AS ENGLISH
PROFICIENT?**

Supreet G. Anand, Doctor of Philosophy, 2018

Dissertation directed by: Dr. Jeff MacSwan, Department of Teaching and Learning,
Policy and Leadership (TLPL)

English Learners (ELs) constitute one of the fastest growing student populations enrolled in K–12 schools. One important factor influencing the academic achievement of ELs is attaining English language proficiency (ELP). Once ELs attain ELP, they are reclassified as English proficient. In practice, exiting or reclassify ELs indicates that they no longer need or receive specialized language supports or regular monitoring using ELP assessments. Reclassification is a high-stakes decision. A change from an EL to a non-EL status may impact state and federal funding. Some federal and state funds are allocated based on the per-pupil count of ELs enrolled in a school district. Understanding the relationships between reclassification policies, reclassification rates, and reclassified student outcomes is necessary to make decisions about resources and expectations for ELs.

This study explored the population of ELs in one state and addressed the following research questions:

1. How long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED)?
2. Does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools?
3. How do select individual student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time?

The analytical sample consisted of ELs who started school for the first time in SY 2010-11 (Cohort 1) or in SY 2011-12 (Cohort 2) and who were continually enrolled during the observation period ending in SY 2015-16. The sample included students enrolled in all grades K-12; no new students were added to the analytical sample. The data were analyzed using a discrete survival analysis. The results indicated that ELs starting school in grades K-8, took on the average of four years to achieve ELP. ELs starting school in upper grades were reclassified faster than ELs who starting in Kindergarten. ELs identified as Hispanic or Latino, received free or reduced lunch, and receiving Special Education services had longer median times to ELP than their peers.

**HOW LONG DOES IT TAKE FOR ENGLISH LEARNERS TO BE
RECLASSIFIED AS ENGLISH PROFICIENT?**

by

Supreet G. Anand

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:

Professor Jeff MacSwan, Chair
Dr. Gary Cook
Dr. Claudia Galindo
Dr. Elisa Gironzetti
Dr. Joan Lieber
Dr. Kellie Rolstad

© Copyright by
Supreet G. Anand
2018

Dedication

This dissertation is lovingly dedicated to my mother, Dr. Jaswaran Grewal and my husband, Sachin Anand, whose unconditional support made this possible; and to Asher and Aavi for whom all of this was accomplished.

Acknowledgements

Completing this dissertation would not be possible without the guidance of my committee members, help from friends, and support from my family.

I would like to begin my thanking my advisor Jeff MacSwan. I made it to the finish line because of his expertise and advice. I would also like to thank Gary Cook, Claudia Galindo, Elisa Gironzetti and Joan Liber for serving as committee members and for their contribution in completing this work. I am deeply grateful to Jeffery Harring for his time and guidance in spite of personal hardships. This work would not have been possible without the data and support provided by the wonderful experts at the New Mexico Department of Education. I am also grateful to Joy Jones for keeping me in track through all the forms and logistics of the program.

I am deeply indebted to my critical friend, Rosalie Grant, for her patience, encouragement and candid feedback. I could not have completed this work without her generosity of spirit and time. I am also thankful to Molly Faulkner Bond for sharing her work and expertise. I am also grateful to Mansour Fahimi for his editing support.

My family deserves a special recognition. Words cannot express how indebted I am to my mother for all the sacrifices she made on my behalf. Her prayers sustained me throughout my life. Last, but not least, I would like express my gratitude to my beloved husband for being my rock. Without his encouragement and support, this would not have been possible.

Contents

Dedication	ii
Acknowledgements	iii
I. Introduction	1
Legislation and Educational Policies for English Learners	2
NCLB Waivers	3
Every Student Succeeds Act	5
English Language Proficiency	6
Problem Statement	10
Purpose of this Study	12
State Context of the Study: New Mexico	13
Summary	14
II. Literature Review	15
Length of Time to Achieve English Proficiency for the Purpose of Reclassification	16
Methodological Challenges to Determining Reclassification Criteria and Rates	25
Student Characteristics and Reclassification Patterns	26
Age and Grade of Entry	26
Initial Levels of English Proficiency	29
ELs with Disabilities	31
Types of Programs and Services for ELs and Reclassification Rates	32
Poverty	35
Gender	36
Home Language	39
Conclusion	40
III. Methodology	42
Classification and Reclassification Policy for ELs in New Mexico	42
Measurement Instrument: The ACCESS	43
Assessment Scores	47
Validity and Reliability of ACCESS	48
Description of the Data Set	49
Building the Analytical Sample	50
ACCESS data set	52
STARS data set	54
Variables	56
Analytical Approach	61
Data Censoring	62
Survival Analysis: Description of the Model	63
Fitting the Model to Data	66
IV. Findings	70
How long does it take ELs to be reclassified as English proficient in New Mexico?: Main Effect of <i>Time</i>	70
Model Estimates of Time to Proficiency	74
Effect of Starting Grade at Which ELs Enter the School System	75
Estimates of Reclassification Rates for Grades K-8	76

Median Rates of Reclassification for Grades K-8	78
Influence of Starting Proficiency Levels on Median Time to Proficiency for Grades K-8	79
Reclassification Rates of High School ELs	82
Influence of Select Student Characteristics on the Probability of Reclassification for ELs	83
Gender	84
Race and Ethnicity	85
Impact of Poverty	87
ELs with Special Needs	89
Characteristics of ELs at Risk of Becoming Long-Term ELs	91
Right Censored Students	92
Limitations of the Study	94
Summary	95
V. Discussion	97
Question 1: How long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED)?	98
Question 2: Does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools?	99
Question 3: How do select individual student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time?	101
Gender	101
Poverty, Race and Ethnicity	102
ELs with disabilities	105
Long-term ELs	110
Conclusions	113
Data collection and use at the local district and State level	113
Reclassification Policy	114
References	119

List of Tables

Table 1:	Reliability of ACCESS scores	49
Table 2:	Percentage of students grades K -12 completing WIDA ACCESS testing for English proficiency by exclusions and school year.....	51
Table 3:	Number of students tested who listed English as home language by ethnicity, grades K-8.....	59
Table 4:	Percentage of students in the sample population by selected student characteristics.....	60
Table 5:	Model Information	68
Table 6:	Life table by school year, combined cohorts (K-8)	72
Table 7:	Analysis of Maximum Likelihood Estimates by Years, Combined Cohorts, K-8.....	74
Table 8:	Base Model Estimate of Years to Achieve Survival Rate of 50%, Combined Cohorts, K-8.....	75
Table 9:	Enrollment County by Grade Band (K-12).....	76
Table 11:	Estimate of Years to Achieve Survival Rate of 50% by Grade Band, Combined Cohorts, K-8.....	78
Table 12:	Frequency distribution of ELs reaching proficiency by year and by proficiency level.....	80
Table 13:	Reclassification Patterns Among High School Grades.....	83
Table 14:	Estimates by Gender	84
Table 15:	Gender differences in likelihood of achieving proficiency for ELs	85
Table 16:	Base Model Estimate of Years to Achieve Survival Rate of 50% for Females, Combined Cohorts, K-8.....	85
Table 17:	Estimates by Ethnicity	86
Table 18:	Likelihood of achieving proficiency for ELs by Ethnicity	86
Table 19:	Estimate of Years to Achieve Survival Rate of 50% by Ethnicity Combined Cohorts, K-8.....	87
Table 20:	Estimates by Ethnicity	88
Table 21:	Likelihood of achieving proficiency for ELs by Ethnicity	88
Table 22:	Base Model Estimate of Years to Achieve Survival Rate of 50% for ELs from Poverty, Combined Cohorts, K-8	88
Table 23:	Estimates by Special Education Services	90
Table 24:	Likelihood of achieving proficiency for ELs by Ethnicity	90
Table 25:	Estimate of Years to Achieve Survival Rate of 50% for ELs with disabilities, Combined Cohorts, K-8.....	90
Table 26:	Characteristics of ELs in grades K-8 who did not become proficient within 5 years.....	92
Table 27:	Number and percentage of ELs in grades K-8 who did not become proficient during the observation period starting with SY 2010-11 and ending in SY 2015-16	93

List of Figures

Figure 1: Hazard Rate (proportion who were reclassified)	73
Figure 2: Survival Rate by School Year	73
Figure 3: Median years to reclassification by grade band	79
Figure 4: Cumulative Proportion of ELs reclassified by year grouped by proficiency level after one year of instruction	81

I. Introduction

English learners (ELs) constitute the fastest growing population in public schools in the United States. Data show that: in SY 2014–15, there were 4,806,662 ELs in the United States, comprising 9.6 percent of all students in kindergarten through grade 12, (U.S. Department of Education, 2017), compared to 4,472,563 in SY 2011–12, or 9 percent (U.S. Department of Education, 2015). School-age children (ages 5–17) who spoke a language other than English at home increased from 4.7 million in 1980 (National Center for Educational Statistics: The Condition of Education, 2011) to 11.9 million in 2015 (United States Census Bureau, 2015). Of the school-age children who spoke a language other than English at home in 2015, about 71 percent spoke Spanish, about 13 percent spoke an Indo-European language other than Spanish, 11 percent spoke an Asian/Pacific Islander language, and 4 percent spoke another language. The remaining 1 percent did not speak English (United States Census Bureau, 2015). While the number of ELs is increasing, the academic gap between ELs and their English-speaking peers is significant and continues to persist (Abedi, 2002; August & Hakuta, 1997; Graham, 1987). The 2011–12 *Civil Rights Data Collection* shows that a disproportionate number of ELs leave school without the reading and mathematical skills needed for our increasingly complex, global economy (ED, 2014).

ELs are a diverse group with different education and social experiences, native cultures, and languages. Interrelated linguistic and sociocultural factors such as English language proficiency (ELP), school program effectiveness, equity, poverty, family background, native culture, and teacher preparedness play a role in understanding the achievement gap. However, ELP is one contributor to the unexplained variance in

achievement of ELs and their native English-proficient peers (Torres & Zeidler, 2001). Not being proficient in English puts ELs at a unique educational disadvantage in U.S. schools, where the medium of instruction and assessment is English (August & Hakuta, 1997; Reardon & Galindo, 2007). Many ELs experience challenges with speaking, reading, writing, or understanding the English language. These challenges present sufficient barriers in their ability to meet the state's proficiency level of achievement on state assessments and successfully achieve in classrooms where the language of instruction is English. Regardless of the challenges or barriers, ELs are expected to acquire the English language at the same time they are learning subject-matter content. Without attention to both language and content learning, ELs may not experience the success with academic standards necessary for all children (Abedi & Gandara, 2008; August & Hakuta, 1997). In the last few decades, legislative policies, such as those discussed below, formally recognized the importance for ELs to attain English proficiency to access the content as a right to equitable education and a civil rights issue.

Legislation and Educational Policies for English Learners

Over the past 40 years, legal and legislative actions have shaped the education of ELs. In the landmark case of *Lau v. Nichols*, the Supreme Court ruled that failing to accommodate the language needs of ELs makes a “mockery of public education” and is a violation of their right to a federally funded education free from national origin discrimination (*Lau v. Nichols*, 1974). The ruling greatly expanded the rights of all children with limited English skills to access additional assistance in learning English. Not being afforded this access violates the civil rights of ELs. The Bilingual Education Act (Title VII) of 1968 acknowledged the educational challenges faced by ELs and

allocated federal funds to provide services to support academic achievement and English proficiency.

Amended and reauthorized numerous times under the Elementary and Secondary Education Act (ESEA), the 2002 English Language Acquisition, Language Enhancement, and Academic Achievement Act (Title III) replaced the Bilingual Education Act. The No Child Left Behind (NCLB) Act instituted a greater focus on standardized testing and held schools accountable for student performance in the content areas of reading/language arts, math, and science (NCLB, 2001, Title I, *Sec 1111*). Title III of NCLB required states to assess all ELs using a standardized ELP test and to report the progress towards the attainment of English proficiency (NCLB, 2001, Title III, *Sec 3113*). Title I of NCLB required states to include EL students in academic achievement assessments in reading/language arts and mathematic (NCLB, 2001, Title I, *Sec 1111*). In a majority of states, the content assessments are administered in English, and ELs typically struggle to meet the law's annual progress requirements, resulting in serious consequences for the students and their schools. Irrespective of the challenges or barriers, the state accountability systems under NCLB expected ELs to reach the same achievement levels as their native English-language peers within the same timeframe.

NCLB Waivers

In 2011, states had the opportunity to seek waivers from key Title I NCLB requirements from the ED. Notable among the NCLB requirements that could be waived included the 2013–2014 timeline for determining Adequate Yearly Progress, and the flexibility to use Title I funds for school and district improvement to identify and support low-performing schools. In a letter dated September 23, 2011, to the Chief State School

Officers, then Secretary Arnie Duncan stated that in exchange for flexibility, states would have to develop and implement comprehensive state plans:

designed to improve educational outcomes for all students, close achievement gaps, increase equity, and improve the quality of instruction. This flexibility is intended to build on and support the significant state and local reform efforts already under way in critical areas such as transitioning to college- and career-ready standards and assessments; developing systems of differentiated recognition, accountability, and support; and evaluating and supporting teacher and principal effectiveness (ED, 2011).

Under the conditions for granting waivers to states, districts were required to: (1) adopt college and career-ready standards in at least reading/language arts and mathematics; (2) implement such standards statewide for all students and schools; and (3) develop and administer annual, statewide, aligned, and high-quality assessments, and corresponding academic achievement standards that measure student growth in at least grades 3–8 and at least once in high school. States were also required to support ELs in reaching these standards by adopting ELP standards that correspond to its college and career-ready standards and that reflect the academic language skills necessary to access and meet the new college and career-ready standards. States also are required to develop and administer ELP assessments that are aligned with college and career-ready standards (ED, 2011). Most states adopted college and career-ready standards that define the literacy expectations for all students to be successful in college and workforce training programs. To measure student learning of the state-adopted content standards, states use standardized assessments aligned to these standards. The accountability systems

developed by states under these waivers were more complex than those developed under the NCLB statutes and captured multiple dimensions of school performance that impact student achievements (Riddle & Kober 2012).

Every Student Succeeds Act

In December 2015, ESEA was reauthorized as the Every Student Succeeds Act (ESSA). The statutory requirements under ESSA continue to reinforce the importance of including ELs in state assessments and accountability systems, and retain the requirements from NCLB that states test all students in reading and math in grades 3 through 8 and once in high school. ESSA also requires that states ensure those tests align with states' college and career-ready standards (ESSA, 2015).

ESSA requires states to adopt ELP standards involving speaking, listening, reading, and writing that address different proficiency levels and align with the states' academic standards (ESSA, 2015). The requirement to align ELP and content standards supports the implementation of ELP standards that focus on acquiring academic language to learn and communicate content (science, mathematics, social studies, English literature) in the classroom.

ESSA significantly modified the accountability requirements of the NCLB. Under Title III of the NCLB, local educational agencies (LEAs) and SEAs were required to define criteria for progress in learning English and to establish performance standards for ELP. ESSA moved the accountability for ELP from Title III into Title I, shifting the accountability determinations from the LEA to the school level. Beginning with implementation in the 2017-18 school year, ESSA now requires academic achievement of ELs to be fully integrated into the state-wide accountability system under Title I. SEAs

must establish long-term goals and measures of interim progress for determining the extent of increases in the percentage of ELs making progress in achieving ELP and include an indicator of this progress in statewide accountability systems.

State accountability systems provide the underlying and guiding structure for schools and districts to identify programmatic strengths and weaknesses, to establish priorities, and to design appropriate programs and services to improve student outcomes. The accountability status of schools and districts can also impact the allocation of resources and services to ELs.

English Language Proficiency

While policy dictates that all SEAs administer an ELP assessment to ELs, there is no homogeneous definition of proficiency across and sometimes within states. The confusion encompassing the reclassification of ELs stems partly from the difficulty in answering the question: What makes a person proficient in a language? Does accuracy matter more than fluency? (Thompson, 2012) Researchers view language proficiency from different perspectives. ELP may be defined from a conventional description of listening, speaking, reading, and writing, or viewed with a focus on how language is used in school to create meaning, particularly within the context of specific subject areas. MacSwan and Pray (2005) defined ELP from a linguistic perspective, with the linguistic forms of language interfacing with linguistic functions of language. Other researchers (Cummins, 1984; Hakuta, Butler, & Witt, 2000) offered a more dichotomous distinction between oral fluency and language needed to be successful in an academic environment. Cummins' (1984) seminal distinction used Basic Interpersonal Communicative Skills (BICS) and Cognitive Academic Language Proficiency (CALP) to distinguish between

social language used in everyday, face-to-face interactions and academic language needed in the classroom. However, critics of Cummins' BICS and CALP model argued that his distinction does not acknowledge the levels of complexity of academic language use and development (Bailey, 2006; Scarcella, 2003). According to Wiley (1996), using a simple distinction between these proficiencies ignores the effect of social practices and power relations (for instance between a linguistic minority and majority communities) and may promote a deficit perspective of language use. In the same vein, MacSwan & Rolstad (2006) pointed out that language learned in the home could be abstract and complex. The academic language used in specific subject areas impacts children's language development and proficiency, but does not amount to qualitative differences from a linguistic point of view (MacSwan & Pray, 2005, p. 657). Collier (1995) added sociocultural processes to the discussion of language proficiency. She argued that social and cultural processes occurring in the lives of students, such as immigration status, effects of poverty, and cultural stereotyping, mitigate students' acquisition of a second language in school. Her argument brings focus to the inherent challenges in language development, as well as the time needed to become proficient.

Another challenge in defining language proficiency lies in specifying *for what purposes*. Language functions in the context of a topic, a particular language task performed while interacting with an audience or interlocutors. Bachman (2002) suggested shifting away from the term 'language proficiency' and creating models of 'language use' and 'language ability.' Acknowledging the sociocultural context, he noted that testing methods and the background characteristics of language learners influence scores as much as the students' language skills. Hakuta, et al. (2000, p. 3) recognized the

difference between conversational language and the complexity of the language used for learning in a more formal school setting, noting:

...linguistic competence is complex, and that even the most privileged second language learners take a significant amount of time to attain mastery, especially for the level of language required for school success.

Bailey (2006) offered a more comprehensive definition of *academic English language* to include the language skills that students need to handle the linguistic demands of the content presented in classrooms.

In ESEA, the statutory definition of ELP has a narrower construct. ELP is defined in terms of the language needed in the classroom to successfully learn the academic content areas. Title III of ESSA defines an EL as an individual who: (1) is aged 3 through 21; (2) is enrolled or preparing to enroll in an elementary school or secondary school; (3) meets one of the following criteria—(a) was not born in the United States, or whose native language is a language other than English; (b) is a Native American or Alaska Native, or a native resident of the outlying areas; and comes from an environment where a language other than English has had a significant impact on the individual's level of ELP; (c) is migratory, whose native language is a language other than English, and who comes from an environment where a language other than English is dominant; and (4) has difficulties in speaking, reading, writing, or understanding the English language, that may be sufficient to deny the individual (a) the ability to meet the state's proficiency level of achievement on state assessments; (b) the ability to successfully achieve in classrooms where the language of instruction is English; or (c) the opportunity to participate fully in

society (ESSA 2015). This statutory definition clearly highlights the relationship between ELP and what ELs must demonstrate in knowledge, skills, and abilities in academic content assessed in English. However, this definition does not recognize language differences from a sociolinguistic or socioeconomic perspective (Rolstad, MacSwan, & Gusman, 2014). To ensure the federally guaranteed right to “participate meaningfully” in public school education programs, it is crucial to address the linguistic, academic, and nonacademic needs of ELs across their entire schooling experience (Rolstad, et al., 2015). A student’s initial designation as an EL should be based primarily on linguistic criteria (Linguanti, 2001; Rolstad, et al., 2015). In current practice, the ELP standards place an emphasis on assessing ELP through the lens of academic language development and may misidentify students as ELs (Rolstad, et al., 2015). Summative assessments used to measure progress and attainment of ELP must articulate how academic language is contemplated in the test design (Boals, et al., 2015). However, despite its importance, researchers have not clearly defined or agreed to academic language construct. Rolstad, et al.’s (2015) proposed assessments for reclassification call for inclusion of English literacy, directed by a theory of language structure and acquisition and targets language as it is specifically used in school contexts.

The statutory definition of ELP also implies that an EL designation is a temporary status. Once an EL demonstrates proficiency, the student is *reclassified* from a limited-English-proficient student to a fluent-English-proficient student. Reclassification is an important milestone for an EL student. In practice, states exit or reclassify ELs as no longer needing specialized language supports or regular monitoring through ELP assessments. Proficiency also may impact funding and resources that a school may

receive based on the enrollment counts. Umansky and Reardon (2014, p. 880) observed that many federal and state policies incentivize “rapid and universal” reclassification of ELs. Schools and districts are under pressure to reclassify students as quickly as possible. If students are reclassified too early, they may not be ready to transition into an unsupported academic environment (Hill, Weston, & Hayes, 2014). Conversely, waiting too long can be problematic too. Students with a designated EL status who receive English-language support services, but who do not achieve proficiency to be reclassified after 6 or more years are termed long-term ELs. The limited research available on long-term ELs shows poor academic outcomes for these students (Olson, 2010).

Problem Statement

There is a paucity of research that specifically addresses the question: How long does it take ELs to be reclassified as English proficient? The National Literacy Panel found very little research on the methods used by districts to classify, track, and reclassify ELs (August & Shanahan, 2006). Previous studies about how long it takes ELs to attain ELP produced estimates ranging from 2–3 year for oral language development and 4–10 years for academic language, depending on a number of factors (Collier, & Thomas, 1989; Conger, Hatch, McKinney, Atwell, & Lamb, 2012; Cook, Boals, Wilmes, & Santos, 2008; Cummins, 1981; Hakuta, et al., 2000; MacSwan & Pray, 2005; Shneyderman & Froman, 2012, Umansky & Reardon, 2014). A report published recently by the National Academies of Sciences, Engineering and Medicine (NAS 2017) noted 5-7 years as the most commonly reported estimates. This variability makes it difficult for states to use existing research to determine a timeframe regarding decisions about reclassification, accountability, and distribution of resources. Additionally, every state

has its own criteria for reclassification, making it difficult to compare results across states (Linguanti & Cook, 2013). Mahoney and MacSwan (2005) examined the results of a national survey of state policies on identification and classification of ELs and concluded that state practices sometimes lead to errors in identification and reclassification of ELs, which may in turn have negative consequences for students.

Research shows a positive relation between ELP and academic performance in content areas (Cook, Linguanti, Chinen, & Jung, 2012; Parker, Louie, & O'Dwyer, 2009). The question of how long ELs take to become proficient is important for educators, because the ELs' achievement is intertwined with language proficiency (Greenberg, 2015; Halle, Hair, Wandner, McNamara, & Chien, 2012; Kieffer, 2011; Linguanti & Cook, 2013). ELs who do not achieve ELP struggle to learn grade-level content generally taught in English. Consequently, they take longer and graduate at lower rates than their peers who are English proficient (Callahan, 2013; Gwynne, Pareja, Ehrlich, & Allensworth, 2012; Kim, 2011). Understanding the variables that influence the time span in which ELs develop ELP may provide educators with a measure of expected progress and may help identify students who are at risk of failing academically, because they are not gaining proficiency at the expected rate. Understanding the relationships between reclassification policies, reclassification rates, and reclassified student outcomes is necessary information in making decisions about resources and expectations for ELs. This knowledge may help educators identify programs and practices that facilitate or delay the development of English proficiency (Greenberg, 2015).

Additionally, ESSA requires that states fully implement the accountability provisions in the statute by the 2017-18 school year. States are required to establish long-

term goals and measurements of interim progress to determine increases in the percentage of ELs making progress toward attaining ELP. To do so, state goals must include student-level ELP targets that establish when each EL will make annual progress toward attaining ELP; and attain ELP within a period of time after the student's identification as an EL.

The implementation of ESSA requires states to answer the question, "How long does it take for ELs to become proficient?" To establish meaningful, fair, and reliable targets, SEAs must understand how long it takes ELs to reach proficiency. States also will be required to report the number and percentage of ELs who have not attained ELP within 5 years of initial classification as an EL (ESSA 2015).

Purpose of this Study

This study is guided by asking: (1) How long does it take for ELs to be reclassified as proficient in English based on a state's reclassification criteria? and (2) What are the variables related to reclassification within the demographics of the state?

Answers to these questions will lead to *Reclassification Guidelines* to help states make informed decisions about determining the length of time it may take students to attain proficiency in the accountability system. The information may inform assessment and accountability systems and help establish targets that take specific factors, such as level of English proficiency at entry, grade, or age of entry, or other school or student characteristics into account. Moreover, the Reclassification Guidelines will help educators establish high and realistic expectations and appropriate educational services for each EL. A deeper understanding of the time ELs need to gain English proficiency can help school districts evaluate the effectiveness of their programs for ELs and adjust those programs accordingly.

Finally, the findings of the study will contribute to the growing literature on establishing a timeframe for how many years it takes for ELs to become proficient. This will provide educators with guidelines for gauging progress towards proficiency and identifying students at risk of becoming *long-term* ELs.

State Context of the Study: New Mexico

The data for this study will include standardized ELP scores from New Mexico. Educational policies in New Mexico encourage a supportive environment for the education of ELs. For example, this was the first state in the nation to pass a law and allocate state funds supporting bilingual education in public schools in 1969; appropriate more funds in 1973 for bilingual education by passing a state law regulating multicultural education; and seek endorsement for teaching English as a second language in 1975. The program goals for this state's Bilingual Multicultural Education are for all children, including ELs, to become bilingual and biliterate in English and a second language and meet the states academic content standards. New Mexico is a member of the World Class Instructional Design and Assessment (WIDA) consortium, which consists of 35 states with a common set of ELP standards and assessment.

New Mexico has used Assessing Comprehension and Communication in English State-to-State (ACCESS) to evaluate student and school performance since school year (SY) 2009-10. The ACCESS is a large-scale ELP assessment given to ELs enrolled in kindergarten through twelfth grade to measure progress and attainment of English. New Mexico has not changed assessments or the reclassification criteria since school year 2009-10 and is therefore able to provide longitudinal data matched over several years.

Summary

ELs are one of the fastest growing populations among students enrolled in K–12 schools. Data clearly show a wide academic gap between ELs and their peers. One important factor influencing the academic achievement of ELs is attaining ELP. Understanding the length of time that ELs take to achieve proficiency is important to improving the type and quality of programs and allocation of resources for the education of ELs. There is a dearth of studies to inform the critical question on how long ELs take to achieve ELP. More studies are needed to understand the complexity of this question. This study will add another dimension to the existing research by examining longitudinal state level data using standardized ELP assessments on students across all grades (K–12).

The next chapter is a review of the literature, including previous research with a focus on answering the question about how long it takes ELs to achieve English proficiency and the influence of select demographic and student-level variables on the time and rate at which ELs are reclassified as fluent ELs.

II. Literature Review

The Supreme Court's interpretation of the Civil Rights Act of 1964 in *Lau v. Nichols*, (1974) mandated school districts and states to provide English-language acquisition support to students who were not proficient in English. *Castañeda v. Pickard* (1981) further reinforced this ruling in determining if English-language support services and programs met the civil rights standards and requirements using a three-pronged test: (1) Is the program based on an educational theory recognized as sound by some experts in the field or considered a legitimate experimental strategy? (2) Are the programs and practices (including resources and personnel) reasonably calculated to implement this theory effectively? (3) Does the program succeed in producing results indicating that students' language barriers are being overcome within a reasonable period of time? These court cases and legislation served as a foundation for educators and policymakers to question how long it takes for ELs to be reclassified as proficient and no longer in need of receiving EL support services (Hakuta, Butler, & Witt, 2000). With the 2006 reauthorization of ESEA, this question is front and center in the educational plan for ELs (ESSA, 2015). SEAs are required to develop comprehensive accountability systems with reasonable expectations in the form of short-term and long-term goals on time to reclassification. However, there remains a paucity of research focused on time to attainment of proficiency or reclassification of ELs. Prior to NCLB, there were a handful of studies (Collier, 1987; Cummins, 1981; Hakuta, Butler, & Witt, 2000) that specifically tried to answer the question of time to reclassification. Conger (2008) noted that these studies relied on now-outdated, small samples of students, often in one or two schools or classrooms, and did not include repeat observations over many years. As a result,

Congress asked the Government Accountability Office (GAO) to conduct a study about the length of time needed to achieve proficiency. The GAO study concluded that no clear consensus exists on the length of time children with limited English proficiency need to become proficient in English. GAO also reported that time to proficiency varies from child to child and is affected by such factors as the child's age, socioeconomic background, and amount of formal schooling already received in another language (GAO, 2001). Subsequently, the NAS 2017 report confirmed the conclusions reached by the GAO (2001) study. NAS 2017 also noted that existing research examining the influence of various factors on time to proficiency continues to be limited.

This literature review will focus on studies conducted after 1980, which investigated how long it takes ELs to achieve ELP in order to be reclassified as English proficient in a school setting. The first half of the paper will focus on the general question on how long it takes for ELs to become proficient in English, and the second half of the paper will review literature on selected student-centered variables that impact the time to reclassification.

Length of Time to Achieve English Proficiency for the Purpose of Reclassification

Cummins (1981) and Collier (1987) were precedent-setting studies that examined a specific timeline to proficiency. Cummins (1981) reanalyzed the data from a 1970 study by Ramsey and Wright involving 1,200 immigrant students in the Toronto school system in grades 5, 7, and 9. Cummins (1981) used the term 'length of residence' to indicate the length of time that a student attended school in Canada. He used the term 'age on arrival' to indicate the age that a student began school in Canada. Using the age on student arrival, Cummins (1981) constructed an average length of residence in the country, according to

the grade level of the student. He found that length of residence, not age of arrival, had a more substantial effect on the rate at which students approached grade norms. In this study, the older learners acquired English academic language proficiency more rapidly than younger learners; thus, the age on arrival did not significantly affect the eventual performance at grade norms. Cummins (1981) noted that the English Competence Test (ECT) used by Ramsey and Wright in the original study assessed pronunciation and a limited number of vocabulary items as academic measures. Based on ECT as a measure, Cummins (1981) reported that it took approximately 2–3 years to reach proficiency in communicative skills in English and an average of 5–7 years to achieve native-level proficiency in academic language required for school. Cummins (1981) used the terms BICS and CALP to differentiate two levels of language proficiency.

Collier (1987) conducted a cross-sectional study involving 1,548 immigrant children from affluent middle-class suburban families. These ‘advantaged’ second-language learners were at an age-appropriate grade level in their primary language when they enrolled in schools in the United States, but had been assessed as non-English-proficient when they entered school. The students in this study received 1–3 hours of second-language support in a well-regarded English as a second-language program. Collier studied the students in three age cohorts on the time to reach native-speaker norms on standardized tests (50 on the normal curve-equivalent scale (NCE)) when taught only in English after arrival to the school. Collier found that to approach the 50 NCE, the time varied by cohort, but on average students needed 4–8 years.

Cummings (1981) and Collier (1987) studied the development of language proficiency within the context of academic achievement by exploring how long it takes

students schooled only in English as the second language to reach the average academic achievement level of native speakers. MacSwan and Pray (2005) cautioned against using these results to provide a clear portrayal of how much time is required to learn English. The authors point to the limitations of a test of academic achievement of English to measure an EL's acquisition of various academic registers in English. They noted:

While higher scores on academic achievement tests in English might reflect mastery of some aspects of English-language proficiency in some populations of students, it is not possible to know whether lower scores indicate that the child does not understand the language of the test or simply that he or she does not know the correct answer (p. 661).

Hakuta, Butler, and Witt (2000) analyzed data from four different school districts to study ELP development as a function of time and exposure to English. Two districts were from California; District A with an enrollment of 3,400 ELs, and District B with 7,000 ELs. District A had a relatively lower poverty rate, with 35 percent of students receiving free or reduced-price lunch (FRL) as compared to District B, with 74 percent of students on free or reduced-price lunch. From District A, the sample consisted of 1,872 ELs in grades 1–6 and residing in the district since kindergarten. From District B, the sample consisted of 122 students classified as ELs and residing in the district since kindergarten, randomly selected from grades 1, 3, and 5 for the purpose of this study. District A used the Individuals with Disabilities Act (IDEA) Proficiency Test (IPT) as a measure for ELP and also administered the MacMillan Informal Reading Inventory to inform decisions on reclassification. District B used the Woodcock Language Proficiency Battery (WLPB) as the measure of ELP. Based on these measures, the authors reported that it took 2–5 years

for students to demonstrate oral language proficiency and 4–7 years to develop academic language proficiency. These test results are deceptively simple to interpret, but researchers cautioned against the consequential validity of the IPT and WLPB as measures of ELP (MacSwan & Pray, 2005; MacSwan, Rolstad, & Glass, 2002; Mahoney & MacSwan, 2005).

For the other two districts, Hakuta, et al. (2000) analyzed the data used by a 1970 study by Ramsey and Wright from approximately 1,200 immigrant students in Toronto, Canada, learning English as a second language in grades 5, 7, and 9, and a study conducted by Klesmer (1993) for the North York School District in Ontario, Canada. The sample used by Hakuta, et al. (2000) consisted of 285 randomly selected ESL students and 43 native English-speaking students as controls. All students were 12 years old and had length of residence from 6 to 71 months. The data represented students at fixed-grade levels but differed in the length of residence. The students were given ELP tests and a test on nonverbal ability. For oral language development, the data from these two sources showed a steep growth up to 5 years, but then began to plateau. Overall, the immigrant students came closer to their English-speaking peers in listening comprehension, but even after 5 years remained at .75 standard deviation units below for oral expression. In academic language development, the immigrant students showed gains over the course of the 5 years, but remained at .5 standard deviation units behind their English-speaking peers. However, this study used cross-sectional data, limiting the information to one period of time and not on how students actually perform over time.

In SY 2015–16, California reported that 1.34 million students attended California schools; making it the state with the largest enrollment of ELs in the United States.¹ The five top languages other than English spoken in the state include Spanish, Chinese, Vietnamese, Tagalog, and Arabic. The ethnic, linguistic, and geographic diversity in the state allows for comparisons. Thus, it is no surprise that the education of ELs is an important topic for discussion in California. There are at least five studies using California data, which found wide variations in the reclassification rate among the districts within the state. Hill, Weston, & Hayes (2014, p. 2) noted:

Because districts determine their own reclassification criteria, it is difficult to compare reclassification rates, the progress of ELs, and the outcomes for ELs and reclassified former English proficient students across school districts throughout the state.

In 2000, the California Department of Education commissioned a legislatively mandated evaluation of the effects of Proposition 227 on the education of ELs. The authors used survival analysis and estimated that after 10 years in California schools, the probability of an EL to be reclassified as fluent in English was less than 40 percent. However, the authors cautioned that this pattern varied dramatically across a set of selected districts enrolling large numbers of ELs. The probability of reclassification ranged from an estimated low of 14 percent in one district to a high of 72 percent in another district (Parrish, Perez, Merickel, & Linqanti, 2006).

The Legislative Analyst's Office (LAO) in California is a nonpartisan office that provides policy information and advice to the California Legislature. In 2002, 1.3 million

¹ Source: California Department of Education's fast fact web page, <http://www.cde.ca.gov/ds/sd/cb/cefelfacts.asp>.

ELs in California took the California English Language Development Test (CELDT). The LAO (2004) reviewed student achievement from the 2002 CELDT and, using a simulation technique, its study reported:

After 6 years of public schooling, about half of EL students who attended California schools since kindergarten gained the English and other academic skills needed to be reclassified. During the next 3 years of schooling, an additional 30 percent of the initial group is reclassified (LAO, 2004, p. 16).

Los Angeles Unified School District (LAUSD) enrolls about one-fifth of the state's ELs. An internal evaluation report prepared for the district showed that approximately 60 percent of all ELs enrolled in LAUSD 6 years or more were reclassified as proficient. The state defines students who continue to be classified as ELs for more than 5 years as not making academic progress in a timely manner. Under the state's criteria, the other 40 percent have fallen behind academically since they have not been reclassified in a timely manner (Salazar, 2007).

Grissom (2004a) conducted a study in the aftermath of Proposition 227, contending that the reclassification rates used by Unz² to support the Proposition's passage and counter opponents who criticize its effectiveness are misleading. Grissom estimated time to proficiency using statewide longitudinal data linked over time. The study used matching scores of three cohorts of students from second through fifth grade

² Proposition 227 was crafted by Ron Unz, a Silicon Valley software entrepreneur, and Gloria Mata Tuchman, a Santa Ana teacher. Also called the English Language in Public Schools Statute, it was on the June 2, 1998, statewide primary ballot as an initiated state statute. It was approved. Proposition 227 changed the way that "Limited English Proficient" (LEP) students are taught in California.

and included only students enrolled in school for all 4 years. He found that only 30 percent of ELs were reclassified as proficient after 4 years.

Abedi (2008) also estimated time to proficiency using survival analysis. He used 6 years of longitudinal data for a cohort of students in a large California district to study factors associated with time classified as ELs. All students were in seventh grade and classified as ELs at the start of data collection. Abedi (2008) found that the median time that these students remained as ELs was between 4 and 5 years. Furthermore, the results indicated that, in addition to the students' level of language proficiency, their background variables (such as ethnicity) and prior achievement in reading appeared to correlate with time to reclassification. One limitation of this approach is that the data used in this study started with students who were already in seventh grade and did not account for their prior English proficiency. A student reported as achieving proficiency after 4 years may have enrolled much earlier and may have actually taken a longer time to achieve proficiency.

The ED-funded Regional Educational Laboratories (RELs) published five studies on the reclassification patterns of ELs in different states. RELs provide school districts and SEAs with technical assistance and research to improve education outcomes for students. ELs are reclassified when they meet the SEA or LEA definition of proficiency. The findings differed for each study. These differences may be due to different measurement tools, reclassification criteria, or demographics in each state (Bailey & Kelly, 2013; Linqunti & Cook, 2013).

Motamedi, Singh, & Thompson (2016) examined data on 16,957 ELs in seven districts in Washington State who entered kindergarten between 2005–06 and 2011–12 in

seven cohorts. The districts participated in the Road Map Project. The Road Map Project is a *cradle-to-career* initiative involving seven districts (Auburn, Federal Way, Highline, Kent, Renton, Seattle, and Tukwila) in the Seattle metropolitan area. These districts have high levels of poverty and low levels of academic achievement. Twenty-two percent of ELs in Washington are enrolled in Road Map districts. This study used discrete-time survival analysis to estimate the time it took ELs to be reclassified in the context of their English proficiency at entry to kindergarten, their gender, and their home language. For reclassification in Washington State, students must score at the transitional level, which is the highest proficiency level on the assessment scale. Data from these districts found that 50 percent of the students were reclassified in 3.8 years after beginning kindergarten. One study limitation is that while all districts in Washington use the same reclassification assessments and criteria to achieve consistency, the state changed the ELP assessments three times while this study occurred. Additionally, the researchers acknowledged that this study did not account for differences in instructional programs and practices for ELs in Road Map Project districts and schools. Therefore, the amount of time for reclassification cannot be interpreted as the result of the efficacy of program or policy in the districts or schools.

Kieffer and Parker (2016) studied the patterns of reclassification in New York City public schools. The study also addressed three student characteristics associated with time to reclassification: grade of entry, initial English proficiency, and disability category. The researchers examined longitudinal data on seven cohorts of students in grades kindergarten through grade 7 who entered New York City public schools in each school year between 2003–04 and 2010–11. The researchers excluded students who

entered in SY 2008–09 because of data quality issues. The analytic sample included 229,249 students initially classified as ELs based on their scores on the Language Assessment Battery–Revised, the diagnostic instrument used by New York City. The New York State English as a Second Language Achievement Test (NYSESLAT) is the summative assessment used to measure progress and attainment of ELP. All ELs are required to score ‘proficient’ on the NYSESLAT and in grade 3 and above, students must also obtain a passing score on the content English/language arts exams. Discrete-time survival analyses estimated the probability of reclassification as it changed over time. The study found that 52 percent of students who entered in kindergarten were reclassified as former ELs by the end of their fourth year in school. Seventy-five percent were reclassified after 6 years, and the remaining 25 percent were considered *long-term* ELs. The estimated median time to reclassification varied by grade, from 3 years for students who entered in grade 2 to more than 5 years for those who entered in grades 6 and 7.

Over 6 years, the researchers conducted a series of studies in Arizona, Utah, and Nevada (Haas, Huang, Tran, & Yu, 2016a & 2016b; Haas, Tran, Huang, & Yu, 2015) by following three cohorts of students in kindergarten, grade 3, and grade 6, from SY 2006–07 until SY 2011–12 to identify reclassification patterns. The researchers found that 90 percent were reclassified in Arizona, 59 percent in Utah, and 65 percent in Nevada. Each state used a different assessment tool and criteria for reclassification. Utah administered the Utah Academic Language Proficiency Assessment and stipulated that to be reclassified, students had to achieve an overall score of advanced or above’ plus a test score of partial (level 2 of 4) on the English/language arts content assessment. Nevada developed a state language proficiency test, the English Language Proficiency

Assessment. To achieve reclassification, ELs in Nevada had to score proficient on the overall test and at least advanced intermediate or higher on the four domains of listening, speaking, reading, and writing. Arizona also developed the Arizona English Language Learner Assessment (AZELLA). During this study, to be reclassified, ELs had to meet a composite score of proficient across all four tested domains. The exploratory REL studies are correlational in nature and do not provide any causation. The unique demographics and data in each state did not allow for generalization, but did provide insight into the reclassification patterns and a road map for other states to conduct their own studies.

Methodological Challenges to Determining Reclassification Criteria and Rates

The existing literature discussed above raises methodological and measurement challenges that limit the generalizability of the research findings (NAS, 2017). The variations in the assessments and reclassification criteria used across states make it difficult to synthesize the evidence. The validity and reliability of the ELP instruments used by states to make reclassification decisions has not been established through the US Department of Education's peer review process. Under Section 1111(e) of the ESEA and 34 C.F.R. § 200.2(b)(5), the US Department Of Education has an obligation to conduct peer reviews of state assessment systems. The purpose of the peer review is to ensure that states are meeting statutory and regulatory requirements under Title I of the 1965 ESEA for implementing valid and reliable assessment systems. While the content assessments used by states have been peer reviewed, USDOE has not yet peer reviewed the ELP assessments.

ELs are a diverse group, with different education and social experiences, native cultures, and languages. Interrelated linguistic and sociocultural variables such as ELP,

school program effectiveness, equity, poverty, family background, native culture, and teacher preparedness play a role in understanding the difference in ELP and growth trajectories of ELs (Bailey & Kelly, 2013; Linqanti & Cook, 2013). This diversity offers a valuable opportunity to study ELs' reclassification patterns to help educators make informed policy and instructional decisions (Kieffer & Parker, 2016). To provide a deeper understanding of the differences in the reclassification patterns among ELs across states, it is important to study these outcomes differences by various student characteristics (Haas et al., 2016a; Haas et al., 2016b; Kieffer & Parker, 2016).

Student Characteristics and Reclassification Patterns

Age and Grade of Entry

Age is a major variable in acquiring a second language (Collier, 1987; Conger, 2008; Haas, et al., 2015; MacSwan & Pray, 2005). The impact of age on the ability to become ELP is a question that researchers have asked for a long time, with inconclusive answers (Conger, 2008). Linguistic theorists (Singleton & Ryan, 2004) extensively debate the validity of the Critical Period Hypothesis (CPH). Some CPH theorists suggested that older students will be slower learners and never obtain native-like proficiency in their second language. However, many studies focused on age differences on basic oral language skills, and not the academic language required in formal schooling (Collier, 1987).

Other researchers observed that older children and adults who receive a supportive environment are motivated and invest sufficient time and attention to learn a second language more quickly than younger children (Marinova-Todd, Marshall, & Snow, 2000). Collier (1987) reported that older students between the ages of 8 and 12

have a faster acquisition rate of second-language skills, which they maintain over younger arrivals in the 4–7 age group. MacSwan and Pray (2005) support the view that older children acquire English faster than younger children. Similarly, Haas, et al. (2015) showed that ELs in the Arizona third grade cohort had the highest cumulative reclassification rates (97 percent) compared to ELs in kindergarten with the lowest cumulative reclassification rates (91 percent).

Cummins (1981) reported that older learners acquired English academic language proficiency more rapidly than younger learners, but the effects of age as a variable seemed to diminish after length of residence of 5 years. Six years later, Collier (1987) observed that while short-term studies showed an initial advantage for younger students, long-term studies show that older students between the ages of 8 and 12 have a faster acquisition rate of second-language skills that maintained over younger arrivals in the 4–7 age group. The same study found that the middle cohort of students (ranging from 8 to 11 years old) took 2–5 years to approach the 50 NCE, while students who entered at age 5–7 were 1–3 years behind the 8–11 age group. Students who entered at ages 12–15 took longer than the other age groups and were projected to reach 50 NCE in 6 to 8 years.

In contrast, other studies on the growth trajectories of ELs showed that students in the earlier grades make more year-to-year progress than students enrolled in higher grades (Cook, Boals, Wilmes, & Santos, 2008; Grissom, 2004b; Kieffer, 2008, 2010, 2011; and Salazar, 2007). Combining “maturational constraints and stabilization of language development” (Cook, et al., 2008, p. 7) used 3 years of longitudinal data from three states and 2 years of data from an additional nine states and found that *lower is faster, higher is slower*. That is, younger ELs in the lower grades with lower proficiency

levels acquired language at faster rates than students in higher grades or peers at higher proficiency. This could be attributed to the language skills required at higher grades and at higher proficiency levels. The breadth and depth of academic language ELs are expected to comprehend and produce increases as they advance in grade and in proficiency level. Specifically the language students need to demonstrate in terms of linguistic complexity, forms and conventions, and vocabulary usage is greater and more complex at higher levels of proficiency level (World-Class Instructional Design and Assessment 2013). Conger (2008) examined extant data records on all ELs enrolled in New York City public schools in grades 1 through 8 from SY 1996-97 through SY 2004-05. The data were organized into four cohorts of EL students who entered school from 1996 through 1999. These four cohorts were observed between 3 to 8 years, depending upon the year that they entered the school district and their age upon entry. This study used discrete-time survival analyses, with time as the primary independent variable. The results indicated that the older a student is when entering the school district, the lower their likelihood of achieving ELP. Using discrete-time survival analysis, the coefficients showed that entrants at age 6 have a probability of reaching proficiency that is 2 percentage points lower than age 5 entrants. As age of entry increases by one year, the probability of becoming proficient falls by roughly two to three points. The research posited that the negative effect of age of school entry on the rate at which ELs become proficient is partially explained by the fact that older students tend to enter the school system with lower proficiency levels. The study also concluded that the entry effect remained unaltered by adjustments for students' social and demographic characteristics or the schools they attend (Conger, 2008). Kieffer and Parker (2016) conducted a

longitudinal study of ELs enrolled in New York City public schools that supported the earlier findings by Conger (2008). Kieffer and Parker (2016) reported using median as a unit of measurement that ELs who entered school in grade 6 or 7 took about a year longer than ELs who entered in kindergarten. Conger (2008) noted that the negative effect of age is partially explained by the fact (based on the demographics of students in his study) that older children entered the school system at lower proficiency levels. The language demands of the curriculum and assessment at higher grades were also factors in how soon ELs could achieve proficiency. However, most studies focused on grades K–8 with very little known about students in high school.

Initial Levels of English Proficiency

ELs that start schools with lower proficiency have less time to learn English and may, therefore, take longer to reach proficiency (Hakuta, et al., 2000). The impact of initial proficiency is more evident in acquiring reading skills. ELs who begin school with limited oral ELP skills demonstrated low-level English reading skills in the primary grades (Kieffer, 2011).

However, the impact of initial levels of English proficiency fades out when ELs reach middle school (Kieffer, 2011). Haas, et al. (2015) studied data from Arizona and noted large differences in cumulative passing rates on the ELP test associated with initial ELP level. Arizona has five levels of ELP based on the AZELLA: pre-emergent (level 1), emergent (level 2), basic (level 3), intermediate (level 4), and proficient (level 5). ELs exit EL services and are reclassified when they achieve level 5. Overall, the students in this sample met this progress expectation at different rates according to their cohort and language proficiency level at entry (Haas, et al., 2015, p. 16). However, the progress

slowed towards the end of the study period, showing that ELs in all three cohorts who started the study at level 4 had the highest cumulative rates of reclassification. In the grade 3 and grade 6 cohorts, ELs who started the study at levels 1 and 2 had higher cumulative reclassification rates than students who began the study at level 3. The authors offered two possible speculations for this unexpected finding. First, the number of students receiving special education services may have confounded the results. Both cohorts from grades 3 and 6 had a higher percentage of students eligible for special education services who were at level 3 and level 5, as compared to ELs who were level 1 and level 2. Second, the authors offered the possibility that the level 1 and level 2 students were newer ELs with possibly stronger education and literacy in their native language and, therefore, were able to make progress toward proficiency at a faster rate.

Using data for ELs enrolled in public schools in Nevada, Haas, et al. (2016b) reported that all three grade cohorts of ELs' cumulative reclassification rates followed a steady progression throughout the study period 2006/07–2011/12. Nevada also uses five levels of ELP: entry (level 1), emerging (level 2), intermediate (level 3), advanced intermediate (level 4), and proficient (level 5). In general, this study reported that ELs, who started the study at higher ELP levels, had higher cumulative reclassification rates than their grade-level peers who started at lower ELP levels. The differences were among the grade levels. For example, ELs in the kindergarten cohort starting at level 4 had the highest cumulative reclassification rate (100 percent). ELs in the grade 3 and grade 6 cohorts who started at levels 1 and 2 had final cumulative reclassification rates of less than 50 percent. The lowest cumulative rates were for the grade 6 cohort; for ELs starting

at level 1, the cumulative reclassification rate was 29 percent and for those starting at level 2, the cumulative reclassification rate was 30 percent.

In Utah, Haas, et al. (2016a) found the same pattern as did the students in Nevada. The Utah patterns also showed the cumulative reclassification rate for all three grade cohorts of ELs followed a steady progression throughout the study period. In 2010–11, Utah changed its ELP assessment to the Utah Academic Language Proficiency Assessment and renamed the five proficiency levels to entering (level 1), beginning (level 2), developing (level 3), expanding (level 4), and bridging (level 5). Overall, ELs who were at level 3 and level 4 had the two highest cumulative reclassification rates. However, in the grade 3 and grade 6 cohorts, ELs starting the study at level 1 had cumulative reclassification rates higher than or similar to students who began the study at level 2. This did not occur in the kindergarten cohort. The authors provided two reasons for the difference across ELs with different ELP levels. First, the anomaly may be due to other risk factors, such as eligibility for special education services or free or reduced-price school lunch program. Second, the actual difference in English proficiency between ELs at level 1 and level 2 was much smaller than between any other level, especially in the higher grade levels.

ELs with Disabilities

The Individuals with Disabilities Education Act (IDEA) requires states and LEAs to ensure that a free appropriate public education is available to all eligible children with disabilities residing in the state (IDEA 2004. 34 CFR §§300.101-300.102). The IDEA also requires that all students with disabilities are included in all general state assessment programs, including ELP, with appropriate accommodations as indicated in their

Individualized Education Programs (IDEA 2004, *Sec 612(a)(16)(A)*, 34 CFR §300.160(a); ESSA 2015, *Sec 1111(b)*).

Research on the difference in cumulative reclassification rates of ELs who are also eligible for special education services showed that ELs with disabilities have longer reclassification periods and are more likely to become long-term ELs as compared to their EL peers without disabilities (Haas, et al., 2015). Data from Nevada showed the difference in reclassification rates of ELs with disabilities and their peers ranged from 41 percentage points in the kindergarten cohort to 63 percentage points in grade 3 (Haas, et al., 2016b). Kieffer and Parker (2016) used New York City Public Schools data and also found that ELs with disabilities took longer to be reclassified. Researchers used a discrete-time survival analysis in which the hazard probability of reclassification in each year was freely estimated, thus allowing differences among disability categories. In doing so, the researchers reported the median time to reclassification was approximately 8 years for students with specific learning disabilities, 6 years for students with speech or language impairments, as compared to 3.5 years for ELs without any disabilities. Sixty-three percent of ELs with specific learning disabilities and 46 percent of ELs with speech or language impairments became long-term ELs (Kieffer & Parker, 2016).

Types of Programs and Services for ELs and Reclassification Rates

Factors, such as age of entry to school, home language, and poverty that impact the time to reclassification are not under the control of SEAs and school districts. However, instructional states and school districts design practices and programs for ELs. This is an important factor because schools and districts may influence the reclassification rates by making programmatic changes (Umansky & Reardon, 2014).

Two studies (MacSwan & Pray, 2005; Umansky & Reardon, 2014) specifically focused on the impact of EL programs and services on rate of reclassification.

MacSwan and Pray (2005) studied the rate of language acquisition among ELs enrolled in a bilingual program. Representing a best-case scenario, this study selected an urban elementary school district in central Arizona with a well-designed bilingual program (p. 663). The district used the Bilingual Syntax Measure (BSM) to assess EL language proficiency. The BSM is an oral language proficiency test designed to elicit the use of specific syntactic structures. The BSM measures proficiency in Spanish and English; however, for this study, only the scores on the English version were analyzed. The researchers examined the records of 89 select ELs who, on enrollment in the district (in grades K–3), had a score a 1 (no English) and when tested every 2 years, had a score of 5 or 6 (proficient) on the BSM. Calculating the English acquisition rate as the time elapsed between the first BSM score of 1 to a score of 5 or 6 (proficient), the study reported that the ELs achieved a score of proficient after an average of 3.3 years. English proficiency, as measured by the BSM, was achieved by about 68.5 percent of the children after 4 years and by more than 90 percent after 5 years.

Umansky and Reardon (2014) compared the reclassification patterns of Latino ELs enrolled in four distinct linguistic instructional environments in a large California school district. These four programs included a traditional English immersion program, a Spanish transitional bilingual program, a Spanish maintenance bilingual program, and a Spanish dual-immersion program. The researchers noted that while they did not observe classrooms, they conducted interviews with teachers and administrators regarding the program models. The descriptions for these program models were typical of such models

across the country. The study involved nine cohorts of 5,423 Latino ELs spanning a 12-year period from fall 2000 to spring 2012. Using discrete-time survival analysis, the researchers modeled the association between instructional program and the timing of reclassification. The results showed that the reclassification rates were meaningfully different between the English immersion and two-language programs. In elementary school, a larger proportion of ELs enrolled in English immersion programs were reclassified, as compared to peers enrolled in two-language programs. However, over time, this early advantage disappeared. Cumulative data shows, towards the end of middle school, students in the bilingual programs surpassed those in the English immersion program. Long-term reclassification rates were highest in the dual-immersion and maintenance bilingual programs. However, estimates from the models, including parental choice, indicated that reclassification rates were highest in the transitional bilingual programs. The authors indicated that the models do not fully control for the demographic differences among students who enroll in these programs. The students enrolled in two-language programs surpassed their peers in English immersion programs by middle school. The authors cautioned that they could not measure bilingualism due to the lack of data on the Spanish proficiency tests and, therefore, could not conclude whether the long-term cumulative reclassification rates were higher due to bilingualism, or due to other benefits of two-language programs, such as increased accessibility to the core content. Overall, the results showed that in over the 8-year period, 50 percent of ELs were reclassified. Of the remaining 50 percent of ELs not reclassified, 60 percent became long-term ELs, and approximately 25 percent were never reclassified. In this study, the number of ELs reclassified tended to follow a predictable pattern; more students were

reclassified at the end of fifth, eighth, and eleventh grades. The authors noted that more students may be reclassified at the end of eighth grade, because teachers want to reclassify students as they transition to high school. Teachers of students in dual-immersion and maintenance bilingual programs may have little incentive to reclassify students prior to fifth grade, since students remain in these programs through fifth grade, regardless of their reclassification.

The findings on reclassification patterns among different program models supports research from several studies (August & Shanahan, 2006; Cheung & Slavin, 2012; MacSwan & Pray, 2005) that showed long-term ELs in bilingual classes demonstrate better academic outcomes than ELs in monolingual classes. In the short term, most often in earlier grades, ELs in English-only classes may outperform their peers in bilingual classes (Conger, 2010; Genesee, 2006; Thompson, 2012). The reclassification rates, in the short-term, may be higher in English immersion programs due to the focused attention on ELP. Conversely, English-only programs may thwart learning academic subject matter, which is detrimental to ELs' overall academic success (Umansky & Reardon, 2014).

Poverty

Students receiving Free or Reduced-Price Lunch (FRL) generally are used as a proxy indicator for poverty. The reclassification patterns studied by Haas, et al. (2015, 2016a, 2016b) for ELs in Arizona, Nevada, and Utah showed that ELs who received FRL when the study began generally had lower reclassification rates among kindergarten students, but the difference narrowed after 5 years. Research (Mulligan, Halle, & Kinukawa, 2012; Roberts & Bryant, 2011) showed that ELs from lower socioeconomic

status generally scored lower on academic assessments than their peers. A longitudinal study by Reardon and Galindo (2007) used a nationally representative sample of 21,400 students with data from the Early Childhood Longitudinal Study–Kindergarten Cohort (ECLS-K) to analyze the scores on a mathematics assessment of first-grade students, as well as gains in scores between first and fifth grades. The students were divided into three groups: (a) language-minority students who were not English proficient; (b) language-minority students who had reached proficiency and no longer required additional language support services; and (c) students whose primary language was English. Researchers reported scores by three student characteristics: race/ethnicity, poverty status, and mother’s education. The results initially showed that ELs generally scored below their English-proficient peers whose primary home language in grade 1 was English. However, longitudinally, there was no measurable difference among the three groups in gains in scores between the first and fifth grade. As students gained ELP, the researchers compared their achievement in mathematics to their English-proficient peers, which indicated that ELP was a greater contributor to academic achievement than poverty or mother’s education.

Gender

Collier (1987) first reported cross-sectional data on advantaged ELs, and later Collier and Thomas (1988), added one more year of data and also examined differences by gender. The 1988 study found no significant differences by gender. However, other studies suggested that there are differences in gender-based patterns (Rojas & Iglesias, 2013). Payne & Lynn (2010) studied 32 male and 41 female native English-speaking students enrolled in Spanish as a second-language classes and found that female students

performed slightly better than males in second-language learning, but not in learning the first language. Analyzing National Assessment of Educational Progress (NAEP) mathematics data, several researchers found that overall, boys performed slightly better than girls (McGraw, Lubienski, & Strutchens, 2006; Perie, Moran, & Lutkus, 2005). When analyzing 2004 NAEP data, Perie, et al. (2005) reported that overall female students had higher average reading scores than their male counterparts. Several studies examined differences in second-language acquisition by gender among ELs (Lapayese, Huchting, & Grimalt, 2014; McGraw, Lubienski, & Strutchens, 2006; Perie, Moran, & Lutkus, 2005; Tong, Irby, Yoon, & Masthes 2010). These studies found disparities by gender in achieving ELP and recommended additional research to understand the disparities of underachievement among Latino boys. Recognizing the vital role of biliteracy on academic achievement, Lapayese, et al. (2014) investigated the interaction between gender and achievement of ELs in bilingual programs. The study sample included 55 Latina/o students from four different schools in Southern California. Using a chi-square analysis, the authors found significant differences across all grades between the male and female study participants. Overall, the Latino girls outperformed the Latino boys each year. However, since this was an exploratory study, no conclusions can be drawn; but it does confirm gender disparities in bilingual education.

Although the differences were small, Haas, et al. (2015, 2016a, 2016b) supported other research findings on female students outperforming their male peers. In all three states, female students had higher cumulative reclassification rates than male ELs. However, there were some variations between the states. In Arizona, the difference in the male and female classification rate was five points less for males after 5 years. The

difference in the final cumulative percentages between female and male ELs achieving reclassification was greatest in the kindergarten cohort, at 5 percentage points. For the grade 3 and grade 6 cohorts, the difference was 1 percentage point. In Nevada, after 5 years, the difference in the cumulative classification rates across all grades showed that female ELs had higher cumulative reclassification rates (seven points or less) than male ELs. The difference in the final cumulative reclassification rate was greatest in the kindergarten cohort, at 7 percentage points, and lowest in grade 6, at 1 percentage point. For the grade 3 cohort, the difference was 6 percentage points. In Utah, in the grade 6 cohort, 61 percent of female ELs achieved reclassification, as compared to 58 percent of male ELs—a difference of 3 percentage points. For the grade 3 cohort, the difference was 4 percentage points. The difference in the final cumulative reclassification rate between female and male ELs was greatest in the kindergarten cohort, at 8 percentage points.

In all three states, the final cumulative reclassification rate was greatest in the kindergarten cohort, which does not support other longitudinal studies on the language growth trajectories of ELs in the early language development (Hammer, Lawrence, & Miccio, 2008; Uchikoshi, 2006). Hammer, et al. (2008) conducted a 3-year longitudinal study examining the impact of maternal language on developing vocabulary or early literacy skills among Spanish-speaking children in Head Start and kindergarten, and found that gender was not a significant factor in vocabulary or early literacy growth in Spanish or English. Uchikoshi (2006) found that EL boys in kindergarten had higher initial levels on growth of English receptive and expressive vocabulary skills than girls. In contrast, Rojas & Iglesias (2013) reported that boys and girls demonstrated similar growth trajectories on English oral-language development; however, girls showed a

growth advantage in Spanish oral language development throughout the study. It should be noted that the methodological differences in these studies make it difficult to compare findings or draw conclusions.

Home Language

Linguistic diversity among ELs varies from state to state. In SY 2014–15, the most common home languages spoken by ELs, in order of popularity, included Spanish/Castilian, Chinese, Arabic, Vietnamese, and Haitian/Haitian Creole (ED, 2017). There are only a handful of studies that examine the differences in reclassification rates among various language groups. The main variable in the Conger (2008) study was age of entry; the covariate estimates revealed students whose home language was Russian or Korean were more likely to become proficient than those who spoke other languages at home, while ELs whose home language was Spanish or Haitian were least likely to attain proficiency. One caveat to this finding was that the home language effects were larger when the race/ethnicity indicators were not included in the model; however, the researcher also noted that the race/ethnicity variables did not wipe out the effects of the home language variable.

Motamedi, et al. (2016) reported significant differences in time to reclassification among ELs in Washington State with different home languages. This study found that Chinese (Cantonese and Mandarin combined), Vietnamese, and Russian or Ukrainian speakers were more likely to be reclassified in their first 8 years of school than Somali and Spanish speakers. The study did not find a significant difference between Somali- and Spanish-speaking ELs in their likelihood of reclassification. Other studies also found that Spanish speakers may take longer to be reclassified than speakers of other languages

(Conger, et al., 2012; Slama, 2014; Thompson, 2015). Motamedi, et al. (2016) found that for the first 2 years of school, Spanish speakers were more likely to be reclassified than speakers of other languages. However, the reclassification rate for speakers of other languages caught up to that of Spanish speakers and even surpassed the rate by year 3. The reclassification rate for speakers of other languages surpassed that of Vietnamese speakers in year 5. However, as noted in the study the differences in the reclassification rates across language groups could have been affected by other factors such as socioeconomic status or parent education, could have affected the results. Additionally, the districts studied were not randomly selected; thus, the results may not be generalizable. Motamedi, et al. (2016)

Similar to the findings in Washington state (Motamedi, et al., 2016), the LAO report (Warren, 2004) found ELs in California that spoke Korean and Mandarin as their primary language were likely to be reclassified in less than 5 years. Students who spoke Hmong, Spanish, or Cambodian took longer than speakers of other languages. Hmong speakers were likely to be reclassified in 7.4 years, while Spanish speakers were likely to be reclassified in 6.7 years, followed by Cambodian speakers at 6.4 years. Vietnamese, Armenian, and Pilipino students were projected to be reclassified in 5 years. This report used a simulation model to project the reclassification rates. Overall, the limitations of this report preclude from further conclusions.

Conclusion

The research shows a wide range of time an EL student may achieve proficiency and be reclassified as a fluent English speaker. The range could be due to the differences in EL student classification assessments and criteria used by various states (Bailey &

Kelly, 2013; Linqunti & Cook, 2013). Additionally, differences in reclassification rates may be due to student demographic characteristics, state policies, and practices (Kieffer & Parker, 2016), or the heterogeneity of the EL population (Cook, Boals, & Lundberg, 2011). To address the achievement gap between ELs and their peers, schools must understand that the academic achievement of ELs is tied to developing academic language within socioculturally appropriate environments (Conger, 2008; Cook, Boals, & Lundberg, 2011). Additionally, research on long- term ELs shows students reclassified in upper elementary, middle grades, or who remain in EL status in high school showed larger academic achievement gaps compared to EL peers reclassified in lower grades (Kim, 2011). Thus, it appears that among long-term ELs, the dropout rate may be higher. Time to reclassification is impacted by many factors such as student poverty, special needs, type of program and services, age, and starting proficiency. Results from studies on time to reclassification also show that unique individual factors may influence how much time a student takes to be reclassified. For example, students who are female, never poor, native-born, White, and not receiving special education services for mild or moderate disabilities are more likely to become proficient faster than other ELs (Conger, 2008). Further research using longitudinal data is needed to investigate the factors that impact these differences and the interaction of student characteristics and time to proficiency. Due to the differences in assessments, policies, and demographics, each state may wish to conduct an investigation to make informed decisions on accountability, instruction, and distribution of resources for the academic success of its EL students.

III. Methodology

The reclassification of a student from an English learner (EL) status, implying limited English proficiency to a proficient speaker of English is a high-stakes decision in their educational process. Once reclassified, ELs no longer receive specialized language development support. Some federal and state funds are allocated based on the per-pupil count of ELs enrolled in a school district. A change from an EL to a non-EL status may impact state and federal funding to schools.

This study explores the population of ELs in one state to address the following research questions:

1. How long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED)?
2. Does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools?
3. How do select individual student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time?

Classification and Reclassification Policy for ELs in New Mexico

New Mexico's Bilingual Multicultural Education Bureau (BMEB) published guidance in a Technical Assistance Manual, which stipulates that EL students must be screened upon registration in grades K–12 to determine eligibility for EL services.

Parents or legal guardians of the students complete a Home Language survey to determine the primary home language other than English. If the parent or legal guardian indicate that the primary home language is not English, schools must administer the

English-language placement test (WIDA-ACCESS Placement Test [W-APT]) within 20 days of student enrollment. If the student scores below the cut score assigned for their grade level cluster, the student is classified as an EL and receives appropriate English-language support services. The services should be appropriate for the student's English-language level in order to receive instruction in both English and the student's home language. If student scores at or above the cut score, then the student is not identified as an EL and does not qualify for additional services. Parents receive notifications that their child qualifies for English-language services and have the option to opt-out of these services. However, according to the BMEB Technical Manual, students identified as ELs must take the ELP assessment, ACCESS, until reclassified.

Students identified as ELs are assessed annually for ELP levels and English-language development progress. As noted in the BMEB Technical Manual, ELs are reclassified as proficient when they attain a composite score of 5.0 on the ACCESS. The BMEB Technical Manual stipulates that a composite score of 5.0 relates to proficiency on the New Mexico Standards-Based Assessment. Once ELs are reclassified using the established criteria, they no longer receive the additional English-language support (New Mexico Public Education Department, 2017).

Measurement Instrument: The ACCESS

New Mexico joined the World-Class Instructional Design and Assessment (WIDA) consortium in 2009. Like many other states in the WIDA consortium, New Mexico uses ACCESS to assess ELP. New Mexico administered the ACCESS for the first time in spring of 2010. ACCESS determines ELP for currently enrolled students, and New Mexico uses the W-APT for incoming students designated as ELs. The ACCESS is

a criterion-referenced assessment that measures the development of ELP in grades K–12. The resulting scores on ACCESS support accountability by measuring progress and attainment of ELP. The W-APT is used to determine placement for English-language instruction education program. Both The W-APT and ACCESS align to the WIDA ELP Standards (Cook, 2007). The WIDA ELP Standards include expectations for student performance for all levels of the language development continuum, starting from the beginning to the advanced level of English proficiency (Yanosky, Yen, Louguit, MacGregor, Zhang, & Kenyon, 2011).

The WIDA ELP Standards contextualize academic language proficiency in five language areas: social and instructional language, English/language arts, mathematics, science, and social studies. All standards are clustered by five grade-levels: prekindergarten–kindergarten; 1–2; 3–5; 6–8; and 9–12. Language domains include listening, speaking, reading, and writing across the six proficiency levels. The standard sequence along the continuum of language development into five proficiency levels starts with the lowest at level 1, or ‘entering,’ to full language proficiency needed for academic success at level 5 or ‘bridging.’ The final stage is level 6, ‘reaching,’ which describes the students who progressed through the other five proficiency levels. These proficiency levels provide a global overview of language acquisition and are embedded in the WIDA ELP standards as performance definitions. The performance definitions have three criteria:

1. Increasing comprehension and production of the technical language required for success in the academic content areas.
2. Demonstrating oral interaction or writing of increasing linguistic complexity.

3. Increasing development of phonological, syntactic, and semantic understanding in receptive skills or control in usage in productive language skills. (WIDA, 2012).

The ACCESS test items reflect the WIDA ELP standard indicators. Test items allow students to demonstrate their language proficiency at their grade cluster and proficiency level. Initially, teachers received online professional learning and wrote standards-based language proficiency assessment tasks using the WIDA ELP standard indicators. Experts at the Center of Applied Linguistics (CAL) reviewed and refined the items, after which a panel of educators from the WIDA consortium states conducted a review of the items for content and linguistic bias before field-testing. ACCESS was originally field-tested in 2004; however, each year, WIDA replaces one-third to one-half of all items. The intent is to continue to replace all items in each test form over a 3-year period to avoid overexposure to items by students taking the test within the same grade span (WIDA, 2012). ELs constitute a heterogeneous group including all ethnicities, the majority being Hispanic. It is important for the test developers of ACCESS to conduct a differential item functioning (DIF) analysis to ensure that a student taking the test is not biased because of their gender or ethnicity. Based on the published technical reports (WIDA 2012; WIDA 2014), two phases of analysis for DIF are conducted on the operational form while operational testing is ongoing. Each item is categorized into three levels of DIF: A, B, or C. An item exhibiting 'A level' DIF shows little or no bias toward a particular group, and an item exhibiting 'C level' DIF is considered to display bias is closely examined by test developers. During the first phase analysis, only ethnicity DIF (Hispanic vs. Non-Hispanic) is investigated. During Phase II analysis, ethnicity and gender DIF are investigated. As with Phase I, for items that show high levels of DIF a team of content

experts investigate the items to determine if any construct-irrelevant factors may contribute to DIF. In terms of DIF by ethnicity (Hispanics vs. Non-Hispanics), special attention is paid to the presence of Spanish-English cognates or false cognates that may affect student performance. That information is provided to the test development team, which makes necessary revisions to the items and keeps a record of such cognates for future reference. The test development team also uses this information to guide the item development and review process for future items. The complete results of the DIF analysis is published in the annual technical reports. (WIDA 2012; WIDA 2014)

The ACCESS test forms consist of five grade-level clusters, divided into three overlapping tiers within each cluster: A (Beginning), B (Intermediate), and C (Advanced) to represent the entire range of ELP (Yanosky, et al., 2011). Listening and reading are assessed using multiple-choice questions that are machine-scored by an outside vendor contracted by WIDA. For grades 1–12, speaking is assessed through scripted face-to-face interviews that allow students to demonstrate proficiency at the different WIDA language proficiency levels. Speaking is scored locally by the test administrator using the “Speaking Rubric.” For writing in grades 1–12, students may receive three or four group-administered tasks. Trained raters at the vendor site use the “Writing Rubric” to score the written responses (Yanosky, et. al., 2011, p. 5).

ELs with IEPs or 504 plans may use accommodations on ACCESS as specified in their IEPs. New Mexico published extensive guidelines applicable to the selection of all assessment accommodations for educators in the *Student Assessment Accommodations Manual*, which is updated periodically (NMPED, 2011a; NMPED, 2013). The accommodations allow ELs to offset challenges caused by a disability and to demonstrate

his or her English-language skills on ACCESS. Some accommodations are allowed for all students on all assessments, such as, additional time between sessions, multiple and frequent breaks, preferential seating, test in location with minimal distraction, visual, verbal, or tactile reminders to stay on task. Other accommodations are selected on a case-by-case basis based on factors such as, a student's current proficiency level, disability, age, grade, and experience with accommodations under consideration. Based on the IEP, students may receive accommodations such as use of assistive technology, use of recording devices, extended time, and large print text. According to State policy, accommodations should not compromise the validity of ACCESS for ELLs as an assessment of ELP (NMPED, 2011a). New Mexico began implementing the *Alternate ACCESS for ELLs* in SY 2012. This assessment is for ELs with significant cognitive disability. The *Alternate ACCESS for ELLs* uses a different scale than the ACCESS. Students taking the alternate ELP assessment were not included in this study.

Assessment Scores

ACCESS reports individual student's scores three ways: raw scores, scale scores, and ELP levels. Raw scores indicate the actual number of items or questions the student answered correctly out of the total number of items or questions. Raw scores are converted to corresponding scale scores, using Item Response Theory (IRT) methods, a process to adjust for differences in the difficulty of the questions on the various forms of the test. As a result, the scale scores allow the results to be reported on a standard scale that adjusts for the developmental growth. Scale scores measure a student's progress over time within a language domain. Scale scores and proficiency levels are reported for the four language domains (listening, speaking, reading, and writing) and four different

combinations of language domains. These combinations include: oral language (listening and speaking), literacy (reading and writing), comprehension (listening and reading); and overall or composite score (a combination of all four language domains) (Yanosky, et al., 2011). The composite score, based on scale scores, reflects all four domains and is weighted as follows: Listening (15%), Speaking (15%), Reading (35%), and Writing (35%). “The weighting of the scores reflects the differential contributions of each language domain required for academic success, with heavier emphasis placed on literacy development” (Yanosky, et al., 2011, p. 9). Proficiency-level scores are grade- and domain-specific interpretations of scale scores and describe the student’s performance on the six WIDA proficiency levels.

Validity and Reliability of ACCESS

The construct, content, and consequential validity of ACCESS is built from a theoretical base, WIDA’s ELP standards, a common ground for curriculum, instruction, and assessment (WIDA, 2012). The test was piloted and field-tested on over 10,000 students, including diverse ELs and proficient English speakers, across the WIDA consortium states. Additionally, for a teacher to administer the test, high inter-rater reliability is required as part of the online training for the speaking section (WIDA, 2012).

A stratified Cronbach’s alpha coefficient is reported as a reliability measure for composite scores on the ACCESS. Reliability of the overall composite scores is very high across all grade-level clusters, as seen in Table 1. Validity of ACCESS is based on the sound technical properties of the assessment as outlined in the technical manual published by WIDA. The test items align to the WIDA standards and evaluate progress

and attainment of ELP of ELs in various stages of language proficiency and placement decisions for students in the proficiency levels defined by the WIDA standards.

Table 1: Reliability of ACCESS scores

Grades	Reliability Scores	Accuracy of decisions about students placement across proficiency level 5 and 6	Accuracy of decisions about cut scores between proficiency level 2 and 3
Kindergarten	.930	Students in KG cannot receive a composite score at level 6	.949
1–2	.949	.975	.943
3–5	.941	.972	.940
6–8	.933	.976	.936
9–12	.936	.977	.921

(WIDA, 2012)

WIDA developers conducted field tests to demonstrate the relationship between student performances. The ACCESS Technical Report contains details of the mean and distribution of ACCESS scale score obtained by students according to their a priori proficiency-level assignment, by grade level cluster (WIDA, 2012, p. 20). Field test results also show that: (a) the items are empirically ordered by difficulty across the five proficiency levels across all domains; (b) the test measures the intended language skills need for academic success; and (c) the test is a valid interpretation of a student's performance on the WIDA standards.

Concurrent validity is established through correlation studies conducted between the ACCESS and other older generation tests. The results showed moderate to high correlation between ACCESS and other ELP assessments such as the Language Assessment Scales, the IDEA Proficiency Test, the Language Proficiency Test Series, and the Revised Maculaitis II (WIDA, 2012, p. 24).

Description of the Data Set

The ACCESS is administered annually between January and February to all ELs across all school districts in the state (NMPED, 2011b). Each district compiles and

forwards the answer sheets to scoring vendors under state contract. The vendor scores the test and sends the test scores to NMPED. The NMPED, using the test scores sent by the vendor, determines the accountability targets and measures for each district, school, and individual student. For this study, achievement data from the ACCESS test score file is compiled at NMPED. New Mexico also uses the Student Teacher Accountability Reporting System (STARS), a comprehensive student, staff, and course information system that provides a standard data set for each student served by New Mexico's public education system. Districts and schools compile student level data and send it to NMPED for aggregation at the state level. Districts and charter schools must use Standardized Reporting dates and templates to submit the required data for each reporting period. The purpose of STARS includes providing a longitudinal data system of student progress and educational history over time. The state provided 6 years of longitudinal data on students in kindergarten through twelfth grade, starting with the first administration of ACCESS in SY 2010-11 until SY 2015-16. The data set includes the population of students from the state's ACCESS and STARS databases.

Building the Analytical Sample

To model change, the longitudinal data set must describe how each student in the sample changes over time; and to posit more flexible models with less restrictive assumptions such as non-linear growth, more than three waves of data are recommended (Singer & Willett, 2003).

The data files received from the State were in multiple spreadsheets and included the end year proficiency levels from the ACCESS dataset and demographic data from the STARS data base. There was a separate file from each database for every year that the

data were available. The first step was to merge all the files to map longitudinal data for each student by linking the student identification numbers from each file. Student ACCESS scores were followed year by year until the last year that the data were available for each student. At that point, it is assumed that the student has left the school system or become proficient and is no longer considered an EL student. The EXCEL files received from the State with data from the STARS database was similarly merged into one file. This resulted in two merged data files; one from the ACCESS data set and one from the STARS data set. The two data files were formatted as a student-level data file in which each student had one record with multiple variables. Descriptive statistics of the population data were used to make a systematic determination of students to include in the final analysis. Table 2 shows the reason and the number of students excluded from the population data set to build the final analytical sample.

Table 2: Percentage of students grades K -12 completing WIDA ACCESS testing for English proficiency by exclusions and school year

	SY 2010-11 53,599 students tested			SY 2011-12 53,789 students tested		
Reason for exclusion¹	Number of students excluded	Percentage excluded from total tested	Population after exclusion²	Number of students excluded	Percentage excluded from total tested	Population after exclusion²
Invalid or missing a means of student identification from ACCESS data file	30	0.06	53,569	21	0.04	53,768
Duplicate student information in ACCESS data file	119	0.22	53,450	108	0.20	53,660
Missing composite overall proficiency level from ACCESS data file	962	1.79	52,488	652	1.21	53,008
Composite overall proficiency levels recorded for previous years from ACCESS data file	37,375	69.73	15,113	40,699	75.66	12,309
Length of time in LEP recorded when tested was more than 0 from STARS data file	941	1.76	14,172	651	1.21	11,658

	SY 2010-11 53,599 students tested			SY 2011-12 53,789 students tested		
Reason for exclusion ¹	Number of students excluded	Percentage excluded from total tested	Population after exclusion ²	Number of students excluded	Percentage excluded from total tested	Population after exclusion ²
Missing 2 or more years of proficiency scores between their first and last proficiency scores from STARS file	500	0.93	13,672	229	0.43	11,429
Missing or duplicate student IDs from STARS file	790	1.47	12,882	69	0.13	11,360
Student has been consecutively enrolled in US public schools for 3 or more years from STARS file	3,933	7.34	8,949	2,440	4.54	8,920
English proficiency code indicates student has exited ELL after 1 or more years from STARS file	55	0.10	8,894	10	0.02	8,910
	Final sample population 8,894			Final sample population 8,910		

¹ Student records excluded in order of appearance in this table.

² This running calculation represents the sample population without the current exclusion and the students listed in the previous cell.

ACCESS data set

ACCESS was first administered in New Mexico during the 2009-10 school year (SY). The ACCESS data files contained the end-of-year proficiency levels from SY 2009-10 to SY 2015-16. Student records in the SY 2009-10 data file were missing the students' dates of entry based on presence of previous years of proficiency scores. Therefore, data from SY 2009-10 were not included in the sample. SY 2010-11 was used as the starting year for analysis in this study. The analytical data sample was limited to two cohorts. Cohort 1 consists of students who took the ACCESS test for the first time in SY 2010-11, and Cohort 2 consists of students who took the test for the first time in SY 2011-12. Cohort 1 is limited to students who took the ACCESS for the first time in SY 2010-11 and had scores available through the end of the data-gathering period. Cohort 2 is limited to students who took the test for the first time in SY 2011-12 and had scores

available until end of the data-gathering period. Limiting the data to these two cohorts assures a minimum of five waves of data for students who took the test for the first time in SY 2010-11 and SY 2011-12 until the end of the data-gathering period. The data set showed a total of 53,599 ELs tested statewide in SY 2010-11, and 53,789 ELs tested statewide in SY 2011-12.

To prepare the analytical data sample a systematic review of missing data was conducted. The first category was students with missing identification numbers in the ACCESS data set. Identification numbers are needed to match their test information across the years. This was a small number; only 30 students (.06 percent) in SY 2010-11 and 21 students (.04 percent) in SY 2011-12 were missing identification numbers. These records with missing identification numbers were excluded from the analytical sample. It was not possible to match the records to build a complete longitudinal history through the data-gathering period without the identification numbers. The second category was students with duplicate records. These students had two testing records assigned to the same identification number, but each testing record had a different score. Of the total sample, there were 119 students (0.22 percent) in SY 2010-11 and 108 students (0.20 percent) in SY 2011-12. Since there was no way to determine which score was correct, these records were not included in the final analytical data set. The third category was students' with missing outcome variables, which excluded 962 students (1.79 percent) in SY 2010-11 and 652 (1.21 percent) in SY 2011-12 of the total sample. These records could not be used in the analytical file because there was no composite overall proficiency levels recorded for these students. The next category resulted in the largest number of students that had to be excluded from the analytical sample. Out of the total

population, 37,375 (69.73 percent) in SY 2010-11 and 40,699 (75.66 percent) in SY 2011-12 had a previous ACCESS test score. To be included in each cohort, the student required a first appearance in the given cohort year. Since these students had a previous ACCESS test score, they were not starting in the SY 2010-11 or SY 2011-12 cohort year and therefore did not fit into the sample category for students who were in the first year of enrollment.

STARS data set

The next step was to examine the records from the merged STARS data file. The STARS data identified a category of students with the length of time for enrollment at the time of testing to be more than zero. This meant that these students enrolled in a New Mexico school and who were not starting in the SY 2010-11 or SY 2011-12 cohort year. These students were also excluded as they also did not fit into the sample category for students who were in the first year of enrollment. A total of 941 (1.76 percent) in SY 2010-11 and 651 (1.21 percent) in SY 2011-12 of the total population was excluded from this category. The next category included students with missing scores for more than two years across the length of the observation period. For example, if a student had a test score in SY 2010-11 but was missing test scores for SY 2011-12 and SY 2012-13, 2 years in a row, the student was dropped from the analysis. However, if a student had test scores for SY 2010-11, was missing a test score for 2011-12 but reappeared in SY 2012-13 with a test score, then the student was included. Students missing 2 years of test scores in a row and then reappearing in the data were also dropped from the analysis. Consecutive scores from year to year are needed to assess which year an EL student meets the reclassification criteria. If a student is missing more than two years of data, it is

not possible to determine the reason whether the student just dropped out of school, moved to a different state or was no longer coded as an EL for some other reason. Students who did not have scores for two years but re-appeared in the data set in the third year were kept in the final analytical data set. The result identified 500 (0.93 percent) students in SY 2010-11 and 229 (0.43 percent) students in SY 2011-12 who fit this category and could not be included. The STARS data files also had duplicate entries assigned to the same identification number. Duplicate entries make it difficult to determine which record actually represents the correct information resulting in 790 (1.47 percent) students from the total number in SY 2010-11 and 69 (0.13 percent) from the total number tested in SY 2011-12 to be excluded. The STARS data file also identified 3,933 (7.34 percent) students in SY 2010-11 and 2,440 (4.54 percent) students in SY 2011-12 as enrolling in schools for more than 3 years prior to the start of the cohort years. These were excluded since the selected sample category set includes only students in the first year of enrollment. Finally, students who exited the program at the start of the data collection period were excluded in the analytical sample. There were 55 (0.10 percent) students in SY 2010-11 and 10 (0.02 percent) in SY 2011-12 who had exited the program in the same year that the data collection period started.

The data from the ACCESS and STARS data files were merged into one file to obtain the analytical data set. From a population of 53,599 students tested in SY 2010-11, the final analytical sample consisted of 8,894 students in Cohort 1. The Cohort 2 analytical sample consisted of 8,910 from a population of 53,789 students tested in SY 2011-12. The exclusion of so many students may result in an underestimate of the years

to proficiency and will be a caveat in interpreting the results. The students in Cohort 1 have six waves of data, while the students in Cohort 2 have five waves of data.

The number of years a student was classified as an EL was determined by a proficiency level score of 4 or lower on the ACCESS given at the end of the school year. Proficiency is indicated by a score of 5 or above on the ACCESS. Students' scores in the analytical sample are followed year by year until the last year of data available, or until the student drops out of the data because of graduation, moving out of state, or obtains a score of 5 or above on the ACCESS. Demographic characteristics were considered static and taken from the base year in the analysis (SY 2010-11 or SY 2011-12); and remained unchanged as students were tracked over the length of the data collection period. These characteristics include gender, race/ethnicity, free and reduced lunch (poverty) status and Special Education status. The analytical data set was organized in a panel structure. No additional students were added. The data set was arranged in a person-person format as described by Singer and Willet (2003) where each student has one row of data for each year that he or she remained an EL.

Variables

Table 4 displays a summary of the variables in the sample population considered for the final model estimates. Descriptive data about the variables available in the State demographic data file was explored to determine viable independent variables. The first variable was gender. The sample population data showed a fairly equal distribution of students by gender, with slightly more males than females. The second variable, grades, was clustered as represented in the ACCESS assessment data file. The frequency count

displayed in Table 4 shows a majority of ELs enrolled in the elementary school grades; however there were enough students in each grade cluster to use K–12 as a covariate.

The third variable considered is race and ethnicity. The procedures for collecting racial and ethnic data involve a self-identification process in which parents or guardians identify the race and ethnicity of their children from a given list of categories during the registration process. The racial distribution showed that three-quarters of ELs identified themselves as Hispanic (78.15% in Cohort 1 and 74.48% in Cohort 2). American Indian/Alaska Native was the second largest ethnic group (15.91% in Cohort 1 and 18.33% in Cohort 2). Additionally, a majority of the ELs identified themselves as Caucasian (81.12 % in Cohort 1 and 78.27 in Cohort 2). Other ethnic groups had relatively small numbers, Asian (2.09% in Cohort 1 and 2.22% in Cohort 2), Black or African American (0.71% in Cohort 1 and 0.79% in Cohort 2), Native Hawaiian or Pacific Islander (0.17% in Cohort 1 and 0.39% in Cohort 2).

For the analysis, the percentage of students receiving free or reduced price lunch is used as a proxy measure for the percentage of students living in poverty. However, it is important to point out that while students receiving free or reduced price lunch can provide some information about relative poverty, it may not be the actual percentage of students in poverty enrolled in school (Harwell & LeBeau 2010). There are multiple ways that a student can become eligible for a free/reduced price lunch. Traditionally, family income is used to establish eligibility for free/reduced price lunch; however, some groups of children such as foster children, children participating in Head Start and Migrant Education Programs, or children receiving services under the Runaway and Homeless Youth Act are eligible for free/reduced price lunch. Additionally, under the Community

Eligibility option, non-poor children may be included if the school district decides that it would be more efficient from an administrative or service delivery perspective to provide the free lunches to all children in the school (Hoffman 2012). This count is available at the school level and the free, and reduced price lunch eligibility is derived from the federal poverty level, and therefore highly related. It is a useful proxy for poverty level from an analytic perspective. In this data set nearly 90 percent of ELs participated in the free and reduced lunch program and assumed to be experiencing the impact of poverty.

Home language was also available as a variable in the data set. Data on language spoken at home is collected at the time of registration and self-reported by parents. In the analytical data set, ELs in Cohort 1 speak 23 different languages and in Cohort 2 the students speak 24 different languages. A majority of parents of ELs reported Spanish (69.19% in Cohort 1 and 68.47% in Cohort 2) as their home language. English was reported as the second most common language spoken (13.12% in Cohort 1 and 11.12% in Cohort 2). English as the second most common language spoken by ELs appeared to be an anomaly. The statutory definition of Title III (ESSA 2015) defines EL as an individual whose native language is a language other than English and who comes from an environment where a language other than English is dominant. This implies that a student who speaks English at home would not fit the definition of an EL. A cross tabulation of ELs in grades K-8 who identified the home language to be English and ethnicity (Hispanic and non-Hispanic) revealed that a majority of Hispanic ELs identified their home language as English. It could be that English is spoken in the household and therefore parents identified the home language to be English or it could an error in the data collection at the school level.

Table 3: Number of students tested who listed English as home language by ethnicity, grades K-8

	SY 2010-11	SY 2011-12	Total Sample
Total (English Home language)	1148	948	2096
Hispanic/ Latino	881	691	1572
Not Hispanic / Latino	267	257	524

American Indian is the second largest community in New Mexico but individual American Indian languages had fewer students when taken separately. To have a sufficient number for analysis, the nine languages represented in this sample would have to be clustered together to create one group called *American Indian/Alaska Native*. An estimate of so many combined languages may not be useful because each language group contributes to the estimates in different ways.

Even though cross linguistic influence is well documented in the field of linguistics, home language as a variable could not be used in this study for two reasons: first the error in data collection showing English as the second largest language spoken by ELs in the State and second the small numbers representing the unique languages of the second largest ethnic group. Therefore no further analysis was conducted using home language as a variable.

The final variable included in this study was the number of ELs with disabilities. Educators follow the guidelines published by NMPED on the process of identifying ELs with disabilities. The data showed enough students (10.23% in Cohort 1 and 9.82% in Cohort 2) in this category to allow for an analysis on how long it was taking EL students with disabilities to become proficient. The numbers of students served by programs for immigrant, homeless, migrant and gifted and talented students were not included as covariates in the final analysis. As shown in Table 4, less than 1% of ELs were identified as migrant or gifted and less than 7% were identified as immigrant or homeless in both cohorts. These variables were not included because the frequency distributions appear

skewed and may not yield significant variance to allow for any meaningful analysis. The data set also does not allow examination of reclassification outcomes by program of instruction because of large number of cells with missing data for this variable. Thus, the predictors used in the final analysis include: grade cluster, gender, race/ethnicity, poverty status indicated by students receiving free or reduced lunch and students receiving special education services.

Table 4: Percentage of students in the sample population by selected student characteristics

Student characteristic	SY 2010-11 (Cohort 1)		SY 2011-12 (Cohort 2)	
	Number of students	Percent	Number of students	Percent
All students ¹	8,894	100.00	8,910	100.00
Gender				
Female	4,251	47.80	4,258	47.79
Male	4,643	52.20	4,652	52.21
Cluster (Grades)				
Kindergarten	5,128	57.66	5,788	64.96
1st–2nd	1,613	18.14	1,243	13.95
3rd–5th	1,092	12.28	862	9.67
6th–8th	454	5.10	470	5.27
9th–12th	607	6.82	547	6.14
Race/Ethnicity				
American Indian/Alaskan Native	1,415	15.91	1,633	18.33
Asian	186	2.09	198	2.22
Black or African American	63	0.71	70	0.79
Caucasian	7,215	81.12	6,974	78.27
Native Hawaiian or Other Pacific Islander	15	0.17	35	0.39
Hispanic or Latino				
No	1,943	21.85	2,274	25.52
Yes	6,951	78.15	6,636	74.48
Free or Reduced-Price Lunch				
Free	7,677	86.32	7,818	87.74
Not Participating	945	10.63	812	9.11
Reduced	272	3.06	280	3.14
Home Language				
English	1,167	13.12	991	11.12
Spanish	6,154	69.19	6,101	68.47
Vietnamese	54	0.61	57	0.64
Cantonese	11	0.12	#	--
Cambodian	#	--	#	--
Korean	12	0.13	15	0.17
Laotian	#	--	#	--
Navajo	1,064	11.96	1,131	12.69
Tagalog	8	0.09	14	0.16
Russian	12	0.13	13	0.15

	SY 2010-11 (Cohort 1)		SY 2011-12 (Cohort 2)	
Student characteristic	Number of students	Percent	Number of students	Percent
Creole (French)	7	0.08	#	--
Arabic	38	0.43	53	0.59
Portuguese	#	--	6	0.07
Japanese	7	0.08	18	0.20
Other	176	1.98	223	2.50
Tiwa	#	--	6	0.07
Tewa	12	0.13	18	0.20
Towa	9	0.10	19	0.21
Keres	106	1.19	150	1.68
Jicanilla Apache	#	--	#	--
Mescalero Apache	0	0.00	#	--
Zuni	43	0.48	71	0.80
American Sign Language	#	--	9	0.10
Homeless Child or Youth				
Eligible and not served	28	0.31	37	0.42
Eligible and served	190	2.14	275	3.09
Not homeless	8,676	97.55	8,598	96.50
Migrant				
No	8,865	99.67	8,870	99.55
Yes	29	0.33	40	0.45
Gifted				
No	8,877	99.81	8,898	99.87
Yes	17	0.19	12	0.13
Special Education				
No	7,984	89.77	8,035	90.18
Yes	910	10.23	875	9.82
Immigrant				
No	8,222	92.44	8,355	93.77
Yes	672	7.56	555	6.23

¹ Please see the sample population table for a list of students excluded from the student demographic table.

Data are censored since there are less than five students in the category.

-- Percentage is not calculated because data are censored.

Analytical Approach

Discrete-time survival analysis methods will be used to estimate the probability for students reclassified as English proficient who started in EL programs in SY 2010-11 and SY 2011-12 (Singer & Willett, 2003). Survival analysis allows for data censorship and the data on all students in the sample population to be included. Students will remain in the data set through the year they are either censored or are reclassified as proficient. The *event* is the point where a student is reclassified based on the state's exit criteria.

Data Censoring

The validity of survival analysis rests on the assumption that censoring is non-informative, either because it occurs at random or it occurs at a time dictated by design (Singer & Willett, 2003). In this study, there are two ways that data are censored. The first way denote students who have experienced the *event* (reclassified) before the start of the data collection period and are not included in the analytical sample. Students contribute to estimates of time to reclassification only for the first time that they are present in the data set. Therefore by definition students who were reclassified before the start of the data collection and have no information to contribute to the estimates. These are known as left censored data. The analytical sample is limited to newcomer ELs starting school in SY 2010-11 or SY 2011-12. The data collection ends in SY 2015-16. All students included in the analytical data set are identified as current ELs who have not yet been reclassified. Students who experience the event during the observed period will not be censored, because they contribute to the model in all time periods up to their experience of the *event*. The second set is right censored data. The analytical sample is grade-heterogeneous. ELs begin in any grade K-12 at the start of the observation period and end in different grades at the close of the observation period. Since the outcome measure is the length of time it takes students to be reclassified or achieve a 5.0 on the ACCESS, *Time* is clocked chronologically in school years. All students who enter the study will be followed until the observation period ends. No new students will be added to the cohorts. However, students who are not reclassified during the observation period and do not experience the *event* are right censored. These students will continue to take the ACCESS and contribute to the EL progress and attainment rate beyond the time that

the data observation ends for this study. To examine possible bias in estimates due to right censorship, descriptive data from students who are classified as right censored student are examined. A comparison of the students who experienced the event and those who are right-censored students yield important information on the reclassification patterns. For example, it may shed light on the characteristics of ELs that may take longer to be reclassified than their peers. A detail description and analysis is presented in the next chapter.

Survival Analysis: Description of the Model

In statistics, the generalized linear model (GLM) is a flexible generalization of linear regression. In the classical regression framework, the interest is in modeling a continuous response variable y as a function of one or more predictor variables. However, in this study the measured outcome of interest binary (either a success or failure), which can be coded as a 1 or a 0. Discrete survival analysis, another form of GLM allows for quantifying time to event data. Survival analysis involves developing probability models based on observed rates of *survival* and *failure* in a data set. *Survival* is a term used to indicate those students who have not experienced the *event* and remain as ELs in the analytical data set. Its counterpart, *failure*, is a term denoting students who experience the *event*, exit EL status, and are no longer included in the data set. The ratio of the *survival* and *failure* rate produces what is known as a *hazard rate*, defined in this study as the likelihood that a student will be reclassified. Probability estimates are based on the number of students who enter each successive time period of the *risk set*. The *risk set* is the pool of students eligible to experience the event during the time interval. An essential feature of the *risk set* is that it is irreversible, once an individual experiences the event or

is censored in one time-period; the student is dropped out of the *risk set* for future time periods. The *risk set* allows the analysis of the event occurrence among the members of each time period's *risk set*, yet generalize results back to the entire population (Singer & Willet, 2003).

There are two quantitative functions of primary interest in survival analysis: (1) the *survivor function* and (2) the *hazard function*. The hazard function denoted by $h(t_{ij})$ is the conditional probability that an individual student (subscript i) will experience the *event* (i.e., experience reclassification as English-proficient, during a particular school year (time period j) – given that the student did not experience the *event* earlier). Each individual in the population has its own discrete-time hazard function. T_i represents the time of the event occurrence, and j represents the current time period. The hazard function is defined as

$$h(t_{ij}) = \Pr[T_i = j \mid T_i \geq j].$$

The *survivor function* provides a way to describe the distribution of event occurrences over time. It represents the probability that some students will not experience the event, even after the observed range of time. The *survivor function* denoted by $S(t_{ij})$ is defined as the probability that student i will survive past time period j given that individual i will not experience the target event in the j th time period or any earlier time period

$$S(t_{ij}) = \Pr[T_i > j].$$

The probabilities computed from the hazard and survivor functions will always sum to 1 in a given time period. Therefore, the *survivor function* can also be expressed as $S(t_{ij}) = 1 - h(t_{ij})$. The odds of experiencing the event of reclassification as English-proficient can be expressed as the ratio of these two probabilities

$$\frac{h(t_{ij})}{1 - h(t_{ij})}.$$

According to Allison (1999), there are three common ways of estimating the coefficients in a survival analysis: (1) ordinary least squares (OLS), (2) weighted least squares (WLS), and (3) maximum likelihood (ML). Since the dependent variables used in this study are categorical (dichotomously scored), ML was preferred. In a similar fashion to logistic regression, a typical way to express the conditional probability that student i experienced the event of being reclassified in time period j is through the logit link function. The baseline hazard model is displayed in the equation below

$$\text{Logit } h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij},$$

where D_1 through D_J represents the series of time indicators for each year that the student was observed through J number of years. This model includes only the main effects of *Time*. The results of the fitted model will be presented in terms of odds ratios and a cumulative failure (reclassification) rate over time, a model-based prediction upon convergence of the estimation algorithm.

Once the basic model is determined, covariates will be added to the hazard function to complete the survival analysis model. The covariates are summarized in Table 4 gender, race/ethnicity, poverty status and special education status. Using the time indicators as well as the selected predictors the full fitted model is represented as

$$\text{Logit } h(t_{ij}) = [\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}] + [\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_P X_{Pij}].$$

The right-hand side of the model is composed of two sets of terms. The first set of terms include α 's representing the baseline logit hazard function when the predictors are equal to zero. Each intercept parameter $\alpha_1, \alpha_2, \alpha_J$ represents the log odds of event occurrence in

that particular time period for individuals in the baseline group. The second set of terms, the β 's, represents the effects of the (selected) predictors on the logit of being reclassified. Each slope parameter $\beta_1, \beta_2, \dots, \beta_p$, assess the effect of one unit difference in that predictor on event occurrence, statistically controlling for the effects of all other predictors in the model. This method is preferred because it allows for a single model containing the main effect (*Time*) and the influences of all the predictor variables to be examined together on time to reclassification (Singer & Willet, 2003).

Fitting the Model to Data

Fitting a model to determine how long it takes ELs enrolled in New Mexico public schools to be reclassified as English proficient was a multi-step process. First, the final analytical data set was converted from single record format to a multi-record format, such that there is a record for each student in each year that the student appears in the data set, with all demographic variables replicated (Singer & Willet, 1993). Data on individual student characteristics, such as gender, race/ethnicity, poverty status and special education status, was obtained from their base years (SY 2011 or SY 2012) and kept constant across the five or six waves of data. High-school students are expected to graduate at the end of the fourth year and there are only four grades (grades 9-12) that students can be placed in. Therefore, they will not have the opportunity to continue into reclassification for the full six-year period. Students in the high school grade cluster were removed from the initial analysis and explored separately. Dummy variables are created for each year students were tracked specified as *D01 – D06*, with *D01* corresponding to the base year of either SY 2011 or SY 2012, depending on the cohort. The outcome

variable ‘Y’ (0 = no event, 1= event) was also coded to indicate whether proficiency was achieved in that year.

The next step was to fit a discrete-time hazard model to the data. Logistic regression analysis was conducted to predict the *event* indicator (reclassification) on the time indicator *D01 – D06* and the selected predictors in the analytical data set. The baseline model included the main effect of *Time*. Maximum likelihood estimation was carried out using a *Fisher scoring optimization algorithm*.

The PROC LOGISTIC procedure in SAS was used to fit the model. The referent groups were coded 0 and all other groups were coded 1. The following predictors: grade cluster, gender, race/ethnicity, poverty status and special education status were added to the model using less-than-full-rank parameterization (or dummy coding). In the full model, kindergarten is used as the referent group for the grade level variable; male is the baseline comparison group for gender; students identified of Native American origin was the baseline comparison group versus all other races, and students of Hispanic origin versus non-Hispanic origin were the comparison groups for race and ethnicity.

The model information in Table 5 shows that the number of observations entered ($n = 65,781$) and the number used were identical ($n = 65,781$) showing no missing data. Model convergence is usually monitored with the gradient of the log likelihood and is said to converge when the largest gradient element meets some numerically small threshold (i.e., close to zero). The model convergence status showed that the maximum likelihood estimation algorithm converged using the default gradient convergence criterion (GCONV) and default precision of 10^{-8} . The model fit statistics showed that the Akaike information criterion (AIC), the Schwarz criterion (BIC), and the deviance (i.e., -

2 times the log likelihood) for the final fitted model (with covariates) are lower than the values of the intercept only (without covariates) model (Akaike, 1974; Schwarz, 1978). AIC, BIC and the deviance are often used to evaluate models of different complexity where smaller values demonstrate better model-data fit. The Global null hypothesis $H_0 : \beta_\rho = 0$ tested that all the predictors' regression coefficient are equal to zero. Small p-values ($p < .05$) indicate that the null hypothesis should be rejected concluding that the regression slope parameters are non-zero in the population. The full model with all the covariates tested against the baseline model showed statistical significance, indicating that the variables as a set reliably distinguished between themselves on time to reclassification. ($\chi^2 = 22746.6976$, $p < .0001$, with $df = 18$). The binary logistic regression indicated that there was a significant association between each predictor variable, namely: grade cluster, gender, race/ethnicity, poverty status and special education status on the dependent variable (Y) of time to reclassification.

Table 5: Model Information

Response Variable		Y
Number of Response Levels		2
Model		Binary logit
Number of Observations Read		65,781
Number of Observations Used		65,781
Response Profile		
Ordered Value	Y	Total Frequency
1	1	8,243
2	2	57,538
Probability modeled is Y=1.		
Model Convergence Status		
Convergence criterion (GCONV=1E-8) satisfied.		
Model Fit Statistics		
Criterion	Without Covariates	With Covariates
AIC	91191.829	43864.466
SC	91191.829	44028.159
-2 Log L	91191.829	43828.466
Testing Global Null Hypothesis: BETA=0		
Test Chi-Square	Chi-Square DF	Pr > ChiSq
Likelihood Ratio	47363.3634 18	<.0001
Score	39514.4843 18	<.0001
Wald	22746.6976 18	<.0001

The final step was to estimate the number of years to achieve a survival rate of .5 (50%) based on the survival probabilities. To obtain the median values, linear interpolation was used to approximate an unknown value from two known values using a linear function (Singer & Willet, 2003). The known functions are the survival probabilities for each year. There are six known survival probabilities for the six years of data observation. To find the median number of years it takes for ELs to be reclassified, linear interpolation was used to fit a piecewise linear function to the line segments (*using four knots*) and then calculating the median value. The detailed results for each predictor are presented in the next chapter.

Additionally, a descriptive analysis was conducted to observe the reclassification patterns of ELs in grades 9-12. A descriptive analysis was also run to examine the characteristics of ELs who were not reclassified after five years; and therefore at risk of becoming long-term ELs.

The next chapter presents the results of the model building process.

IV. Findings

This chapter presents the findings based on the methodology described previously. The first section discusses the results for assessing the first two research questions: how long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED); and does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools? The findings for the third research question on how select student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time, is reported in the second section. Data on individual student characteristics, such as gender, race/ethnicity, poverty status and special education status, was obtained from their base years (SY 2011 or SY 2012) and kept constant across the five or six waves of data. This section also includes additional descriptive information on the characteristics of ELs who were right censored as well as ELs who were not reclassified after five years of receiving additional language support services in language instructional education programs is described.

How long does it take ELs to be reclassified as English proficient in New Mexico?:

Main Effect of *Time*

The analytical sample for this study consists of ELs who started school in the State for the first time in SY 2010-11 (Cohort 1) or in SY 2011-12 (Cohort 2) and were continually enrolled during the observation period ending in SY 2015-16. The sample includes students enrolled in all grades K-12, no new students were added to the analytical sample. In this section, descriptive statistics are presented followed by a

summary inferential statistics based on survival data analyses conducted in SAS. In particular, a useful tool for understanding how likely events are to occur over time is a Life Table, which is summarized in Table 6. This table displays a simple summary of the number of students from the original risk set who have not been reclassified for each time period (column A), the number of students reclassified during the year (column B), the number of students who are censored in each time interval (column C), and the number of students who did not become proficient at the end of the year and will most likely move into the risk set for next year (column D). Using these basic counts, Table 6 also displays hazard rate of reclassification in each year, the proportion of ELs who are right censored, and the cumulative survival rates. It should be noted that the number in the risk set for each year does not exactly add up to the number reclassified and the number censored minus the beginning total. This may be because some students who (for different reasons) did not have test scores for one year, may appear in the not proficient group for that year, or have dropped out the next year and not appear in the risk set for that year but show up in the risk set again the following year. The model relies on discrete hazard analysis and provided meaningful estimates to account for censorship and the in and out movement of students in the original risk set.

Table 6: Life table by school year, combined cohorts (K-8)

School Year	Beginning Totals (Risk set)	Reclassified	Censored	Number not proficient	Hazard rate (proportion who were reclassified)	Proportion who were right censored	Survival rate **
	A	B	C	D	B/A	C/A	(A-B-C)/A
1	17,133	875	2,441	16,258	0.05	0.14	0.81
2	13,817	1,039	1,165	12,778	0.08	0.08	0.84
3	12,066	1,410	811	10,656	0.12	0.07	0.82
4	10,319	2,915	541	7,404	0.28	0.05	0.67
5*	7,365	1,720	351	5,645	0.23	0.05	0.72
6*	2,714	421	314	2,293	0.16	0.12	0.73

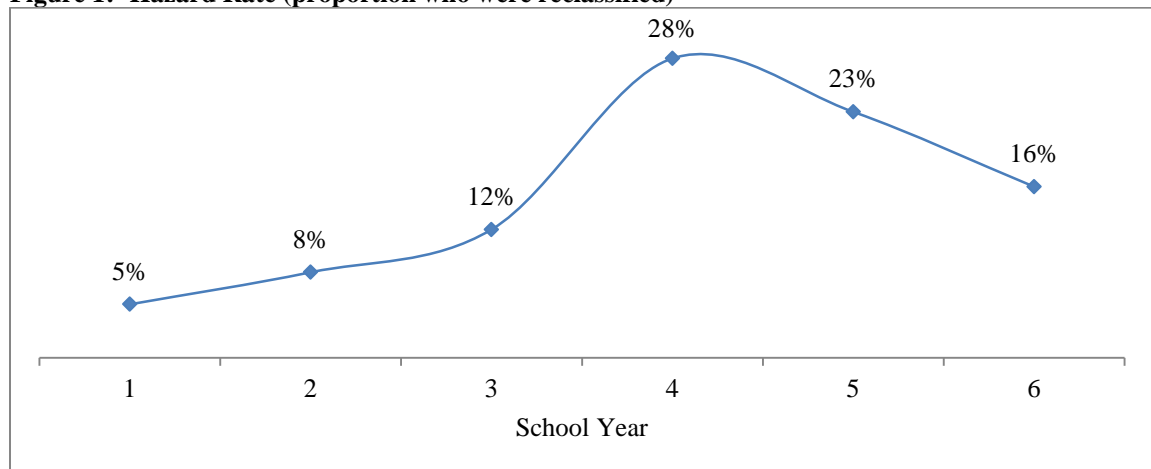
* The two cohorts are observed over different but overlapping years. School years have been renumbered as 1 through 6 for both cohorts. Data collection for Cohort 1 ended after 6 years. Data collection for Cohort 2 ended after 5 years.

** Survival rate is the proportion of students who remain in the risk set after each wave

Table 6 shows important statistics about the reclassification patterns each year.

The hazard rate appears to increase every year, peaking at year four and then decreases towards the end of the observation period. Of the ELs who experienced reclassification, the largest proportion (28 percent; n= 2,915) were observed in year four of the observation period. The second large proportion (23 percent; n=1,720) of ELs who experienced reclassification were observed in in year five of the observation period. Overall a quarter of the students in the initial cohort were reclassified, the rest did not become proficient and either moved to the next grade, transferred, or dropped out of school.

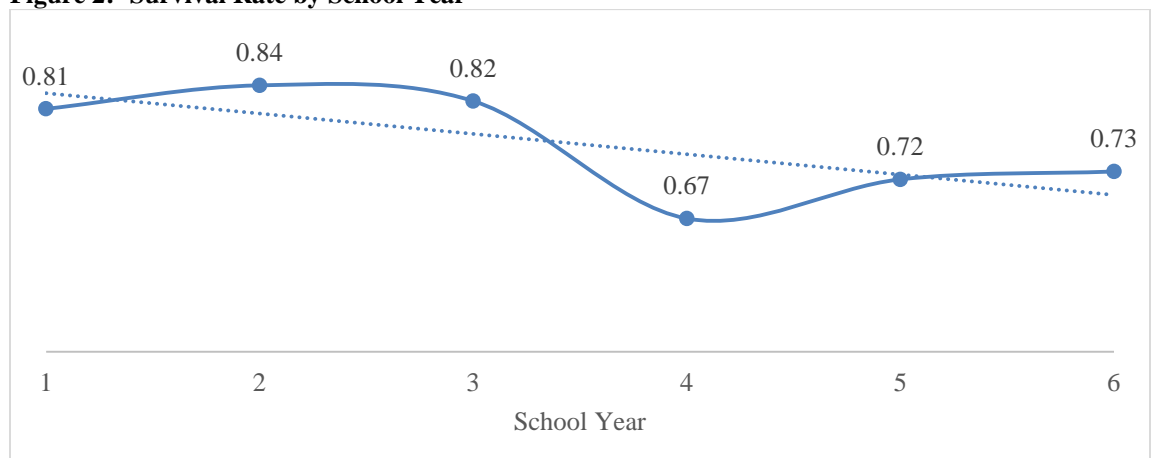
Figure 1: Hazard Rate (proportion who were reclassified)



The risk set (Column A, Table 6) suggests that approximately half the students have been reclassified after five years. Approximately 30 percent of ELs were censored after five years. The proportion of students remaining in the initial cohort decreases substantially over time during the first four years. For instance, in Year four, 65% from the initial risk set of 17,133 ELs remained as such, while in Year five 43 percent remained in the risk set and in Year six 16 percent.

As depicted in Figure 2, the survival rate shows an overall declining trend as a function of school year. The largest decline is between Year 3 and 4. However, there is a subsequent rise till the end of the observation period of six years.

Figure 2: Survival Rate by School Year



Model Estimates of Time to Proficiency

Each year of students' start year was used as a covariate in a regression model to control for the possible differences between the cohorts. In this model, reclassified or not reclassified status was used as the dependent variable. High school students (grades 9-12) were removed from the model as high school students do not have the opportunity to continue into reclassification for the full six year period. Table 7 shows the model estimates for the effect of time on reclassification for grades K-8. Accordingly, an odds ratio of 1 indicates that two groups have the same probability of experiencing reclassification at each time point. Odds ratios greater than 1 indicate that a particular group is more likely to experience reclassification, while odds ratios less than 1 indicate that a particular group is less likely to experience the event. As shown in the highlighted column three (Log Odds Estimates), and column eight (odds ratios estimates) in Table 7 below, the odds of reclassification varies for each year of the observation period. The results show the odds of reclassification go up over the years (Year 1 – Year 4) plateaus in Year 5 and drops in Year 6. The drop in Year 6 could be an artifact of the sample because only one cohort is contributing data to the estimate in Year 6.

Table 7: Analysis of Maximum Likelihood Estimates by Years, Combined Cohorts, K-8

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi- Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp (β))	95% CI for exp (β)
Year 1	1	-2.7093	0.0845	1027.231	<.0001	-0.6556	0.067	[0.056, 0.079]
Year 2	1	-2.3499	0.0841	781.5256	<.0001	-0.538	0.095	[0.081 0.112]
Year 3	1	-1.8434	0.0826	497.5795	<.0001	-0.4023	0.158	[0.135 0.186]
Year 4	1	-0.5857	0.0803	53.1311	<.0001	-0.1203	0.557	[0.476 0.652]
Year 5	1	-0.5335	0.0819	42.474	<.0001	-0.0941	0.587	[0.500 0.689]
Year 6	1	-0.742	0.09	67.9568	<.0001	-0.0814	0.476	[0.399 0.568]
Start Year	1	0.1268	0.0259	24.0381	<.0001	0.0349	1.135	

To obtain the median value, an interpolation was used to approximate the median values using a linear function. Using the TRANSREG Procedure in SAS, Table 8 below displays the years to achieve a survival rate of .5 (50 percent). This generates a survival rate for one half of the sample, or how long it takes half the students to achieve proficiency. Row 7 in Table 8 shows that it takes 4.02 years to get to a survival rate of 0.5. In other words the median number of years it takes for ELs to be reclassified in New Mexico is four years.

Table 8: Base Model Estimate of Years to Achieve Survival Rate of 50%, Combined Cohorts, K-8

Name	Time Period*	Period	Intercept	Survival
Row 1	1	1.36	1	0.94
Row 2	2	1.85	1	0.86
Row 3	3	2.56	1	0.74
Row 4	4	4.17	1	0.48
Row 5	5	5.24	1	0.30
Row 6	6	5.82	1	0.20
Row 7		4.02	1	0.50

*Model iterations converged at R-square= 0.97398

Effect of Starting Grade at Which ELs Enter the School System

The variable, grade band, included the grades at which the student entered the analytical sample at the start of the observation period (in SY 2010-11 and SY 2011-12). The grades are clustered by the five grade bands used in the ACCESS test forms. The data (see Table 9) shows approximately 60 percent of ELs entered the school system in kindergarten at the start of the observation period in SY 2010-11 and SY 2011-12. The proportion of (newly enrolled) ELs decreased as the grades increased up till the middle school grades 6-8. There is an upward spike in the number of newly enrolled in grade 9 (n= 519) but the numbers drop off again for the rest of the high school grades. Overall, there were more ELs that started school in grades 9-12 than in the middle school grades 6-8. However, the high school grade band could not be included in the survival analysis because ELs in grades 9-12 timed out of school before the end of the five or six year

observation period. As there are a large number of ELs that start school in grades 9-12 (see Table 9), it is important to pay attention to the rates of reclassification among high school ELs. In this study a descriptive analysis of reclassification of ELs starting school in grades 9-12 is discussed separately. First the estimates of the rate of reclassification for grades K-8 are presented followed by the analysis of grades 9-12.

Table 9: Enrollment County by Grade Band (K-12)

	2011		2012		Total	
Entry Grade	Number	Percentage	Number	Percentage	Number	Percentage
Kindergarten	5,283	57.3%	5,876	64.8%	11,159	61.0%
1st - 2nd	1,684	18.3%	1,284	14.2%	2,968	16.2%
3rd - 5th	1,164	12.2%	887	9.8%	2,051	11.2%
6th - 8th	476	5.2%	479	5.3%	955	5.2%
9th - 12th	615	6.7%	549	6.1%	1164	6.4%
Total Sample	9,222	100.0%	9,075	100.0%	18,297	100.0%

Estimates of Reclassification Rates for Grades K-8

The estimates for the effects of grade band (K-8) on reclassification show statistical differences in the median time to reclassification between grades bands as compared to Kindergarten. The analysis of maximum likelihood estimates obtained through logistic regression (Table 10) show the log-odds of being reclassified increase by 0.57 units for ELs who entered in grades 1-2 compared to ELs who entered in Kindergarten. This result is significant, Wald chi-square (1 df) = 281.84, $p < .0001$). Table 10 also displays the odd ratio estimates. The odds ratio coefficient for grades 1-2 is 1.77 with a 95% confidence interval of [1.66, 1.90]. This suggests, during the observation period, ELs who started school in grades 1-2 are between 66 and 90 percent more likely to become proficient than students who entered in Kindergarten. For students entering in grades 3-5, the analysis of maximum likelihood estimates (Table 10) show the log-odds of being reclassified increase by 0.64 as compared to ELs who started school in

Kindergarten. This result is significant, Wald chi-square (1 df) = 251.24, $p < .0001$. The odds ratio coefficient for grades 3-5 is 1.90 with a 95 percent confidence interval of [1.76, 2.06]. This suggests, during the observation period, ELs starting in grades 3-5 are between 76 percent and 106 percent more likely to become proficient by the end of the observation period than students who entered in Kindergarten. For ELs entering in middle school grades 6-8, the maximum likelihood estimates (Table 10) show the log-odds of being reclassified increase by 0.80 for ELs entering in middle school grades 6-8 as compared to ELs entering in Kindergarten. This result is significant, Wald chi-square (1 df) = 190.75, $p < .0001$. The odds ratio coefficient for grades 6-8 is 2.22 with a 95% confidence interval of [1.99, 2.49]. This suggests that ELs starting school in grades 6-8 are between 99 percent and 149 percent more likely to become proficient than students who start school in Kindergarten.

Table 10: Analysis of Maximum Likelihood Estimates by Grade Bands

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp β)	95% CI for exp (β)
G1to2	1	0.57	0.03	281.84	<.0001	0.11	1.77	[1.66 1.90]
G3to5	1	0.642	0.04	251.24	<.0001	0.10	1.90	[1.76 2.06]
G6to8	1	0.80	0.06	190.75	<.0001	0.09	2.22	[1.99 2.49]
Start Year	1	0.13	0.03	24.04	<.0001	0.04	1.14	

Overall, there is statistically significant variance in the likelihood estimates of time to reclassification among the grade bands. Based on the estimates presented above ELs who enter school in middle school are likely to have higher rates of reclassification as compared to students who start in Kindergarten.

Median Rates of Reclassification for Grades K-8

Table 11 shows estimates of the time it takes half the ELs to be reclassified in a given grade band. Row 7 shows the number of years to achieve a survival rate of 0.5 (or 50 percent).

ELs entering school in Kindergarten had the longest median time to reclassification at 4 years. Kindergarten was used as the baseline group. ELs entering school in middle school (grades 6-8) show the lowest median time of reclassification (Period) at 3.1 years as compared to Kindergarten at 4.0 years. ELs entering school in other elementary grades also had lower median time to reclassification as compared to the Kindergarten comparison group. Data presented in Table 11 shows the median time to reclassification for ELs entering in grades 1-2 is 3.3 years as compared to Kindergarten with a median rate of 4 years. The time to reclassification for one half of ELs entering in grades 3-5 was less than ELs entering in Kindergarten and grades 1-2 but more than ELs entering in grades 6-8. Specifically, it takes 3.3 years to get to a survival rate of 0.5 for ELs entering in grades 3-5.

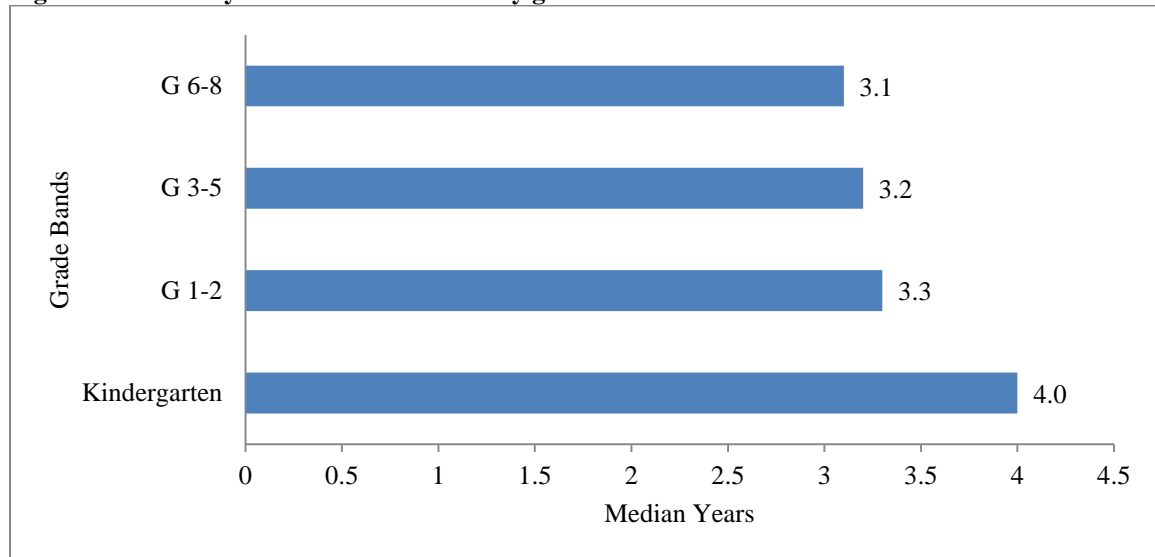
Table 10: Estimate of Years to Achieve Survival Rate of 50% by Grade Band, Combined Cohorts, K-8

Name	Intercept	Time Period*	Grades 1-2		Grades 3-5		Grades 6-8	
			Period	Survival	Period	Survival	Period	Survival
Row 1	1	1	1.14	0.89	1.11	0.89	1.05	0.87
Row 2	1	2	1.85	0.76	1.86	0.75	1.87	0.72
Row 3	1	3	2.77	0.60	2.80	0.58	2.89	0.53
Row 4	1	4	4.40	0.30	4.42	0.28	4.48	0.24
Row 5	1	5	5.23	0.15	5.23	0.13	5.21	0.10
Row 6	1	6	5.60	0.08	5.57	0.07	5.50	0.05
Row 7	1		3.30	0.50	3.23	0.50	3.06	0.50

*Model iterations converged at R-square = 0.97351 (Grades 1-2); R-square = 0.97225 (Grades 3-5); R-square = 0.96805 (Grades 6-8)

Figure 3 displays the estimated median time to reclassification among grade bands K-8 based on the estimates presented in Table 11. A lower median rate means it takes less time for one half of the students in that grade band to be reclassified.

Figure 3: Median years to reclassification by grade band



Influence of Starting Proficiency Levels on Median Time to Proficiency for Grades K-8

Literature cited in this study reports that the starting proficiency is an important factor in assessing the probability of reclassification by grade band (Hakuta, et al., 2000; Kieffer, 2011; Haas, et al., 2015). A descriptive analysis of the distribution of frequencies at each proficiency levels nested by grade bands is presented to gauge the impact of starting proficiency levels on time to reclassification. Table 12 below displays the details of the frequency distribution for each year of the observation period. It is important to note that the actual starting proficiency level of the students at the start of the school was not available. Therefore the proficiency recorded at the end of the first year of enrollment was used as the baseline. The proficiency levels were combined into three categories: Level 1 and Level 2 were combined and are represented by ‘beginning’ levels of proficiency at the end of year 1. Level 3 and 4 are combined to represent

‘intermediate’ level of proficiency and the remaining ELs who scored at a level 5 or above are marked ‘advanced or proficient’. The advanced or proficient category includes students who were reclassified at the end of the first year of enrollment. The distribution shows that a majority of ELs in Kindergarten (86 percent) started at the beginning proficiency level at the end of the first year (n=9,643); and 12 percent started at the intermediate level (n=1308). Proportionately, in grades 1-2, about 30 percent started at the beginning level compared to 65 percent at the intermediate level by end of year 1. Similarly more students were at the intermediate level in grades 3-5 (60 percent) and grades 6-8 (56 percent) than at the beginning level at the end of the first year of enrollment. The proficiency level with the highest frequency count is highlighted in the column marked ‘number of students’. Grades 1-8 had the highest proportion of ELs at the intermediate level while Kindergarten had the highest proportion of ELs at the beginning level at the end of the first year of instruction.

Table 11: Frequency distribution of ELs reaching proficiency by year and by proficiency level

Grade Cluster*	Proficiency Level End of Year 1	Number of Students	Collection year proficient:					
			1	2	3	4	5	6
Kindergarten	0 - 2.9	9643 (.86)	--	58	364	2,059	1,422	357
	3 - 4.9	1308 (.12)	--	164	211	426	102	24
	5 +	208 (.02)	208	--	--	--	--	--
1st-2nd	0 - 2.9	893 (.30)	--	12	70	87	55	11
	3 - 4.9	1932 (.65)	--	390	464	194	63	**
	5 +	143 (.05)	143	--	--	--	--	--
3rd-5th	0 - 2.9	402 (.20)	--	--	32	14	20	**
	3 - 4.9	1235 (.60)	--	314	136	42	38	14
	5 +	414 (.20)	414	--	--	--	--	--
6th-8th	0 - 2.9	312 (.33)	--	**	28	31	**	**
	3 - 4.9	533 (.56)	--	93	105	62	13	**
	5 +	110 (.12)	110	--	--	--	--	--
Total (ALL)		17133	875	1,039	1,410	2,915	1,720	421

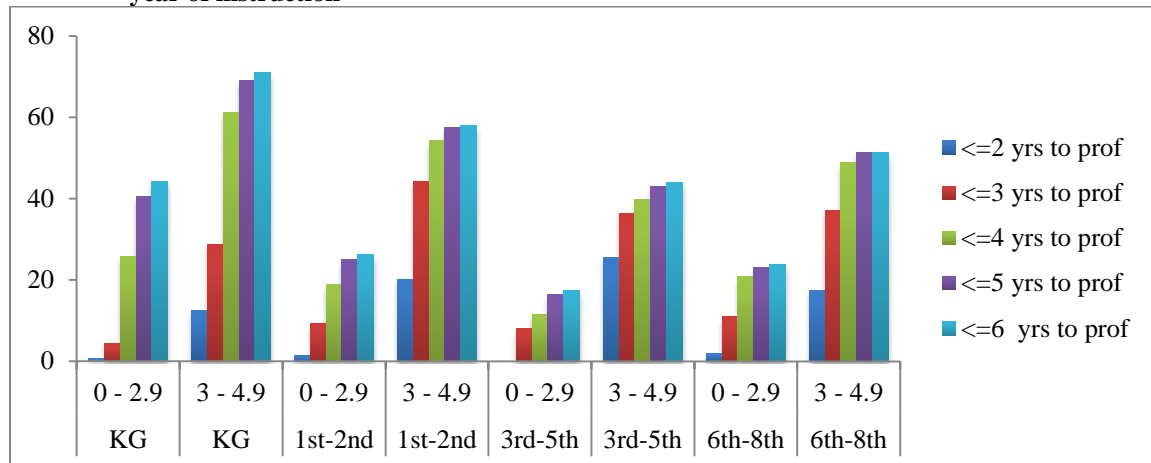
* this data table includes both cohorts

** cells with n<10; -- cells with n=0

The figure below also shows the cumulative proportions of ELs who were reclassified each year. On the x-axis are the grade level and proficiency levels clustered

by beginning (0-2.9) and intermediate (3-4.9) levels. On the y-axis is the cumulative number of ELs reclassified each year represented by a unique color.

Figure 4: Cumulative Proportion of ELs reclassified by year grouped by proficiency level after one year of instruction



Taken together the results of time to reclassification by grade level and starting proficiency level suggest that the starting proficiency level does impact time to reclassification. Students starting at lower proficiency level take longer than ELs starting at higher proficiency level, which is supported by literature. The patterns of reclassification for ELs starting at the beginning or intermediate level look similar by grade band. The median time to proficiency was different for each of the grades bands as shown in Figure 3. However, there may not be much practical impact between a median of 3.1 for grades 6-8 and 3.3 for grades 2-3 on the actual school year that ELs become reclassified. ELs will actually stop receiving services and are reclassified either after year three or after year four in school. A difference of 0.1 between the grades 2-8 when compared to Kindergarten may have little practical significance. However, Kindergarteners reclassified after year 4 may actually be receive services as an EL for an additional year as compared to all other grades.

Reclassification Rates of High School ELs

Analyzing data on the rate of reclassification of ELs entering in high school (grades 9-12) is challenging. Many ELs may leave high school without actually achieving proficiency and many may not become proficient before they leave high school. For the purpose of this study, ELs are considered new entries based on having no prior data for 1 year, and having "0" in the field indicating whether they had been in U.S. schools for 3 or more years. ELs who started ESL programs in states other than New Mexico, and ELs with missing data for the year previous to the baseline cohort year may be included in the data. Ninth graders with at least 4 years of data, 10th graders with at least 3 years of data, and 11th graders with at least 2 years of data are all considered to have reached the end of high school. Table 13 shows the percent of students who were classified as ELs and entered programs in High School in New Mexico in SY 2011 and SY2012, who left before becoming proficient, became proficient, or did not become proficient by the end of High School, by entry grade.

Table 13 shows that a majority of students entered high school in grade 9 (n= 519). The numbers drop off as the grades increase with only 114 students starting in grade 12. The data also shows that a majority of ELs who entered high school in SY 2011(Cohort 1) and SY 2012 (Cohort 2) either left high school for some reason before becoming proficient or continued to enroll in school but were not proficient by grade 12. 44.32 percent of students who entered high school in 9th grade (n= 519) became proficient by the end of high school, however, almost an equal proportion 43.93 percent left before the end of high school without becoming proficient. 11.75 percent of students who entered high school in 9th grade did not become proficient before the end of high

school. The descriptive analysis show 36.20 percent ELs entering high school in grade 10 (n=337) became proficient by end of high school, nearly 8 percentage points below those starting in grade 9. Less than half (43.81 percent) of ELs entering high school in grade 11 (n=194) became proficient by end of high school, 0.51 percentage points below those starting in 9th grade. However, the number of students who started in 12th grade (n=114) and became proficient at the end of high school is much lower than all of the other grades, nearly 18 percentage points less than those who started in 9th grade. A quarter of the students (26 percent) who started in grade 12 met the proficiency criteria by the end of the year.

Table 12: Reclassification Patterns Among High School Grades

Entry Grade	2011				2012				2011 and 2012			
	Students	A*	B**	C***	Students	A	B	C	Students	A	B	C
9	261	48.28 %	10.34 %	41.38 %	258	39.53 %	13.18 %	47.29 %	519	43.93 %	11.75 %	44.32 %
10	179	45.25 %	24.02 %	30.73 %	158	37.34 %	20.25 %	42.41 %	337	41.54 %	22.26 %	36.20 %
11	109	36.70 %	23.85 %	39.45 %	85	25.88 %	24.71 %	49.41 %	194	31.96 %	24.23 %	43.81 %
12	66	.	77.27 %	22.73 %	48	.	68.75 %	31.25 %	114	.	73.68 %	26.32 %

* Students who left before end of HS without becoming proficient.

** Students not proficient by end of HS.

*** Students who became proficient by end of HS.

Influence of Select Student Characteristics on the Probability of Reclassification for ELs

In this section the effects of various predictors on the odds of reclassification in the final fitted logit hazard model are displayed and interpreted. Data on select individual student characteristics, namely gender, race-ethnicity, poverty status and special education status, was obtained from their base years (SY 2011 or SY 2012) and kept constant across the five or six waves of data. The model estimates presented in the section are based on ELs enrolled in grades K-8 at the start of the observation period. In

the full model, kindergarten is used as the referent group for the grade level variable; male is the baseline comparison group for gender; students identified of Native American origin was the baseline comparison group versus all other races, and students of Hispanic origin versus non-Hispanic origin were the comparison groups for race and ethnicity. To find the median number of years it takes for ELs to be reclassified, an interpolation was used to fit a piecewise linear function to the line segments. Each subsection presents a table showing the odds ratios as well as the confidence intervals associated with the predictor. Since the indicators are binary, the odds ratios represent the difference in odds between the comparison group and the alternative groups. Results on each predictor are presented in a series of subsections.

Gender

The distribution of males (n= 8,945) and females (n= 8,188) in the analytical sample for grades K-8 is approximately 52 percent female and 48 percent male. The maximum likelihood estimates in Table 14 show the log-odds of being reclassified increase by 0.26 units for females compared to males, which is statistically significant ($p<.0001$). Moreover, the odds ratio estimates suggest that females are 29 percent more likely to become proficient than males.

Table 13: Estimates by Gender

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi- Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp β)	95% CI for(Exp β)
Female	1	0.26	0.03	106.00	<.0001	0.07	1.29	[1.23, 1.36]

Table 14: Gender differences in likelihood of achieving proficiency for ELs

Parameter	2011		2012		Total		Median Years to reclassification	Likelihood of Proficiency within six years
Male	4,49	52.21%	4,451	52.21%	8,945	52.21%	4	<i>Comparison group</i>
Female	4,11	47.79%	4,075	47.79%	8,188	47.79%	3.7	29 percent <i>more likely</i> to become proficient than male students*

* Significant at $p < 0.01$

Time to reclassification for one half of females using males as the comparison group was calculated using linear interpolation. Row 7 in Table 16 shows it took 3.7 years for females to achieve a survival rate of 0.5 (or 50 percent). This means that females had a median survival rate of 3.7 years as compared to males at 4 years.

Table 15: Base Model Estimate of Years to Achieve Survival Rate of 50% for Females, Combined Cohorts, K-8

Name	Intercept	Time period*	Period	Survival
Row 1	1	1	1.268	0.92
Row 2	1	2	1.85	0.82
Row 3	1	3	2.64	0.68
Row 4	1	4	4.267	0.395
Row 5	1	5	5.24	0.23
Row 6	1	6	5.73	0.14
Row 7	1		3.67	0.50

*Model iterations converged at R-square = 0.97561

Race and Ethnicity

Non-Hispanic or non-Latino (n= 3,963) was used as the racial comparison group for students identified as Hispanic (n= 13,170). The results displayed in Table 17 show that the log-odds of being reclassified decreased by 0.18 units for students identified as Hispanic compared to students classified as not Hispanic or Latino, which is statistically significant ($p < .0059$). Moreover, the odds ratio estimates suggest that Hispanic ELs are 17 percent less likely to become proficient than ELs classified as non-Hispanic or non-Latino.

In another analysis, ELs who identified their ethnicity as American Indian/Alaskan Native (n= 2,863) were compared to ELs who were not American Indian/Alaskan Native (n=14,270). The results displayed in Table 17 show that the log-odds of being reclassified as proficient decreased by 0.31 units for students identified as American Indian/Alaskan Native compared to students classified as not American Indian/Alaskan Native, which is statistically significant ($p < .0017$). Moreover, the odds ratio estimates suggest that American Indian/Alaskan Native are 26 percent less likely to become proficient than not American Indian/Alaskan Native ELs.

Table 16: Estimates by Ethnicity

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp (β))	95% CI for exp (β)
Native	1	-0.31	0.1	9.8092	0.0017	-0.06	0.74	[0.61, 0.89]
Hispanic	1	-0.18	0.07	7.5899	0.0059	-0.04	0.83	[0.73, 0.95]

Table 17: Likelihood of achieving proficiency for ELs by Ethnicity

Parameter	2011		2012		Total		Median Years to reclassification	Likelihood of Proficiency within six years
Not Hispanic or Latino	1,855	21.55%	2,108	24.72%	3,963	23.13%	4	Comparison group
Hispanic or Latino	6,752	78.45%	6,418	75.28%	13,170	76.87%	4.3	17 percent less likely to become proficient than students who are not Hispanic or Latino*
Not American Indian/Alaskan Native	7,267	84.43%	7,003	82.14%	14,270	83.29%	4	Comparison group
American Indian/Alaskan Native	1,340	15.57%	1,523	17.86%	2,863	16.71%	4.5	26 percent less likely to become proficient than students who are a race or ethnicity other than Native*

* Significant at $p < 0.01$.

The estimates from linear interpolation are displayed in Table 19. The estimates show it takes 4.3 for ELs who are Hispanic to achieve a survival rate of 50 percent, which is to say the median years to reclassification for Hispanic ELs were 4.3 years as compared to the median years to reclassification at 4 years for the comparison group of ELs who are not Hispanic or Latino. ELs entering school in grades K-8, who identified themselves as American Indian/Alaskan Native, seem to take longer to be reclassified than any other ethnic group. Estimates show the median years to reclassification for American Indian/Alaskan Native ELs were 4.5 years as compared to the median years to reclassification for not American Indian/Alaskan Native ELs at 4 years.

Table 18: Estimate of Years to Achieve Survival Rate of 50% by Ethnicity Combined Cohorts, K-8

Name	Intercept	Time period*	Hispanic		Native American	
			Period	Survival	Period	Survival
Row 1	1	1	1.41	0.95	1.44	0.95
Row 2	1	2	1.86	0.88	1.87	0.89
Row 3	1	3	2.52	0.78	2.50	0.80
Row 4	1	4	4.10	0.53	4.06	0.57
Row 5	1	5	5.23	0.36	5.22	0.40
Row 6	1	6	5.88	0.25	5.91	0.29
Row 7	1		4.29	0.50	4.51	0.50

*Model iterations converged at R-square = 0.97170 (Hispanic); 0.96979 (Native American)

Impact of Poverty

Qualifying to receive free and reduced price lunch is used as a proxy indicator for poverty. Approximately 90 percent of ELs in the analytical sample qualified for free or reduced price lunch programs and likely to be experiencing poverty (n=15,483). The peer group used for comparison was ELs who did not qualify to receive free and reduced price lunch (n= 1,650). The likelihood analysis (Table 20) showed that the log-odds of being reclassified as proficient decreased by 0.53 units for ELs experiencing poverty as compared to peers who are not experiencing poverty (p<.0001). The odds ratio estimates

displayed in Table 20 show that ELs experiencing poverty are 41 percent less likely to become proficient than their peers.

Table 19: Estimates by Ethnicity

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp β)	95% CI for Exp β
Free and Reduced Lunch (Poverty)	1	-0.53	0.04	155.24	<.0001	-0.08	0.59	[0.542, 0.64]

Table 20: Likelihood of achieving proficiency for ELs by Ethnicity

Parameter	2011		2012		Total		Median Years to reclassification	Likelihood of Proficiency within six years
No Free and Reduced Price Lunch	902	10.48%	748	8.77%	1,650	9.63%	4	Comparison group
Free and Reduced Price Lunch	7,705	89.52%	7,778	91.23%	15,483	90.37%	5	41 percent less likely to become proficient than students who do not receive free or reduced price lunch *

* Significant at $p < 0.01$

The median time to reclassification for ELs experiencing poverty is 5 years displayed in Table 21 as compared to the median time to reclassification of 4 years for ELs who are not experiencing poverty. Row 7 (Table 22) estimates show it takes 5 years for ELs from poverty to achieve a survival rate of 0.5 (or 50 percent).

Table 21: Base Model Estimate of Years to Achieve Survival Rate of 50% for ELs from Poverty, Combined Cohorts, K-8

Name	Intercept	Time period*	Period	Survival
Row 1	1	1	1.50	0.96
Row 2	1	2	1.88	0.91
Row 3	1	3	2.46	0.83
Row 4	1	4	4.00	0.63
Row 5	1	5	5.20	0.47
Row 6	1	6	5.96	0.36
Row 7	1		4.95	0.50

*Model iterations converged at R-square = 0.96598

ELs with Special Needs

In the analytical sample used in this study ELs (grades K-8) receiving special education services (n= 1,725) were approximately 10 percent (n= 15,408). The Individuals with Disabilities Education Act (IDEA) provides definitions of the thirteen disability categories that guide how states define who is eligible for a free appropriate public education. However, to protect the privacy of students, the aggregated total number of ELs with disabilities identified to receive special education services is used. Additionally, ELs with severe cognitive disabilities were not included in the sample because those ELs took the Alternate ACCESS assessment. The alternate ACCESS is administered to only those ELs who are identified with severe cognitive disabilities in grades 1-12. The Alternate ACCESS for ELs amplifies Proficiency Level 1 and provides students with severe cognitive disabilities a chance to demonstrate progress within Level 1. Since the Alternate ACCESS is on a different scale than the ACCESS, the scores could not be merged with the results of the ACCESS. This may underestimate the results for this variable.

ELs not receiving special education services were used as a reference group to students receiving special education services. Table 23 shows the log-odds of being reclassified as proficient decreased by .83 units for students receiving special education services compared to students who were not receiving special education services, significant at $p < .0001$. Moreover, the odds ratio estimate suggests that ELs receiving special education services are 56 percent less likely to become proficient than ELs who are not receiving special education services.

Table 22: Estimates by Special Education Services

Parameter	DF	Estimate (Log Odds)	Standard Error	Wald Chi- Square	Pr > ChiSq	Standardized Estimate	Odds Ratio (Exp (β))	95% CI for exp (β)
Special Education Services	1	-0.83	0.051	262.33	<.0001	-0.14	0.44	[0.40, 0.48]

Table 23: Likelihood of achieving proficiency for ELs by Ethnicity

Parameter	2011		2012		Total		Median Years to reclassification	Likelihood of Proficiency within six years (based on Odds ratio estimates)
No Special Education	7,742	89.95%	7,666	89.91%	15,408	89.93%	4	Comparison group
Special Education	865	10.05%	860	10.09%	1,725	10.07%	5.7	56 percent less likely to become proficient than students who do not attend special education classes part- time*

* Significant at $p < 0.01$

The median years to reclassification for ELs who are receiving special education services is 5.7 years as compared to the median years to reclassification at 4 years for ELs who are not receiving special education services (Table 25).

Table 24: Estimate of Years to Achieve Survival Rate of 50% for ELs with disabilities, Combined Cohorts, K-8

Name	Intercept	Time period*	Period	Survival
Row 1	1	1	1.56	0.97
Row 2	1	2	1.90	0.93
Row 3	1	3	2.43	0.87
Row 4	1	4	3.92	0.70
Row 5	1	5	5.18	0.56
Row 6	1	6	6.02	0.46
Row 7	1		5.68	0.50

*Model iterations converged at R-square = 0.96

Characteristics of ELs at Risk of Becoming Long-Term ELs

The data show, on the average, ELs who entered school and began receiving EL services in SY 2010-11 and SY 2011-12 (in grades K-8) achieved proficiency within four years. However, approximately 30 percent (n=5, 068) did not reach proficiency within five years, and may be at risk of becoming long-term ELs. The term long-term ELs is generally used to distinguish between students who are newly identified as ELs and achieve proficiency within a reasonable time frame- usually less than five or six years, from those who been educated in U.S. schools for more than six years and have not met English proficiency criteria (Thompson 2015). The definition of a long-term EL varies from State to State. New Mexico does not have a published definition of a long-term EL; however, starting with SY 2017-18, the State is required under ESSA to report the number of students that have not achieved proficiency within five or more years from initial identification as an EL.

In New Mexico, of the sample of students at risk of becoming long-term ELs (see Table 26), a majority entered into the cohorts in Kindergarten (72.73 percent). Less than 4 percent of ELs at risk of becoming long-term ELs entered the cohorts in middle school (3.18 percent). A majority of ELs in the sample who had not achieved proficiency in five years are Hispanic or Latino (78.37 percent) and those experiencing poverty (94.97 percent). ELs identified to receive special education service constituted 10 percent of total data sample used at the start of the observation period; however, ELs identified to receive special education service represented 15 percent of the ELs in the data set representing ELs who had not at risk of becoming long-term ELs. This is congruent with the findings above showing ELs with disabilities are 56 percent less likely to become

proficient than peers. Table 26 displays the detailed breakdown of the characteristics of students at risk of becoming long-term ELs in New Mexico.

Table 25: Characteristics of ELs in grades K-8 who did not become proficient within 5 years

Student Subgroup	2011		2012		Total¹	
Total Sample	2,385	100.00%	2,683	100.00%	5,068	100.00%
Entry Grade						
Kindergarten	1,588	66.58%	2,098	78.20%	3,686	72.73%
1st - 2nd	408	17.11%	266	9.91%	674	13.30%
3rd - 5th	322	13.50%	225	8.39%	547	10.79%
6th - 8th	67	2.81%	94	3.50%	161	3.18%
Gender						
Female	991	41.55%	1,130	42.12%	2,121	41.85%
Race / Ethnicity						
Hispanic or Latino	1,896	79.50%	2,076	77.38%	3,972	78.37%
American Indian/Alaskan Native ¹	429	17.99%	512	19.08%	941	18.57%
Special Populations						
Free and Reduced Price Lunch	2,227	93.38%	2,586	96.38%	4,813	94.97%
Special Education	380	15.93%	391	14.57%	771	15.21%

¹Number and Percentage of students who did not become proficient within 5 years for ELs in grades K - 8th, who enrolled for the first time in New Mexico SY 2010-11 & SY 2011-12

Right Censored Students

Right Censored ELs in this context is comprised of students who were not reclassified during the observation period. As such, there were a total of 4,812 ELs who did not achieve proficiency at the end of the observation period. This number reflects a six year observation period for ELs in Cohort 1 starting in SY 2010-11 and five year observation period for ELs in Cohort 2 starting in SY 2011-12. These students could have impacted the estimates because they continue to be enrolled in the school system and contribute to the pool of ELs in the analytical data sample. Table 27 summarizes the characteristics of ELs who were right censored. The proportion of right censored shows 95 percent were students who came from poverty (n=4,579). The second highest number of ELs who were not reclassified within the observation period were ELs identified as Hispanics or Latino (n= 3,780), accounting for nearly 80% of right censored ELs. Table 27 also shows

that approximately 74 percent of ELs who did not become proficient during the observation period entered the cohort in Kindergarten. On the other hand, the lowest proportion of ELs who did not become proficient was ELs who entered the cohort in middle school grades 6-8 (2.91 percent). This seems congruent with the findings in this study of the characteristics of ELs who did become proficient. Taken together, the information presented on right censored ELs during the observation period suggests that the exclusion of these students from the main analysis may have changed the point estimates but not the overall conclusions that ELs who are Hispanic, Native American or Alaska/Native, come from poverty, or receive special education services take longer to be reclassified as compared to their peers.

Table 26: Number and percentage of ELs in grades K-8 who did not become proficient during the observation period starting with SY 2010-11 and ending in SY 2015-16

Student Subgroup	SY 2010-11		SY 2011-12		Total	
	Number	Percentage	Number	Percentage	Number	Percentage
Total Sample	2,129	100.00%	2,683	100.00%	4,812	100.00%
Entry Grade						
Kindergarten	1,453	68.25%	2,098	78.20%	3,551	73.79%
1st - 2nd	358	16.82%	266	9.91%	624	12.97%
3rd - 5th	272	12.78%	225	8.39%	497	10.33%
6th - 8th	46	2.16%	94	3.50%	140	2.91%
Gender						
Female	881	41.38%	1,130	42.12%	2,011	41.79%
Race / Ethnicity						
Hispanic or Latino	1,704	80.04%	2,076	77.38%	3,780	78.55%
American Indian/Alaskan Native ¹	370	17.38%	512	19.08%	882	18.33%
Special Populations						
Free and Reduced Price Lunch	1,993	93.61%	2,586	96.38%	4,579	95.16%
Special Education	335	15.74%	391	14.57%	726	15.09%

Note: Observation periods are 6 years for SY 2010-11, and 5 years for SY 2011-12

Limitations of the Study

The current study utilized a limited sample of newcomer ELs starting in two school years in one State. Generalization of these findings to other States assessments cannot be made without further investigation. Moreover, this study was challenged by a few data limitations. First, there was no initial proficiency score recorded for each student, since the starting point for the analysis was the ACCESS test score given at the end of the student's first recorded school year. This means that it was not possible to determine the student's true initial proficiency when they began school in the State. The initial proficiency is recorded after one year of instruction. This limited the use of initial proficiency level as a variable in survival analysis to determine the time to proficiency. Second, there was no record for the date of entry to schools, leading to the assumption that all students started on the same day. This means that ELs who may have started at the start of the school year were considered on the same footing as those who came in at the end of the school year. Third the analytical sample only included those newcomer ELs who had been in the United States for zero to two years. The study did not examine those ELs who might have been born in the U.S. but who came from families that predominately spoke another language at home. Of the ELs included in the sample, the study could not determine the impact of additional English exposure on students in U.S. schools for 1-2 years versus 0 years. Fourth, data was missing or unclear on the types of Language Instructional Education Program (LIEP) that ELs were enrolled in. Research cited in this study shows that LIEPs are an important factor in analyzing the rate and time to reclassification. Without this data, this study is missing an important variable in understanding the differences in reclassification of ELs. Finally, limiting the analysis

dataset to only those students for whom good data has been available can introduce concerns about eliminating problematic cases. Arguably, students with poor performance/attendance as well as schools with poor administrative practices constitute a systematically different subgroup – most likely clustering towards the end tail of performance scale.

Summary

The results of this study show time to proficiency varied significantly by grade band with students who entered in a grade other than kindergarten. Students who entered in Middle School were 119 percent more likely to become proficient than those who entered in kindergarten (median time to proficiency 3.1 years). Students who entered in Elementary School also had a higher likelihood of becoming proficient than those who entered in Kindergarten. Students who entered in grades 1-2 being 77 percent more likely to become proficient (median time to proficiency 3.3 years), and those who entered in 3rd to 5th grade being 90 percent more likely to become proficient (median time to proficiency 3.2 years). Female students were 29 percent more likely to become proficient than male students (median time to proficiency 3.7 years).

The following groups were less likely to become proficient within the time period analyzed. ELs who listed their ethnicity as Hispanic or Latino were 17 percent less likely to become proficient than ELs who listed their ethnicity as not Hispanic or Latino (median time to proficiency 4.3 years). ELs who listed their racial group as American Indian/Alaskan Native were 26 percent less likely to become proficient than students who listed other races (median time to proficiency 4.5 years). ELs who receive a free or reduced price lunch were 41 percent less likely to become proficient than those who did

not receive a free or reduced price lunch (median time to proficiency 5 years). ELs who were in special education programs were 56 percent less likely to become proficient than those who were not in special education programs (median time to proficiency 5.7 years).

On a final note, the results provide some perspective on impact of individual student and family characteristics. The results provide insights for identifying at-risk groups that could benefit from interventions, modification to curriculum and other resources. The interpretations of the findings, policy considerations and possible actions are addressed more deeply in the next chapter.

V. Discussion

This study examined time to reclassification and the impact of select student and family³ level factors on achieving proficiency in English for students classified as English learners (ELs) enrolled in grades K-12. The analytical sample of ELs from New Mexico included longitudinal data on two cohorts of ELs who started school in SY 2010-11 and SY 2011-12. Based on the data analyzed, ELs starting school in grades K-8 took on the average of four years to become proficient in English. ELs who entered school in grades 6-8 were 119 percent more likely to become proficient within the observation period than ELs who started in Kindergarten. ELs identified as Hispanic or Latino, those who received free or reduced lunch, as well as ELs receiving Special Education services had longer median times to proficiency than their peers. This chapter will focus on the results in response to the research questions and then branch out to more general policy implications and observations about reclassification of ELs. Relevant recommendations are noted throughout the chapter.

Three questions guided the study: (1) How long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED)? (2) Does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools? (3) How do select individual student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time? Below is a discussion of the findings for each question.

³ Variables such as poverty and race were considered as family based characteristics. These variables provide some indication of the family environment that the student may be living in. Variables such as grade, special education services were considered student level characteristics.

Question 1: How long does it take ELs to be reclassified as English proficient based on the criteria established by the New Mexico Public Education Department (NMPED)?

This study shows it takes an average of four years for an EL to be reclassified based on the assessments and criteria used by the State of New Mexico. This finding is within the range found in other studies mentioned in the literature review using longitudinal state level data on English proficiency assessments (Abedi (2008); Motamedi, Singh, & Thompson (2016); Abedi (2008); Kieffer and Parker (2016)). The median time ranged from 3.8 years Motamedi, Singh, & Thompson (2016) and between 4 and 5 years Abedi (2008). Generally speaking, the probability of reclassification did change over time, depending on when the EL started school. Overall, the of reclassification probability peaked in year four, before decreasing thereafter. This implies that the average time for ELs (included in this sample) to become proficient is within the first four years of entering the school. Research on long-term ELs (Olson, 2010; Thompson, 2015) shows a negative correlation between longer time to reclassification and academic achievement. This information is helpful for the State in setting long-term targets and short- term goals for the local and state accountability system. ESSA (NCLB, 2001, Title I, Sec 1111) requires States to set long-term expectations for the number of ELs that will reach proficiency within a set time frame. ESSA also requires that Sates establish short-term goals about how much progress ELs should make every year to reach proficiency. Based on the results, it is recommended that New Mexico set a target based on the finding that at least half of the ELs reach proficiency in four years.

Question 2: Does the probability of reclassification vary by grade cluster when first enrolled in the state's public schools?

Estimates from the analytical sample data by grade band (K-8), show that ELs who entered school in Kindergarten took longer to be reclassified than ELs in grades 1-8. In fact, the median years to reclassification for ELs, who entered in middle school (grades 6-8) is 3.1 years, as compared to 4 years for ELs who entered in Kindergarten. This finding is supported by literature that older children may acquire English faster than younger children (Cummins, 1981; Collier, 1987; Marinova-Todd, Marshall, & Snow, 2000; MacSwan and Pray, 2005). Haas, et al. (2015) also found that ELs in kindergarten had the lowest cumulative reclassification rates. However, Motamedi & Thompson (2016) found ELs who entered the schools in lower grades were reclassified in less time than those who enrolled in higher grades. Students who enrolled in kindergarten became proficient a year earlier than students who enrolled in grades 2–5. However, the study by Motamedi & Thompson (2016) was conducted with ELs enrolled in a very specific program, who may have come to school better prepared through pre-school programs. This study also found that the level of English proficiency at entry had a different association with time to reclassification, depending on the grade that students entered school.

Based on results from this study, the observed differences in the median rate to reclassification between Kindergarten and grade bands 1-2 (3.3 years), grades 3-5 (3.2 years) and grades 6-8 (3.1 years) is only 0.1 years. This difference may not have concrete implications for implementing reclassifications policy since ELs are reclassified only at the end of the school year, and based on the median differences, one half of ELs in grades

1-8 are ready to be reclassified after third year in school. One half of ELs who start in Kindergarten, will remain ELs for one more year, as compared to ELs who start in grades 1-8 and will be reclassified after 4 years.

Research works cited (Conger 2008; Conger 2009; Motamedi & Thompson 2016) show that starting proficiency levels also influence time to proficiency. As such, it is important to take the starting proficiency levels into account when interpreting the estimates for grade level effect. Better yet, starting proficiency level should be used as a covariate in regression-based analyses so that its impact could be parceled out. This study included a descriptive analysis of the distribution of ELs reclassified each year who started at beginning or intermediate proficiency levels embedded in grade bands. These analyses indicate that regardless of grade, ELs who started at lower proficiency levels took longer to be reclassified than peers who started at higher proficiency levels. This suggests that the model estimates in this data set may be influenced by the variations of proficiency levels in the grade bands. For example, in Kindergarten, a majority of ELs started at a lower starting proficiency level than ELs in upper grades. This may have contributed to Kindergarten ELs taking longer to achieve proficiency as compared to ELs starting in grades 1-8. Information on previous educational experiences of ELs was missing from the data set. Students start school with varying experiences and may have been exposed to school in previous grades but students starting school in Kindergarten who came from poverty and may have lacked the opportunity to attend preschool programs that generally prepare students for school. Thus, it is recommended that the State consider measuring both school readiness and English language proficiency (ELP) for incoming ELs in Kindergarten to evaluate the gaps in academic readiness between

ELs and their peers when school begins, and apply the appropriate early interventions to help mitigate the gaps.

Question 3: How do select individual student-level and family-level characteristics impact the probability of reclassification for students in the same grade cluster or who have been ELs for the same amount of time?

The first two research questions focused primarily on the effects of time and age on the probability of reclassification. The third question examined the effects of predictors other than time on ELs probability of meeting the state's reclassification criteria. As discussed in the previous chapter the effects on reclassification probability of all substantive predictors were statistically significant. Below is a discussion of other factors that may impact the time ELs require for reclassification. The variables include: gender; poverty, race and ethnicity; presence of a disability.

Gender

The study data show that girls were 29 percent more likely to be reclassified than boys, however, it should be noted that research focusing on gender differences in language acquisition are inconsistent. Findings from this study support research using other State data (Grissom 2004; Haas, et al., 2015, 2016 a, 2016b; Thompson, 2015) that found females have a slight advantage over males in attaining ELP. However, other studies (Collier, 1987) did not find any significant gender differences. Studies conducted using academic achievement assessments, such as, NAEP show that boys perform better in mathematics, while girls perform better in reading language arts. Several studies (Lapayese, Huchting, & Grimalt, 2014; McGraw, Lubienski, & Strutchens, 2006; Perie, Moran, & Lutkus, 2005; Tong, Irby, Yoon, & Masthes 2010) found disparities by gender

in achieving ELP and recommended additional research to understand the disparities of underachievement among Latino boys. The data analyzed for this study did show disparity in ELP achievement by gender with girls having a slight advantage over boys. Recognizing the vital role of language proficiency on academic achievement, it is recommended that New Mexico examine the gender differences in EL achievement.

Poverty, Race and Ethnicity

This study shows that time to reclassification varied across the various racial groups. The findings indicate Hispanic and Native American/Alaska Native ELs are less likely to achieve proficiency compared to peers in other ethnicities. This should be interpreted with caution because race or ethnicity of a student is not, in and of itself, a contributor for learning English at a faster or slower rate. There is no causal relationship that indicated if an EL is Hispanic or Native American/Alaska native the student will take longer to achieve proficiency. There are a number of other factors at play, such as disadvantaged homes and community environments, economic status, lack of opportunities and bias (Harry & Klingner 2007).

In this study, poverty has emerged as an important indicator of performance. The demographic distribution of the analytical sample showed a majority of ELs are Hispanic and come from poverty. The study estimates show ELs experiencing poverty are 41 percent less likely to become proficient than their peers. This finding supports the research (Haas et.al (2015, 2016a, 2016b); however, these studies also found that the differences in reclassification rates between ELs from poverty and their peers narrowed after 5 years. Since the observation period for this study ended after five years for Cohort 2, the long-term impact on the gap between ELs from poverty and their peers was not

observed. The narrowing of the performance gap suggests that achievement ELP could be confounded by academic achievement. Furthermore, the theory of language transfer explains the influence of the first language (L1) and how it interacts in some way with the acquisition of the second language (L2). Understanding the underlying organizational principles of L1 and having a metalinguistic awareness of that knowledge can facilitate the learning of L2 by transference of cognitive and language skills.

Cummins (1981) theorized that transfer of academic skills across languages does not occur automatically; academic language skills must be taught explicitly using both languages. Noting the interdependence of language and language related academic content matter, MacSwan and Rolstad (2005) proposed an approach to transfer theory drawing on neurocognitive research. In this approach, language and the conceptual understanding of school subject matter is specifically differentiated. Conceptual knowledge is distinct and independent from linguistic knowledge MacSwan and Rolstad (2005). Multilingual speakers may learn conceptual knowledge through various languages and apply all of it to learning academic English. ELs access academic concepts through the languages they know including language they may have acquired in their communities. ELs from poverty may not have the same opportunities to learn conceptual knowledge, thereby impacting performance on ELP assessments that measures academic language. Schools continue to use standardized testing as an essential basis for major school reform while research clearly shows that there are weaknesses in the content assessments (Abedi, 2002) and English language proficiency assessments (MacSwan & Rolstad, 2006).

Under ESSA, state and local school authorities have more flexibility to design the educational plan for their students. Thus, it is recommended that New Mexico consider using the results of the standardized assessments to continuously inform teachers and families about EL achievement relative to their peers. Additionally, the State should consider empowering staff at the local school and district level to utilize the resources, and use multiple data points with particular attention to mitigating the impact of poverty resulting in increasing and enhancing ELP and EL achievement.

It should be noted that ELs from poverty might demonstrate lower performance on ELP and academic achievement than their EL peers not experiencing poverty. However, poverty should not be the predictor of students' educational destiny. Reardon and Galindo (2007) also found that longitudinally, there was no measurable difference in academic achievement among ELs from poverty and the comparison groups overtime as ELs gained English language proficiency. The amount of money that a family has, or the ethnicity of a child, should not influence the opportunities afforded to achieve their full potential. Since the passage of the Elementary and Secondary Education Act (ESEA) in 1965, federal funds helped low-income students in educational programs such as Title I. However, data show a persistent academic gap between ELs and their peers, which demonstrates that schools may lack an understanding of the root causes of this gap and how to identify the resources needed for programs (e.g., bilingual classes and assessment). Schools may also lack resources to provide services (e.g., bilingual counselors and other wrap around supports) needed to help ELs succeed and close the gap.

ELs with disabilities

The study found that EL with disabilities (ELSWD) had the longest median time to reclassification, with a reclassification median time of 5.7 years compared to the 4 year median reclassification rate of their EL peers without disabilities. Such low numbers of ELSWD reclassifying as English proficient indicates that the ELSWD are the most academically at-risk students, performing at the lowest levels in English proficiency. Moreover, in this study, ELSWD represented approximately 10 percent of total analytical sample but were represented in higher numbers (15 percent) in the subset of ELs at risk of becoming long-term ELs. Elimination of this potential confounder may require application of weights to balance the representation of ELSWD to a standard set of benchmarks. Understanding the time to proficiency of ELs with severe cognitive disabilities is an area that could benefit from further research. Overall, the findings in this study are supported in literature showing ELs with disabilities are at risk of becoming long-term ELs. Haas et.al (2015, 2016a, 2016b) and Kieffer & Parker, (2016) found that ELSWD are less likely than those without disabilities to be reclassified, resulting in large proportions of dually identified students in public schools.

There are multiple factors that may impact the reclassification rates of ELSWD; the most important of which is the appropriateness of the assessments used to measure ELP. Students classified ELSWD can fall under any one of the thirteen disability categories specified in the 2004 Individuals with Disabilities Education Act (IDEA). Each disability category exhibits potential difficulties in many different areas and manifests a unique set of characteristics that can impact the measurement of language skills. Thus, no single ELP measure can account for the unique characteristics of each

disability. Related to this, there could be varying programs available for ELSWD across different schools, further contributing to this extraneous variability.

Determining reclassification criteria for ELSWDs is particularly complex, especially when a student's language production and comprehension are affected by a disability. New Mexico uses a composite score including weighted values of the four domains of listening, speaking, reading and writing. WIDA does not provide an overall composite score without all four domain scores and weights the domains of reading and writing more than listening and speaking. For example, reading and writing disabilities are considered specific learning disabilities, defined by IDEA (2004) as disorders in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write or spell. An EL with a reading and writing disability is at a disadvantage when reading and writing domains are weighed more than the other domains. The IDEA (2004) defines "speech or language impairment" as a communication disorder, such as stuttering, impaired articulation, language impairment, or a voice impairment that adversely affects a child's educational performance. ELs "speech or language impairment" may demonstrate lower-than-expected oral language performance on the ELP test resulting in lower overall proficiency scores.

The statutory definition of EL (cited in the first chapter) states that the deciding factor in making reclassification decisions for all ELs is to meet the state's established proficiency level on state English language proficiency (ELP) assessments. Additionally, The Department of Justice and Office of Civil Rights' (U.S. Department of Education 2015) issued guidance related to valid and reliable assessment of English proficiency for

all ELs, stating: “the English language proficiency assessment must meaningfully measure student proficiency in each of the language domains, and, overall, be a valid and reliable measure of student progress and proficiency in English” (p. 33). While states are required to provide accommodations and alternate ELP assessments as determined by a student’s IEP team (Section 612, IDEA, 2004), states cannot develop a different set of ELP standards or achievement standards for English learners with disabilities (ELSWD). ELSWD receive accommodations designed to offset challenges resulting from a student’s disability but the accommodation cannot invalidate the test measures. The testing manuals published by New Mexico describe the attributes of the recommended users of an accommodation but do not index the specific IDEA disability category or categories associated with the user (e.g., specific learning disabilities). This lack of guidance on suitable accommodations by specific disability may contribute to the variation among the schools on how these accommodations must be chosen and implemented for each individual student.

IEP teams may sometimes allow ELSWDs to omit parts of the ELP test related to their disability (e.g., the listening and speaking subtests if they are deaf). States have a legal obligation to create a valid and reliable record of ELs’ proficiency in English, and ED’s regulations require that for an EL who is not able to participate in all domains of an English proficiency test because there are no appropriate accommodations for the affected domain, an overall score must be provided from the remaining domains, 34 CFR 200.6(h)(4)(ii). In some states, students who do not participate in all four domains are not assigned composite scores or performance levels. This is no longer permitted under the ED regulations, which, effectively require that for an EL who was not able to

participate in all domains of an English proficiency test, a composite score be created that reflect only the tested domains. There is no clear consensus on how to generate a composite score for ELs whose disability necessitates their selective participation in the ELP assessment. The Department of Justice and Office of Civil Rights' in a letter has states that to demonstrate proficiency on the ELP assessment, States can use either a conjunctive score (minimum proficiency scores in each language domain) or a composite score (derived from weighting the domain scores). The conjunctive scores or composite score must "overall, be a valid and reliable measure of student progress and proficiency in English" (U.S. Department of Education, 2015 p. 33).

To make language proficiency assessments valid and reliable for all ELs and particularly ELSWD is to develop organic assessments that adhere to Universal Design for learning (UDL) principles. The underlying belief of UDL principles is that curricula, instruction, and assessment must account for individual variation across all learners, rather than adopting an inflexible "one-size-fits all" approach designed for English proficient students without disabilities (CAST, 2011). Provided the same construct is measured, computerized testing formats may facilitate UDL. The cohorts in this study began school in SY 2010-11 or SY 2011-12 and were administered the old ACCESS but the new version of the test, ACCESS 2.0, applies UDL principles to test items during the development phase to move them from paper formats to an online form. As part of this process, the WIDA team tries to ensure that the items are user friendly and balance the accessibility enhancements with usability concerns. For example, all information students need to answer a question is on the screen; navigation components always appear in the same place on the screen; and stimulus pictures and text, item stems, and response

options appear in predictable locations, with limited variation allowed to accommodate differences in text length, number of response options, and degree of graphic support (World-Class Instructional Design and Assessment 2017-18). While computerized testing formats based on UDL design principles may result in more valid assessments, the State may encounter other challenges. For example, data showed 90 percent of ELs in the data sample came from poverty. ELs may experience a lack of opportunity to work with a computer at home due to the family socioeconomic status. Additionally, schools may experience technological issues with school computer hardware and software that may impact ELP assessment outcomes.

New Mexico does not provide a composite score without all four domain scores. States will likely require technical assistance to ensure that LEAs can generate valid and reliable assessment composite scores for students who have missing domain scores on the ELP assessment due to their disability. It is recommended that New Mexico should consider developing policy and guidance related to the reclassification of ELs with disabilities. The guidance should align with students who are dually identified and to their disabilities. A “one-size-fits all” approach to assessment and instruction would not be appropriate. For example, using alternate methods of assessment such as portfolios may be more suitable for ELs whose standard ELP test cannot yield valid results. To improve outcomes for ELs with disabilities specific disability characteristics should be carefully considered in both the assessment and instruction of these students. Another recommendation is for the State is to improve the guidance and training for teachers on the use of accommodations for each disability category for ELs. Accommodations help improve ELSWD access academic content and assessments. Without the appropriate

accommodations ELSWD also may have challenges accessing assessments. The guidance published by the State should also provide detailed information about appropriate and specific accommodations for particular disabilities. Teachers may benefit from precise information about which accommodation is suitable for each specific disability.

Long-term ELs

The ELP measure used by New Mexico, the ACCESS, assesses the language of content areas. Developing language and academic skills is essential to academic success. The test results indicate that half the ELs take 4 years to reach the proficiency. This study highlighted that there were enough students at risk of becoming long-term ELs that require different, specialized, and targeted instruction to expedite proficiency.

Researchers (Olson, 2010; Thompson, 2015) cited potential adverse consequences to ELs who remain in a limited English proficient status for extended periods of time (e.g., more than 5-6 years). For example, they lack oral and literacy skills needed to master academic content at the middle and high school levels. Often LTELs do not receive adequate, appropriate, and specialized instruction needed for both academic achievement and engagement in content classes. Often students at the middle and high school levels receive interventions that may supplant other coursework resulting in delaying on-time high school graduation.

Thus, it is important to focus on ELs who are not reclassified as proficient within a 4-5 year time frame especially if they start at lower proficiency levels. Prematurely exiting ELs from an English language development program may have detrimental effects because students may stop receiving the additional language support needed to access academic content.

As a group, long-term ELs are more likely to experience academic failure than their peers. Menken et al., (2012) found that even simply increasing awareness about this student population among educators seems to positively impact their educational outcomes. For the purpose of this study ELs in both cohorts not reaching proficiency after five years were included in this subset.

This study found that approximately one-third of ELs in the data sample were not reclassified as proficient after five years and were in danger of becoming long-term ELs. A large proportion of potential long-term ELs were classified as Hispanic or Latino (78.37 percent) and experienced poverty as identified by the free or reduced priced lunch program (94.97 percent). This demographic is not surprising given a majority of Hispanic ELs in this study were also experiencing poverty, however the proportion of Hispanic and ELs from poverty in the set of long-term ELs was higher as compared to this population in the sample. The negative impact of poverty on academic achievement is well documented (Keiffer, 2008). However, these results must be interpreted with caution because other studies found that language proficiency, not ethnicity, was a greater indicator of student success in academic achievement than poverty Reardon and Galindo (2007).

Among the pool at risk of becoming long-term ELs, 15 percent were also identified as students with disabilities. Literature cited in this study (Haas et.al, 2015, 2016 a, 2016 b; Kieffer & Parker 2016) report ELs with disabilities took longer to be reclassified and are at a higher risk of becoming long-term ELs than their peers. In this study, ELs with disabilities represented approximately 10 percent of total analytical sample but were represented in higher numbers (15 percent) in the subset of ELs at risk

of becoming long-term ELs. Representation of ELs with disabilities in the sub set of ELs at risk of becoming long-term ELs could have possibly been underestimated because ELs with severe cognitive disabilities, who took the alternate language proficiency test and were not included in this data sample. Overall, the findings in this study are supported in literature showing ELs with disabilities are at risk of becoming long-term ELs.

Thompson (2015) cautions that using the term long-term ELs implies that it is the student's responsibility to achieve proficiency in a certain timeframe and failure to do is a lack of ability on the part of the student. The fact that certain ELs do not achieve proficiency in a reasonable timeframe may be an aspect of a faulty classification system at work (Thompson, 2015; Rolstad, MacSwan, & Guzman, 2015) or quality of programs and services afforded to ELs (Menken, Kleyn & Chae, 2012). Long-term ELs may have oral language proficiency but are usually behind their peers in literacy skills. Not achieving proficiency is exacerbated because of inconsistent and haphazard service delivery and educational programming year-from-year (Menken, Kleyn & Chae, 2012). To illustrate the inconsistent provision of educational programs for ELs Menken, Kleyn & Chae, (2012) exemplify the experience on a student who in history class received bilingual instruction but in seventh grade she was switched to English only instruction. In sixth grade math, she received English only instruction for part of the year but bilingual instruction for the rest of the year and then English only again in seventh grade. Additionally, many school districts may not have programs especially designed to meet their unique needs. Thus, it is recommended that the State publish a clear definition of long-term ELs, and conduct a more in-depth analysis of their long-term EL population to determine the growth trajectories by individual schools, and service delivery models to

identify consistent, targeted, and evidence-based educational programs and services.

This is aligned with new ESSA mandates that starting with SY 2017-18, States must to report the number of ELs who have not become English proficient after five years.

Conclusions

This section summarizes recommendations in two areas: data collection and use at the local district and State level, and general reclassification policies.

Data collection and use at the local district and State level

Under NCLB and ESSA States strived to implement standardized policies for EL education. Under ESSA States have to use uniform statewide procedures for reclassification. New Mexico purposefully developed and implemented centralized, consistent and uniform criteria for determining if a student meets the definition of proficiency established at the State's ELP assessment to ensure some degree of equity in services across the State. For example, regardless of the school or district ELs are guaranteed their EL classification status and can continue to receive services or remain reclassified. To continue this progress, it is recommended that the State evaluate the academic effects of the reclassification policies. While ESSA mandates States monitor former ELs for four years, the State should consider monitoring ELs who are reclassified throughout their school enrollment to ensure that ELP is not impacting academic success.

Variability in instructional programs and supports for ELs is an important contributor to understanding the variations among students time to proficiency (Menken, Kleyn & Chae, 2012). However, lack of EL enrollment data on the programs resulted in a key analytical impediment. It is recommended that the State establish and implement clear policies and guidance on the program models to ensure equity in services and

improve programs and services for ELs in every school and district across the state. Additionally, it is recommended that the State modify data collection methods and system to record the ELs' educational services or program, initial proficiency level and entry date into the New Mexico school system. This would provide the necessary analytical platform to examine the accuracy the length of time to proficiency.

Despite the limitations of the data available, it is recommended that New Mexico consider replicating the current analysis for additional years to fulfill ESSA requirements for monitoring statewide improvement in LIEP programs. If the median time to proficiency decreases for future cohorts, this would indicate increased programmatic effectiveness in supporting EL students to proficiency. In summary, a deeper analysis of other factors, such as, the previous educational experiences, starting English proficiency levels, quality of educational services and programs is needed to provide a clearer interpretation of the data on ELs entering high school

Reclassification Policy

Reclassification is an important milestone in the academic journey of ELs. Once reclassified, ELs must access the content in mainstream classes without specialized language instructional services or linguistic assessment accommodations. In policy decisions, the relative effectiveness of programs and services is gauged by how quickly these programs reclassify ELs and move them to mainstream classes. Reclassification rates are a proxy for program quality and effectiveness of the teaching and learning in the classroom. Used in accountability systems, the rate of attainment of proficiency or reclassification may have some unintended negative consequences. Schools may not want to enroll ELs who impact the accountability negatively by lower academic scores or

longer exit rates. States may use lower exit criteria scores to boost the reclassification rates.

It is important to look carefully behind the reclassification rate, and consider theoretical issues. The initial classification of a student as an EL, and reclassification as a proficient speaker of English rests on the construct of ELP. However, researchers view the construct of ELP from different perspectives: What marks a speaker as proficient in a language? Which matters more, accuracy or fluency? In what contexts should an individual be able to demonstrate proficiency in the language?

A student may use English to communicate in the home and community but is marked as an English learner in school, signaling that the student has yet to master the academic language needed to succeed in the content classes. Many native English speakers may also have low proficiency in the content academic language, but are not coded as learners of (academic) English. If native speakers were administered the English language proficiency (ELP) tests, they may as well be classified as ELs too. This calls into question the interpretation and implementation of the policies and procedures used to identify and exit EL students. Many ELP tests use discrete point language skills, while emerging second language acquisition theories suggest that language competence does not progress from one step of a sequence to the next in an orderly fashion. Many ELP tests use a standardized approach of setting cut points that reflect progression in acquiring English. However, no one test can account for the considerable individual variability in features of learners' interlanguage while progressing from one stage to the next.

New Mexico, like many other States uses the ACCESS to measure ELP. The ACCESS 2.0 is part of the next generation ELP assessments developed to align better

with college and career-ready standards, lending hope that the resultant reclassification scores provide valid inferences of ELP in academic settings. Nonetheless, no amount of well-designed item types will address the shortcomings of erroneous classifications. Language assistance programs based on ELP standards are designed on the belief that language proficiency drives academic competence; that is -- as ELs become more proficient in English, achievement in the content areas will increase. However, the research and our understanding of the relationship between ELP and academic achievement continue to evolve. We know there is a correlation or association between the two competencies but a direct causal relation has not been established. Additionally, content assessments and ELP assessments normed on dissimilar populations provide inadequate information about the relationship between ELP and content achievement of ELs. Researchers may wish to design studies that focus on obtaining a better understanding of the relationship between ELP and academic achievement.

Ever since the landmark Supreme Court decision on *Brown v. Board of Education*, researchers focus on “between-group” differences in terms of academic achievement, which resulted in policies and practices designed to reduce the “between-group” differences in educational achievement Ramirez and Carpenter (2005). However, little attention is directed to “with-in” group differences that are as important as those between groups and may be more relevant in determining how to narrow the achievement gap between groups. This is especially true of the EL population because ELs are a heterogeneous group and yet policies reflect a one size fits all approach. This study showed significant difference in time to proficiency between ELs experiencing poverty and their peers, difference by grade and starting proficiency level, difference between

ELs with special needs and their peers, difference by gender and ethnicity. Additional longitudinal studies are needed to understand how to better interpret the demographic and individual student difference on time to reclassification.

Also, data show a great deal of variability in reclassification rates based on individual student factors. There may also be the same variability among the achievement of ELs after reclassification. Some students may experience academic success and some may not. Using broad simplistic categories of EL and former EL mask an enormous amount of variation among the students within each category. Educators must identify processes and strategies for overcoming the ‘services or no services’ dichotomy that often hampers reclassification policies and decisions. There is an urgent need to create a flexible system that is responsive to the changing needs of EL. It is recommended that States decouple services based on classification and reclassification that focus on “services or no services’ and evaluate them based on need so ELs have full access to the needed resources and services.

The study results show approximately 33 percent of ELs take longer than five years to meet the State’s definition of proficiency. A single-minded focus on attainment of the proficiency standards ignores the “opportunity gaps” that may exist among schools. Researchers (Ramirez & Carpenter 2005; Kim, 2011), point to disparities among schools in factors such as school funding, class size, and the percentages of credentialed teachers. To overcome these disparities, it is recommended that Federal, state and local policies focus on classroom- and school-level factors that shape the experiences of ELs. Additionally, researchers found that although language support services were typically designed with newly arrived immigrants in mind, one-third to

one-half of ELs in secondary school is actually long-term ELs (Menken & Kleyn, 2010; Olsen, 2010). Thompson (2015), cautions against using the label Long-Term ELs, because it may blind educators to students' abilities. Many, not all, ELs struggle and take longer to achieve proficiency. This does not mean that all ELs classified as such for five or more years have low literacy skills nor are passive, disengaged learners. More research is needed to understand the root causes and structural forces that contribute to unequal outcomes. First, and foremost, we all must recognize and foster the unique assets, interests, talents, and cultural experiences that ELs contribute to our society, country, workplace and schools.

References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometric issues. *Educational Assessment*, 8:231–257.
- Abedi, J. (2008). Classification system for English language learners: issues and recommendations. *Educational Measurement: Issues and Practice*, 27(3):17–31.
- Abedi, J., & Gandara, P. (2008). Performance of English language learners as a subgroup in large scale assessments: Interaction of research and policy. *Educational Measurement: Issues and Practice*. 25(4), 36–46.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. AC-19, 716–723
- Allison, P. D., & SAS Institute. (1999). Logistic regression using the SAS system: Theory and application. Cary, N.C: SAS Institute
- August, D., & Hakuta, K. (1997). Improving schooling for language-minority children: A research agenda. Washington, DC: National Academy Press.
- August, D., & Shanahan, T. (2006). Developing Literacy in Second-language Learners: Report of the National Literacy Panel on Language Minority Children and Youth. Mahwah, NJ: Lawrence Erlbaum.
- Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing*, 19, 453-476.
- Bailey, A. (Ed.). (2006). The language demands of school: Putting academic English to the test. New Haven, CT: Yale University Press..

- Bailey, A.L., & Kelly, K.R. (2013). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy*, 27(5):770–804. Retrieved from <http://doi.org/10.1177/0895904811432137>.
- Boals, T., Kenyon, D. M., Blair, A., Cranley, M. E., Wilmes, C., Wright, L. J. (2015). Transformation in K-12 English Language Proficiency Assessment: Changing Contexts, Changing Constructs. *Review of Research in Education*, v39 n1 p122-164.
- Butler, Y.G., Orr, J.E., Bousquet Gutierrez, M., & Hakuta, K. (2000). Inadequate conclusions from an inadequate assessment: What can SAT-9 scores tell us about the impact of Proposition 227 in California? *Bilingual Research Journal*, 24(1-2):141–154. Retrieved from <http://eric.ed.gov/?q=Inadequate+Conclusions+from+an+Inadequate+Assessment%3a+What+Can+SAT9+Scores+Tell+Us+about+the+Impact+of+Proposition+227+in+California%3f&id=EJ623772>.
- Callahan, R. M. (2013). The English learner dropout dilemma: Multiple risks and multiple resources (Report No. 19). Santa Barbara, CA: University of California Santa Barbara, Gevirtz Graduate School of Education, California Dropout Research Project. Retrieved from <http://www.cdrp.ucsb.edu/download.php?file=researchreport19.pdf>.
- Carroll, P.E., & Bailey, A.L. (2016, January). Do decision rules matter? A descriptive study of English language proficiency assessment classifications for English-language learners and native English speakers in fifth grade. *Language Testing*, 33(1):23–52. Retrieved from <http://dx.doi.org/10.1177/0265532215576380>.

- Castañeda v. Pickard (1981). 648 F.2d 989, 5th Cir.
- Cheung, A., & Slavin, R.E. (2012). Effective reading programs for Spanish-dominant English language learners (ELLs) in the elementary grades: a synthesis of research. *Review of Educational Research*, 82(4):351–395.
- Collier, V.P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21:617–641.
- Collier, V.P., Wayne P. T. (1989). How quickly can immigrants become proficient in school English? *Journal of Educational Issues of Language Minority Students* 5:26-3.
- Conger, D. (2008). Testing, time limits, and English learners: Does age of school entry affect how quickly students can learn English? *Institute for Education and Social Policy, IESP Working Paper #08-04*. Retrieved from http://steinhardt.nyu.edu/iesp/archived_working_papers.
- Conger, D. (2009). Testing, time limits, and English learners: Does age of school entry affect how quickly students can learn English? *Social Science Research*, 38(2):383–396. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/19827181>.
- Conger, D. (2010). Does bilingual education interfere with English-language acquisition? *Social Science Quarterly*, 91(4):1103–1122.
- Conger, D., Hatch, M., McKinney, J., Atwell, M.S., & Lamb, A. (2012, January). Time to English Proficiency for English Language Learners in New York City and Miami-Dade County. Institute for Education and Social Policy, IESP Policy Brief #01-12. Retrieved from <http://steinhardt.nyu.edu/iesp/briefs>.

- Cook, G H. (2007). Alignment Study Report: The WIDA Consortium's English Language Proficiency Standards for English Language Learners in Kindergarten through Grade 12 to ACCESS for ELLs® Assessment. Retrieved from www.wida.us/get.aspx?id=193
- Cook, H.G., Boals, T., Wilmes, C., & Santos, M. (2008). Issues in the Development of Annual Measurable Achievement Objectives for WIDA Consortium States (WCER No. 2008–2). Wisconsin Center for Education Research. Retrieved from <http://eric.ed.gov/?id=ED501340>.
- Cook H. G., Boals T., Lundberg L. (2011). Developing informed expectations for the academic achievement of English learners: What can we reasonably expect? *Phi Delta Kappan*, 93(3), 66–69.
- Cook, H.G., Linquanti, R., Chinen, M., & Jung, H. (2012). National Evaluation of Title III Implementation Supplemental Report: Exploring Approaches to Setting English Language Proficiency Performance Criteria and Monitoring English Learner Progress. Office of Planning, Evaluation and Policy Development, U.S. Department of Education. Retrieved from <http://eric.ed.gov/?id=ED531979>.
- Collier, V. P. (1995). Promoting academic success for ESL students; Understanding second language acquisition for school. Trenton, N.J.: New Jersey Teachers of English to Speakers of Other Languages-Bilingual Educators.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. *Schooling and Language Minority Students: A Theoretical Framework*, pp. 3–49. Retrieved from <http://eric.ed.gov/?id=ED249773>.

- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, UK: Multilingual Matters.
- Cummins, J., & Nakajima, K. (1987). Age of arrival, length of residence, and interdependence of literacy skills among Japanese immigrant students. In Harley, B., Allen, P., Cummings, J., and Swain, M. (Eds.). *The Development of Bilingual Proficiency: Final Report Volume III: Social Context and Age* (pp. 183–202). Toronto, Canada: Modern Language Centre, Ontario Institute for Studies in Education. (ERIC Document Reproduction Service No. ED291248)
- ESSA (2015). Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).
- Genesee, F. (2006). *Educating English language learners: A synthesis of research evidence*. Cambridge, UK: Cambridge University Press.
- Graham, J. G. (1987). English language proficiency and the prediction of academic success. *TESOL Quarterly*, 21(3), 505-521.
- Greenberg Motamedi, J. (2015). Time to reclassification: How long does it take English learner students in Washington Road Map districts to develop English proficiency? (REL 2015–092). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Greenberg M. J., Singh, M., & Thompson, K. D. (2016). English learner student characteristics and time to reclassification: An example from Washington state (REL 2016– 128). Washington, DC: U.S. Department of Education, Institute of

- Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Government Accountability Office. (2001). Meeting the needs of students with limited English proficiency. (GAO Publication No. 01-226). Washington, D.C.: U.S. Government Printing Office.
- Grissom, J.B. (2004a). Reclassification of English learners. *Education Policy Analysis Archives*, 12(36). Retrieved from <http://epaa.asu.edu/epaa/v12n36/>.
- Grissom, J.B. (2004b). Age and achievement. *Education Policy Analysis Archives*. Retrieved from <http://dx.doi.org/10.14507/epaa.v12n49.2004>.
- Gwynne, J., Pareja, A S., Ehrlich, S.B., & Allensworth, E. (2012). What matters for staying on-track and graduating in Chicago public schools: a focus on English language learners. *Consortium on Chicago School Research*. Retrieved from <http://eric.ed.gov/?q=ED532513&id=ED532513>.
- Haas, E., Huang, M., Tran, L., & Yu, A. (2016a). *The Achievement Progress of English Learner Students in Utah* (REL 2016–155). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Haas, E., Huang, M., Tran, L., & Yu, A. (2016b). *The Achievement Progress of English Learner Students in Nevada* (REL 2016–154). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education

Evaluation and Regional Assistance, Regional Educational Laboratory West.

Retrieved from <http://ies.ed.gov/ncee/edlabs>.

Haas, E., Tran, L., Huang, M., & Yu, A. (2015). *The Achievement Progress of English Learner Students in Arizona* (REL 2015–098). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.
Retrieved from <http://ies.ed.gov/ncee/edlabs>.

Hakuta, K., Butler, Y., & Witt, D. (2000). *How Long Does It Take English Learners To Attain Proficiency?* Santa Barbara, CA. University of California, Linguistic Minority Research Institute.

Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly*, 27(1):1–20. Retrieved from <http://dx.doi.org/10.1016/j.ecresq.2011.07.004>.

Hammer, C.S., Lawrence, F.R., & Miccio, A.W. (2008). The effect of summer vacation on bilingual preschoolers' language development. *Clinical Linguistics & Phonetics*, 22:686–702.

Harry, B. & Klingner, J., (2006). Why are so many minority students in special education?: Understanding race and disability in schools. New York: Teachers College Press, Columbia University.

Harry, B., & Klingner, J. (2007). Discarding the Deficit Model. *Educational Leadership*, 64(5), 16-21

- Harwell, M., & LeBeau, B. (2010). Study of eligibility for free lunch as a SES measure in education research. *Educational Researcher*, 39(2), 120-131.
- Hill, L.E., Weston, M., & Hayes, J.M. (research support: Ezekiel, D., & Chavez, Belen). (2014). *Reclassification of English Learner Students in California*. Public Policy Institute of California. Retrieved from <http://www.ppic.org/main/publication.asp?i=1078>.
- Individuals With Disabilities Education Act, (IDEA) 20 U.S.C. § 1400 (2004).
- Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100(4):851–868. Retrieved from <http://eric.ed.gov/?id=EJ823716>.
- Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Researcher*, 39(6):484–486. Retrieved from <http://dx.doi.org/10.3102/0013189X10378400>.
- Kieffer, M.J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal*, 48(5):1187–1225. Retrieved from <http://dx.doi.org/10.3102/0002831211419490>.
- Kieffer, M. J. & Parker, C. E. (2016). *Patterns of English learner student reclassification in New York City public schools* (REL 2017–200). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

- Kim, J. (2011). *Relationships Among and Between Ell Status, Demographic Characteristics, Enrollment History, and School Persistence* (CRESST Report No. 810). National Center for Research on Evaluation, Standards, & Student Testing. Retrieved from <http://eric.ed.gov/?id=ED527529>.
- Klesmer, H. (1993) Development of ESL Achievement Criteria as a Function of Age and Length of Residence in Canada. A Special Leave Project for The Board of Education for the City of North York. Retrieved from <http://files.eric.ed.gov/fulltext/ED408862.pdf>
- Lapayese, Y., Huchting, K., & Grimalt, O. (2014). Gender and Bilingual Education: An Exploratory Study of the Academic Achievement of Latina and Latino English Learners. *Journal of Latinos and Education*, 13 (2), 152-160.
- Lau v. Nichols, 414 U.S. 563 (U.S. Supreme Court January 21, 1974).
- Legislative Analyst's Office. (2004). A Look at the Progress of English Learner Students. Sacramento, CA: Author. Retrieved from http://www.lao.ca.gov/2004/english_learners/021204_english_learners.htm.
- Lindholm-Leary, K., & Hernandez, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: native English speakers, fluent English/previous ELLs, and current ELLs. *Journal of Multilingual and Multicultural Development*, 32(6):531–545. Retrieved from <http://dx.doi.org/10.1080/01434632.2011.611596>.
- Linquanti, R. (2011). Strengthening assessment for English learner success: How can the promise of the common core state standards and innovative assessment systems be realized? In D. Plank and J. Norton (Eds.), *The road ahead for state*

- assessments (pp. 13-25). Palo Alto & Cambridge: Policy Analysis for California Education and Rennie Center for Education Research & Policy.
- Linguanti, R., & Cook, H. G. (2013). Toward a “common definition of English learner”: A brief defining policy and technical issues and opportunities for state assessment consortia. Washington, DC: Council of Chief State School Officers.
- MacSwan, J., & Pray L. (2005). Learning English bilingually: age of onset of exposure and rate of acquisition among English language learners in a bilingual education program. *Bilingual Research Journal*, 29(3):653–678. Retrieved from <http://www.terpconnect.umd.edu/~macswan/MacSwan&Pray2005.pdf>.
- MacSwan, J., & Pray, L. (2005). Learning English bilingually: Age of onset of exposure and rate of acquisition of English among children in a bilingual education program. *Bilingual Research Journal*, 29(3), 687-712.
- MacSwan, J. & Rolstad, K. (2005). Modularity and the facilitation effect: Psychological mechanisms of transfer in bilingual students. *Hispanic Journal of the Behavioral Sciences*, 27(2), 224-243.
- MacSwan, J., & Rolstad, K. (2006). How Language Proficiency Tests Mislead Us About Ability: Implications for English Language Learner Placement in Special Education. *Teachers College Record*, 108, 11, 2304-2328.
- MacSwan, J., Rolstad, K., & Glass, G. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-LAS Español. *Bilingual Research Journal*, 26(2):213–238.

- Mahoney, K.S., & MacSwan, J. (2005). Re-examining identification and reclassification of English Language Learners: a critical discussion of select state practices. *Bilingual Research Journal*, 29(1):31–42.
- Marinova-Todd, S.H., Marshall, D.B., & Snow, C.E. (2000). Three misconceptions about age and L2 learning. *Tesol Quarterly*, 34(1):9–34. Retrieved from <http://eric.ed.gov/?ID=EJ603514>.
- McGraw, R., Lubienski, S.T., & Strutchens, M.E. (2006). A closer look at gender in NAEP mathematics achievement and affect data: Intersections with achievement, race and socio-economic status. *Journal for Research in Mathematics Education*, 37(2):129–150. Retrieved from <http://eric.ed.gov/?id=EJ765478>.
- Motamedi, J.G., Singh, M., & Thompson, K.D. (2016). *English Learner Student Characteristics and Time to Reclassification: An Example from Washington State* (REL 2016– 128). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Mulligan, G., Halle, T., & Kinukawa, A. (2012). *Reading, Mathematics, and Science Achievement of Language-Minority Students in Grade 8*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Issue brief, NCES No. 2012–028. Retrieved from <http://eric.ed.gov/?id=ED531335>.

- National Academies of Sciences, Engineering, and Medicine (NAS). 2017. *Promoting the Educational Success of Children and Youth Learning English: Promising Futures*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/24677>.
- New Mexico Public Education Department (2017). *Assessment Accountability, Evaluation*. Retrieved from http://www.ped.state.nm.us/ped/Assessment_index.html.
- New Mexico Public Education Department. (2017). *Bilingual Multicultural Education Bureau*. Retrieved from http://ped.state.nm.us/ped/Bilingual_index.html.
- New Mexico Public Education Department. (2013). *Student Assessment Accommodations Manual*. Retrieved from www.ped.state.nm.us.
- New Mexico Public Education Department. (2011a). *Student Assessment Accommodations Manual*. (Technical Report: Version 7.13.2011[36]). Retrieved from <http://www.ped.state.nm.us/AssessmentAccountability/AssessmentEvaluation/dl11/1/Accommodations%20Manual%20%209.1.11.pdf>.
- New Mexico Public Education Department. (2011b). *New Mexico Standards Based Assessment: 2010-11 Technical Report*. Retrieved from http://ped.state.nm.us/AssessmentAccountability/AssessmentEvaluation/dl11/2010_11_NM_SBA_Tech_Report.pdf.
- New Mexico Public Education Department. (1990). *New Mexico State Board of Education's Consolidating Initiatives for Tomorrow's Education (CITE) Plan*. Santa Fe, NM. No Child Left Behind Act (NCLB) of 2001, Pub. L. No. 107-110 § 115, Stat. 1425 (2001).

- Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational Opportunity for California's long term English learners*. Long Beach, CA: Californians Together.
- Parker, C. E., Louie, J., & O'Dwyer, L. (2009). New measures of English language proficiency and their relationship to performance on large-scale content assessments (Issues & Answers Report, REL 2009–No. 066). Washington, DC: Institute of Education Sciences.
- Parrish, T., Perez, M., Merickel, A., & Linqunti, R. (2006). *Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12: Findings from a 5-year Evaluation*. Palo Alto, CA: American Institutes for Research.
- Payne, T.W., & Lynn, R. (2010). Sex differences in second language comprehension. *Personality and Individual Differences* 50, 434–436 Retrieved from <https://doi.org/10.1016/j.paid.2010.10.026>
- Perie, M., Moran, R., & Lutkus, A.D. (2005). NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics (NCES No. 2005– 464). National Center for Education Statistics Working Paper. Retrieved from <http://eric.ed.gov/?id=ED485627>.
- Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools*, 36: 172-187.
- Ramirez, A., & Carpenter, D. (2005). Challenging assumptions about the achievement gap. *Phi Delta Kappan*, 86(8), 599-603.

- Ramsey, C.A., & Wright, E. N. (1974). Age and second language learning. *The Journal of Social Psychology*, 94:115–121.
- Reardon, S.F., & C. Galindo. (2007). Patterns of Hispanic students' math skill proficiency in the early elementary grades. *Journal of Latinos and Education*, 6(3), 229-252.
- Roberts, G., & Bryant, D. (2011). Early mathematics achievement trajectories: English-language learner and native English-speaker estimates, using the early childhood longitudinal survey. *Developmental Psychology*, 47(4):916–930. Retrieved from <http://dx.doi.org/10.1037/a0023865>.
- Rojas, R. & Iglesias, A. (2013). The Language Growth of Spanish-Speaking English Language Learners. *Child Development*, 84 (2), 630–646.
- Rolstad, K., MacSwan, J., & Guzman, N. (2015). Bilingual learners and the purposes of language assessment. *Miríada Hispánica*, 10, 207-224.
- Riddle, W., & Kober, N. (2012). *What Impact Will NCLB Waivers Have on the Consistency, Complexity, and Transparency of State Accountability Systems?* Washington D.C: Center on Education Policy. Retrieved from <http://files.eric.ed.gov/fulltext/ED535956.pdf>
- Salazar, J.J. (2007). *Master Plan Evaluation Report for English Learners—2005/2006*. Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Assessment Branch. Retrieved from http://notebook.lausd.net/pls/ptl/docs/PAGE/CA_LAUSD/FLDR_ORGANIZATIONS/FLDR_PLCY_RES_DEV/PAR_DIVISION_MAIN/RESEARCH_UNIT/P

U BICATIONS/REPORTS/MASTER%20PLAN%20REPORT%20(05-06)--
FINAL%20REPORT.PDF.

- Saunders, W., & O'Brien, G. (2006). *Oral language. Educating English Language Learners: A Synthesis of Research*. New York: Cambridge University Press.
- Scarcella, R. (2003). Accelerating academic English: A focus on the English learner. Oakland, CA: Regents of the University of California.
- Shneyderman, A., & Froman, T. (2012). *Time to English Reading Proficiency* (Research Brief No. 1201). Research Services, Miami-Dade County Public Schools. Retrieved from <http://eric.ed.gov/?q=ED536602&id=ED536602>.
- Singer, J. D., & Willett, J. B. (2003). Describing discrete-time event occurrence data. In *Applied longitudinal data analysis: modeling change and event occurrence* (pp. 325–356). Oxford ; New York: Oxford University Press.
- Singleton, D. and Ryan, L. (2004). *Language Acquisition: The Age Factor*. Clevedon, UK: Multilingual Matters Ltd.
- Slama, R.B. (2014). Investigating whether and when English learners are reclassified into mainstream classrooms in the United States: A discrete-time survival analysis. *American Educational Research Journal*, 51(2):220–252. Retrieved from <http://aer.sagepub.com/content/51/2/220.short>.
- Snow, S., & Hoefnagel-Höhe, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development* 49:1114–1128.
- Schwarz, Gideon E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6 (2): 461–464. Retrieved from <https://projecteuclid.org/euclid.aos/1176344136>

- Thompson, K. (2012). Are we there yet? Exploring English learners' journey to reclassification and beyond. Stanford University. Retrieved: <http://purl.stanford.edu/qd054zq1735>.
- Thompson, K.D. (2015). Questioning the long-term English learner label: how categorization can blind us to students' abilities. *Teachers College Record*, 117(12). Retrieved from <http://www.tcrecord.org/Content.asp?ContentId=18152>.
- Tong, F., Lara-Alecio, R., Irby, B. J., & Mathes, P. G. (2011). The effects of an instructional intervention on dual language development among first-grade Hispanic English-learning boys and girls: A two-year longitudinal study. *Journal of Educational Research*, 104(2), 87–99.
- Torres, H. N., & Zeidler, D. L. (2001). The effects of English language proficiency and scientific reasoning skills on the acquisition of science content knowledge by Hispanic English language learners and native English language speaking students. *Electronic Journal of Science Education*, 6(3). Retrieved from <http://ejse.southwestern.edu/article/download/7683/5450>
- Uchikoshi, Y. (2006). English vocabulary development in bilingual kindergartners: What are the best predictors? *Bilingualism: Language and Cognition*, 9:33–49.
- Umansky, I.M., & Reardon, S.F. (2014). Reclassification Patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal*, 51(5):879–912. Retrieved from <http://dx.doi.org/10.3102/0002831214545110>.

- U.S. Census Bureau (2015). American Community Survey 1-year estimates. Retrieved from <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2015/1-year.html>.
- U.S. Department of Education (2015). U.S. Departments of Education and Justice Release Joint Guidance to Ensure English Learner Students Have Equal Access to High-Quality Education. Retrieved from <http://www.ed.gov/news/press-releases/us-departments-education-and-justicerelease-joint-guidance-ensure-english-learner-students-have-equal-access-high-quality-education>
- U. S. Department of Education (2017). Ed data express: Data about elementary and secondary schools in the U.S. Retrieved from <http://eddataexpress.ed.gov>
- U.S. Department of Education, Office of Civil Rights. (2014). Civil rights data collection. Retrieved from <http://ocrdata.ed.gov/>
- U.S. Department of Education. (2011). Letters from the Education Secretary or Deputy Secretary. Retrieved from <http://www2.ed.gov/policy/gen/guid/secletter/110923.html>
- U.S. Department of Education. (2015). *Fast Facts: Profiles of English Learners*. Retrieved from: http://www.ncela.us/files/fast_facts/OELA_FastFacts_ProfilesOfELs.pdf
- U.S. Department of Education. (2017). *Fast Facts: Profiles of English Learners*. Retrieved from: http://www.ncela.us/files/fast_facts/05-19-2017/ProfilesOfELs_FastFacts.pdf
- Warren, P. (2004). A Look at the Progress of English Learner Students. Sacramento, CA. Legislative Analyst's Office. Retrieved from www.lao.ca.gov

- World-Class Instructional Design and Assessment. (2012). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 202, 2010-2011 Administration*. Annual Technical Report No. 7. Volume 1 of 3: Description, Validity, and Student Results. Retrieved from <https://www.wida.us>
- World-Class Instructional Design and Assessment. (2013). *ACCESS for ELLs Growth Reports: User Guide*. Retrieved from <https://www.wida.us>
- World-Class Instructional Design and Assessment. (2014). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 302, 2013-2014 Administration*. Annual Technical Report No. 10. Volume 1 of 3: Description, Validity, and Student Results. Retrieved from <https://www.wida.us>
- World-Class Instructional Design and Assessment. (2015). *FAQs: ACCESS for ELLs Test Administration*. Retrieved from https://www.wida.us/assessment/FAQs_ACCESS2%200_Accommodations.pdf.
- World-Class Instructional Design and Assessment. (2017-18). *Test Administrator Manual ACCESS for ELLs 2.0 Kindergarten ACCESS for ELLs Alternate ACCESS for ELL*. Retrieved from <https://www.wida.us/ACCESSupdates>
- Wiley, T. (1996). *Literacy and language diversity in the United States*. McHenry, IL: Center for Applied Linguistics and Delta Systems, Co.
- Yanosky, T., Yen, S.J., Louguit, M., MacGregor, D., Zhang, Y., Kenyon, D.M. (2011). *Annual Technical Report for ACCESS for ELLs® English Language Proficiency Test, Series 201, 2009–2010 Administration*. Board of Regents of the University of Wisconsin System on behalf of the WIDA Consortium.