# UC Irvine
## UC Irvine Electronic Theses and Dissertations

### Title
Language Classification and its Consequences for English Language Learners in Diverse Middle Schools

### Permalink
https://escholarship.org/uc/item/1fv7r5ft

### Author
Reyes, Marcela Raquel

### Publication Date
2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Language Classification and its Consequences for English Language Learners in Diverse Middle Schools

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Education

By

Marcela Raquel Reyes

Dissertation Committee:
Distinguish Professor Greg Duncan, Chair
Associate Professor Thurston Domina, Co-chair
Assistant Professor Maria Rosales-Rueda
Professor Jamal Abedi

2016

© 2016 Marcela Raquel Reyes

## DEDICATION

To

my loving husband, Konrad,

for his support and encouragement

"It always seems impossible until it is done"
Nelson Mandela

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## ACKNOWLEDGEMENTS

I would like to thank and acknowledge my dissertation committee. Greg Duncan, Chair, I thank for his valuable constructive criticism and feedback which made my dissertation work more rigorous. Thad Domina, Co-Chair, I thank for his constant support and mentorship throughout the years and reassurance with the highest levels of expectation throughout my graduate experience. I greatly appreciate his confidence in my abilities. I also thank him for providing access to the districts data through the Evaluating the Quality of Universal Algebra Learning (EQUAL) project. Maria Rosales-Rueda, I thank for her constant one-on-one feedback on the data analysis and conceptualizing of the results as I worked on my dissertation. Jamal Abedi, I thank for his expertise and assistance on situating my work in the grand literature on English language learners. Without their guidance and persistent help this dissertation would not have been possible.

I would also like to thank my husband, Konrad, for his support and encouragement throughout my PhD program and the dissertation process. He encouraged me to work on my passion to improve adolescent's academic opportunities in secondary school. He created a conducive learning environment at home and he also cared for our son, Xavier, as I worked on my dissertation.

I also thank UCI Irvine Graduate Division for offering me the Eugene Cota-Robles (ECR) Fellowship. ECR funding allowed me to explore my research interests, publish articles, and ultimately successfully complete my dissertation.

Finally, I thank the two Southern California School Districts and all the administrators (that must remain anonymous) that provided me with the district student-level data and provided valuable information regarding their districts. Thank you for participating in my research study and helping me understand English language learners classification process and course placement in middle school. I also want to acknowledge the Spencer foundation that provided funding to collect district data through the EQUAL project.

## CURRICULUM VITAE

Marcela Raquel Reyes

## RESEARCH INTEREST

Secondary education, adolescents' experience, English language learners, course enrollment, students' motivation, teacher expectations, college aspirations, career aspirations, and college access

## EDUCATION

Ph.D. Education, specialization in Educational Policy and Social Context, 2016
University of California, Irvine
*Dissertation Title:* English Language Learners Opportunities to Learn: Language Proficiency Classification and its Consequences in Diverse Middle Schools
*Dissertation Committee:* Greg Duncan (Chair) Thurston Domina (Co-Chair), Maria Fernanda Rosales-Rueda, and Jamal Abedi

M.A. Education, specialization in Educational Policy and Social Context, 2014
University of California, Irvine
Qualifying Paper: *Motivational Predictors of Math Course Persistence*
*Advisor:* Thurston Domina

M.A. Education, specialization in Social Research Methodology, 2009
University of California, Los Angeles
*Advisor:* Frederick Erickson

B.A. Sociology & minor in Mathematics, 2005
Mount St. Mary's College
*Advisor:* Pam Haldeman

## PUBLICATIONS

**Reyes, M**. and Domina, T. (forthcoming). "Track placement and the motivational predictors of math course enrollment." *Teachers College Record.*

Hwang, N., **Reyes, M.**, and Eccles, J. (2016) "Implicit theories of intelligence by gender, race, and achievement: Who holds a fixed mindset and who does it harm in high school mathematics?" *Youth and Society.*

Simzar, R., **Martinez, M.,** Rutherford, T., Domina, T., and Conley, A.M. (2015). "Raising the stakes: How students' motivation for mathematics associates with high- and low-stakes test achievement." *Learning and Individual Differences.*

## WORK UNDER REVIEW

**Reyes, M**., Powell, M. and McDaniel, A. (under review) "Tracking high school students' pathways to male- and female-dominated STEM careers: The influence of career expectations, math course-taking and self-efficacy"

## WORK IN PROGRESS

**Reyes, M.** "A mixed-method study: Districts' implementation of language classification policies and the implications for male, Hispanic, and low-income middle school students"
**Reyes, M. "**A mixed-method study: Language classification, and English and math course placement in middle schools"
**Reyes, M.** "Middle school language classification effects on high school students' achievement and behavioral outcomes"

## REPORT

Herman, J., Dai, Y., Htut, A., **Martinez, M.**, Rivera, N. (2010). *Evaluation of the Enhanced Assessment Grant (EAG).* Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from: http://www.cse.ucla.edu/products/reports/R791.pdf

## RECENT AWARDS

Keith Curry Award 2015-2016
Ford Foundation Honorable Mention Fellowship 2015-2016
Faculty Mentor Program Honorable Mention Fellowship 2014-2015
Eugene Cota-Robles Fellowship 2011-2015
Leadership Award (Scholar Mentor Club, Tutor) 2005
Outstanding Scholar in Theory and Research in Sociology 2005

## CONFERENCES/PRESENTATIONS

**Reyes, M.,** Powell, M. and McDaniel, A. (2016, April). Tracking male and female high school pathways to male- and female-dominated STEM careers: The influence of career expectations, math course-taking and self-efficacy. Paper presented at the annual meeting of the American Educational Research Association.

Hwang, N., **Reyes, M.**, and Eccles, J. (2016, April). Implicit theories of intelligence by gender, race, and achievement: Who holds a fixed mindset and who does it harm in high school mathematics? Paper presented at the annual meeting of the American Educational Research Association.

Simzar, R.M., **Martinez, M.,** Sanabria, T., Rutherford, T., Domina, T., & Conley, A.M. (2014, August). *Student motivation for mathematics and high-stakes versus low-stakes test achievement.* Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.

**Martinez, M**. (2014, April). *Motivational predictors of math course persistence*. Paper presented at the annual symposium of UC Irvine Associated Graduate Students, Irvine, CA.

**Martinez, M.** (2014, March). *Motivational predictors of math course persistence*. Paper presented at the biennial meeting of the Society for Research on Educational Effectiveness, Washington, D.C.

**Martinez, M.** (2012, October). *Math track: Parent's expectations and theories of intelligence*. Poster presented at the annual UC Irvine Interdisciplinary Conference on Researching Equity, Irvine, CA

**RESEARCH EXPERIENCE**

**Graduate Student Researcher** 2014-Present
Program Director: Thurston Domina
Location: UC Irvine, School of Education
Project: Evaluating the Quality of Universal Algebra Learning (EQUAL) Project

**Graduate Student Researcher** 2012-2015
Program Director: Greg Duncan
Location: UC Irvine, School of Education
Project: Child Development Meta-Analysis 1960-Present

**Public Administration Analyst** 2009-2011
Program Director: Noelle Griffin, Eva Baker, Joan Herman
Location: UCLA, National Center for Research on Evaluation, Standards, & Student Testing
Projects: Green Dot, Evaluation of the Enhanced Assessment Grants, CDE after school, etc.

**Field Interviewer** 2007-2008 & 2010-2011
Program Director: Jennifer Hawes-Dawson
Location: RAND
Project: Healthy Passages

**Graduate Student Researcher (Project Coordinator)** 2008-2009
Program Director: Peggy Toy
Location: UCLA Center for Health Policy Research
Project: Health Data (*Datos y Democracia*)

**Professional Expert** 2006-2008
Program Director: Katherine Hayes, Deborah Oliver, Alicia Fernandez, Ebrahim Maddahian
Location: Los Angeles Unified School District
Projects: Spanish Language Arts, Kindergarten project, Modified Consent Decree, etc.

**Consultant/Research Intern** 2004-2006
Program Director: Don Haviland
Location: Mount St. Mary's College, Institutional Research and Assessment
Project: MSMC Internal Reviews

## TEACHING EXPERIENCE

**Teacher Assistant** EDU 50: Origins, Purposes & Central Issues in K-12 Education 2012
**Project Coordinator** UCLA Center for Health Policy Research 2008-2009
**Math Tutor** Mount St. Mary's College (The Scholar Mentor Club) 2003-2005
**K-12 Lead Tutor** Century Lift 2003-2006

## INVITED/GUEST LECTURES

Spring2015: "High School Curriculum/Higher Education" EDU 50: Origins, Purposes and Central Issues in K-12 Education
Fall2014: "Math Policies and Inequalities" EDU 50: Origins, Purposes and Central Issues in K-12 Education
Summer2014: "Math Policies and Inequalities" EDU 50: Origins, Purposes and Central Issues in K-12 Education
Winter2014: "Math Policies and Inequalities" EDU 50: Origins, Purposes and Central Issues in K-12 Education
Winter2013: "Achievement Motivation and Social Context" EDU108: Adolescent Development and Education

## LANGUAGE

Spanish (Full Professional Proficiency)

## ACADEMIC SERVICE

Fall 2012-Spring 2015: **Chicano/Latino Graduate Student Collective** at UC Irvine (Member and Web Designer) Website: http://clubs.uci.edu/clgsc/?page_id=182
Fall2013/Winter 2014: **PhD Admissions Committee in Educational Policy and Social Context** at UC Irvine, SOE
Fall 2013: **California Forum for Diversity in Graduate Education** at Saint Mary's Forum (SOE Representative)
Summer 2013: **Competitive Edge Summer Research Program** at UC Irvine (Mentor)
Spring 2013: **California Forum for Diversity in Graduate Education** at UCI (SOE Representative)

## PEER REVIEWER

*American Educational Research Journal, 2014*
*Asian Social Science Journal, 2015*
*Education Researcher, 2013-2014*
*Gender and Society, 2015-2016*
*Journal of Engineering Education, 2016*
*Sociology of Education, 2014-2015*

**PROFESSIONAL ASSOCIATIONS**

American Educational Research Association (AERA)
American Psychological Association (APA)
Society for Research on Educational Effectiveness (SREE)

**SPECIALIZED TRAININGS**

AERA Institute of Statistical Analysis: Math Education and Equity on May 20-23, 2013 (Washington, DC)
University of California Evaluation Center (UCEC) program on July 24-26, 2012 (Santa Barbara, CA)

**STATISTICAL PROGRAMS**

STATA, SPSS, CMA, Atlas.ti, and Dedoose

**STATISTICAL SKILLS**

Applied Regressions, Structural Equation Models (SEM), Hierarchical Linear Model (HLM), Regression Discontinuity (RD), RD with Multiple Rating-Score Variables

## ABSTRACT OF THE DISSERTATION


Language Classification and its Consequences for English Language Learners
in Diverse Middle Schools
By
Marcela Raquel Reyes
Doctor of Philosophy in Education
University of California, Irvine, 2016
Professor Greg Duncan, Chair
Professor Thurston Domina, Co-Chair



**Executive Summary**

California policy defines English Language Learners (ELLs) as students who have not reached English proficiency and who speak another language at home. The state sets minimum requirements to determine students' English proficiency, but individual districts have the freedom to be more rigorous. Therefore, a student can be considered ELL in one district and Reclassified Fluent English Proficient (RFEP) in another. My work examines how districts establish their language classification policies and the extent to which the classification itself determines students' opportunities to learn.

My three-study dissertation addresses the following research questions: 1) How do school districts establish and implement their language classification policies for middle school students? 2) How does language classification affect eighth-grade English and math course placement? 3) How does language classification affect high school students' English and math achievement and behavioral outcomes? I use data from two Southern California districts, including interviews, documentation and student-level data, to answer these questions. I focus on middle school because it is less frequently studied and it determines adolescents' high school outcomes.

In study one, administrators from Manzanita Unified School District (MUSD) and Granada Unified School District (GUSD) developed individual language classification to equitably determine their students' English proficiency. However, in both districts, assessments served as gatekeepers that students must pass to be reclassified. Male, Hispanic, and low-income students were less likely to be reclassified, even if we consider only those who passed the district's or state's requirements. In study two, MUSD administrators stated language classification itself did not determine eighth-grade English and math placement. However, data show under what circumstances language classification influences students' course placement. The year ELL started to receive specialized English courses, they were placed in lower math courses. In study three, in most instances middle school language classification itself does not affect achievement or behavioral outcomes in high school. At MUSD, ELL may be receiving appropriate English and math; therefore, there is no effect. Alternately, reclassification may not have impact since neither ELL nor RFEP may be receiving adequately rigorous content. The few times classification does affect outcomes, RFEP perform worse than ELL students.

# INTRODUCTION

In California, 25% of K-12 students are English Language Learners (ELLs), compared with 11% of K-12 students nationally (Hill, Betts, Chavez, Zau, & Bachofer, 2014; Kohler & Lazarín, 2007). ELLs are designated students who have not yet reached full English proficiency and who speak another language at home. They are one of the fastest growing student groups in the country, yet are also among the lowest performers on a broad range of educational outcomes (Capps et al., 2005; Maxwell, 2014). ELL students, on average, score lower than non-ELL students in English reading, writing, and comprehension, as well as in less language-intensive subject areas, such as mathematics and science (Edwards, Leichty, & Wilson, 2008; Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hampden-Thompson, Mulligan, Kinukawa, & Halle, 2008). ELLs are also more likely to be more disengaged from school and to have more school absences, school suspensions, and higher dropout rates (Losen & Martinez, 2013; Uriarte, Lavan, Agusti, & Karp, 2009). Yet, the ELL classification itself may have unintended consequences if classified students are not given sufficient opportunities to learn rigorous educational content or if they are not suitably integrated with their non-ELL peers. Thus, it remains unclear whether this group's language skills, the stigma associated with the ELL label, or a lack of access to rigorous courses drives the achievement gap between ELL and non-ELL students.

In this three-study dissertation, I use data from two Southern California districts to examine the processes educators use to classify the English language proficiency of middle school language minorities. For one district, I also examine how language classification itself affects student course placement, achievement, and behavioral outcomes in secondary school. Study one reveals a disconnect between administrative language classification policies and their implementation. Moreover, although district administrators approached their language classification policies differently, student-level data indicate that both districts rely heavily on the assessments students must pass to be reclassified. Further, this work shows that male, Hispanic, and low-income students are less likely to be reclassified, even if we consider only those who pass the district's or state's minimum requirements. Study two focuses on one district; here administrators stated prior student performance determined English and math course placement rather than the language classification itself. However, regression discontinuity (RD) models show language classification itself can have unintended consequences under specific course placement policies. RD models find that in the first two cohorts, classification did not affect English and math placement. On the other hand, in the third cohort the district began to provide English courses specifically for ELL students, but for different reasons they also started to explicitly distinguish between algebra and pre-algebra. Unintentionally, ELL students are more likely to placed in the latter, slower math track. The third study, which also focuses on one district, finds regressions show RFEP students had higher California Standards Test (CST) in English Language Arts, California High School Exit Examination (CAHSEE) in English Language Arts, CAHSEE math, and advanced math courses, as well as lower attendance and suspension rates than ELL students. RD models, however, provided less biased estimates, revealing that these differences are due to unmeasured factors, and not language classification itself. In most instances, language classification itself did not affect student achievement and behavior. Only in a few instances do differences exist, and RFEP students are less likely to pass the CHASEE ELA portion and more likely to be suspended compared with ELL students.

I specifically focus on California because of its large representation of ELL students and because the state's current demographics are projected to become the demographic makeup of the nation as a whole in the coming decades (Kucsera, Siegel-Hawley, & Orfield, 2014). The California Department of Education (CDE) requires children who speak another language at home to take and pass the California English Language Development Test (CELDT) to assess their English proficiency, as well as the CST in English Language Arts, which assesses their English skills.[1] ELL students are those students who fail at least one of these two tests. Each school year, ELLs are eligible to retake these tests to become what is known as Reclassified Fluent English Proficient (RFEP), sufficiently proficient in English. There are also those students classified as Initially Fluent English Proficient (IFEP), meaning students who do not speak English at home but who are able to pass the language classification assessments on their first attempt. Finally, the label of English Only (EO) is given to monolingual English students who are not required to take these tests for language classification purposes.

Often researchers use socially constructed categories to compare students' academic achievements without questioning the categories themselves. We need to investigate how educators construct their language classification categories, particularly when school districts have the freedom to establish their own language classification policies. California sets minimum requirements on how to determine students' language classification, but 90% of California's districts set even higher reclassification requirements than those of the state (Hill, Weston, & Hayes, 2014; Linquanti & Cook, 2013). These higher requirements make it more difficult for students in these districts to reclassify. This can also lead to disparities, as a particular student can be considered ELL in one district and RFEP in another. Using an organizational model framework I examine how administrators establish the district's language classification policies

[1] In these studies, I examine the years prior to California's recent implementation of the Common Core State Standards (CCSS), the Smarter Balanced Assessment Consortium (SBAC), and the increase in per-pupil funding for ELL students, each of which will almost certainly affect future ELL policies and practices (Hill, Weston, et al., 2014; Umansky et al., 2015). In the discussion section of each study, I address the implications of my findings as they pertain to the new policy changes.

and how its policies are implemented. These policies are important since studies show, male, Hispanic, and low-income students are more likely to be categorized ELL than female, Asian American, and high-income students (Grissom, 2004; Halle, Hair, Wandner, McNamara, & Chien, 2012; Kohler & Lazarín, 2007; Slama, 2014). Yet, these studies mainly focus on elementary school grades or they do not account for students' prior achievement.

Language classification policies can be critical, as elementary and high school ELL students often tend to be placed in classes that are separate from their non-ELL peers, and these classes tend to be less rigorous (Callahan, 2005; Callahan, Wilkinson, & Muller, 2010; Gandara et al., 2003; Hahnel, Wolf, Banks, & LaFors, 2014; Mayer, 2008). We currently know less about middle school students' circumstances, however, here, the potential inequalities in students' opportunities to learn may compound certain disadvantages facing ELL students. ELL students, on average, score lower than RFEP students in English and math on every measure examined (Edwards et al., 2008; Grimssom, 2004; Gandara & Rumberger, 2009; Mosqueda & Maldonado, 2013; Slama, 2014; Saunders & Marcelletti, 2013), and that ELL students are more likely be suspended, absent, and dropout of school altogether (Uriarte et al., 2009). However, most of the literature on language reclassification, course placement, and behavioral outcomes has selection bias limitations: students are not randomly reclassified RFEP, and they are not arbitrarily placed in advanced English and math courses. It is unclear if these differences are explained by the different opportunities to learn or if they can be explicated by the way language classification categories are created. It may be that students who are being reclassified are those who are more academically inclined, and therefore, their higher achievement scores may be spurious, due to some other unmeasured variables (e.g., motivation, cognitive abilities, parental expectations, etc.) that capture their academic abilities.

In my own work, I specifically compare current ELL middle school students with peers who reclassified in middle school. Middle school ELL students are less frequently studied than ELL elementary students, yet are more likely to be Long-Term English Language Learners (LTELL) and foreign-born children (Hahnel et al., 2014; Olsen, 2010). Indeed, middle school is a significant schooling stage that often determines the educational foundation of an adolescent's high school, college, and, ultimately, career outcomes (Walqui et al., 2010). Moreover, researchers who focus on high school ELL students compare them with all RFEP students, regardless of when reclassification occurred. However, researchers have found that students who reclassify in elementary school have stronger academic outcomes compared with students who reclassify in secondary school (Halle et al., 2012). For these reasons, I focus on middle school language minorities.

Of course, ELL middle school students are a diverse and complex group, which stems from differences in English proficiency, national origin, socioeconomic status, previous schooling, and the number of years in the U.S. system (Callahan, 2005; Krashen & Brown, 2005). For this reason, there are many aspects to consider when discussing middle school ELL students' academic needs and how educators might improve academic and behavioral outcomes. I conduct two mixed-methods studies and one quantitative study to investigate the specific situation of middle school learners. The following introductory sections provide a short summary of the three studies, and include a brief overview of the motivation, theoretical framework, and key results of my work.

**Language Classification Policies**

In the first study—a mixed-method study—I used the concurrent triangulation strategy (Creswell, 2013) where I collected both quantitative and qualitative data concurrently, and then compared the results to determine if there was convergence, differences, or some combination. I used data from two Southern California districts to examine the processes educators use to classify middle school language minorities. I accessed district data through the Spencer-funded Evaluating the Quality of Universal Algebra Learning (EQUAL) project. I chose to focus my investigation on districts I refer to pseudonymously as Manzanita Unified School District (MUSD) and the Granada Unified School District (GUSD) because they had large percentages of ELL students, and, more specifically, large numbers of Hispanics and Asians. Furthermore, these two districts provided insight into practices that prevail in relatively low-income communities. Here, I addressed the following research questions:

1) How do school districts establish language classification policies? (Qualitative)
2) How do school districts implement the language classification policies? (Quantitative)
    a. What roles do student gender, race, and socioeconomic status play in the language classification process?
3) How would school districts' reclassification rates change if the state's minimum language classification policies were implemented? (Quantitative)

As previously mentioned, most of California's school districts set higher reclassification requirements than those mandated by the state (Hill, Weston, et al., 2014). This is troubling when male, Hispanic, and low-income students are more likely to be classified ELL (Grissom, 2004; Halle, Hair, Wandner, McNamara, & Chien, 2012; Kohler & Lazarín, 2007), even when accounting for their prior English performance (Slama, 2014). Both of the districts I investigated set higher minimum requirements than the state's, but their requirements and rationales differed. According to interviews with administrators, MUSD and GUSD intended to implement nondiscriminatory policies that would correctly categorize students in their proper ELL and RFEP categories. At the same time, both districts believed it was essential that ELL students not be reclassified too early because, once reclassified, the student would lose language support.

MUSD requires ELL students to score a 556 or higher on the CELDT, and a 325 or higher on the CST; however, they also considered English teacher recommendations, and allowed parents to request their child's classification be changed, regardless of the test scores. On the other hand, GUSD claimed to make their language classification objectively, based on a student's CELDT, CST, and essay exam scores. In GUSD, district administrators expected English teachers to reclassify all students who passed the CELDT, CST, and essay, believing teachers seldom use their discretion when the student has not passed all three requirements. As discussed above, regardless of administrative classification philosophies, male, Hispanic, and low-income students were less likely to reclassify in both districts.

During the course of my research, it became clear that district administrators have good intentions and want to reclassify students properly. For example, MUSD is willing to make exceptions and consider special requests, while, for its part, GUSD attempts to make decisions on assessments that will objectively tell them if students are English proficient. With an organizational model framework I examine the language classification process that describes districts as "fragmented centralized" organizations that require different actors to construct and implement policies (Meyer, 1983, p. 181). In both districts, the imperfect language classification process prevents many ELL students from reclassification. Manzanita's administrators may have encouraged reclassification be based on teacher and parent recommendations instead of assessments, but, in most instances, ELL students had to meet the district's higher than state standards to be considered for reclassification. Granada's administrators may have encouraged English teachers to base reclassification on assessments, but teachers made exceptions for certain students who failed the CELDT or CST, but passed the essay exam (that they themselves scored). Moreover, there are demographic differences regarding those who pass the district's essay exam and who takes the essay exam, which, in the end, is another unfortunate mechanism preventing a number of male, Hispanic, and low-income students from reclassifying. The educational language classification process in both districts is loosely coupled, with administrators setting language classification rules that English teachers are not necessarily following.

In Manzanita, one-quarter of the sixth graders, and, in Granada, two-fifths of the sixth graders were classified ELL, but only 28% and 26%, respectively, were reclassified by the eighth grade. Complicated and changing reclassification criteria can make the pathway out of the ELL category extremely difficult, thus placing an undue burden on students, particularly given the methods of entry into the system. Significantly, more students would be reclassified in both districts if the districts based reclassification solely on the CELDT and CST; the same is true if the districts followed the state's minimum language classification requirements. Most prior research demonstrates that the ELL classification can have negative implications on student course placement, achievement, and behavioral outcomes (e.g., Edwards et al., 2008; Gandara & Rumberger, 2009; Grissom, 2004; Mosqueda & Maldonado, 2013; Uriarte et al., 2009). However, there are two studies that show, in a few instances, RFEP students perform worse than ELL students in high school (Robinson-Cimpian & Thompson, 2015; Robinson, 2011). We should also consider if the RFEP classification has unintended consequences.

**Language Classification and Course Placement**

In study two, which was a mixed-methods study, I used the concurrent embedded strategy (Creswell, 2013) in which I used qualitative data as a supporting role to the quantitative data. I focused on the Manzanita Unified School District (MUSD) to estimate the effects of seventh grade language classification on eighth-grade English and math course placement, as well as their eighth-grade classroom peers' English and math achievement distributions. As most literature on language reclassification and course placement is limited by selection bias (e.g., Callahan et al., 2010; Gandara et al., 2003), I used the regression discontinuity (RD) method that accounts for students not being randomly reclassified RFEP and not arbitrarily placed in advanced English and math courses. On average, MUSD ELL middle school students had a 40% probability of being reclassified if they passed the CELDT at 556 and CST at 325. ELL students below the cutoff had about 10% probability of being reclassified. This fuzzy RD allows us to compare students near the cutoff threshold. I did not use the Granada Unified School District (GUSD) data in this investigation because it did not meet one of the RD assumptions: RD requires that the assignment variables include continuous variables, and GUSD's essay exam was scored on a scale of 1-4. Furthermore, a frontier RD could not be carried out because the essay exam was more difficult to pass than either the CELDT or CST; I could not conduct an RD with only those students who passed the essay exam. Therefore, using the MUSD data, I addressed the following research questions:

1) How is English and math course placement determined for ELL and RFEP middle school students? (Qualitative)
2) How does language classification (ELL and RFEP) affect middle school English and math course placement? (Quantitative)
    a. What is the peer achievement composition in ELL and RFEP students' English and math classrooms? (Quantitative)

ELL students should receive academic support to become proficient in English, but all too often are put into largely ineffective remedial and support English courses (Menken & Kleyn, 2010; Olsen, 2010). Furthermore, ELL students can be placed into an overarching track, such that students who are exposed to low-level instruction in one subject tend to be subjected to low-level instruction in all areas. For example, high school ELL students are tracked in this way and are less likely to be enrolled in college preparatory coursework in math, science, or the social sciences (Callahan, 2005; Callahan et al., 2010; Gandara et al., 2003; Hahnel et al., 2014). ELL students can also be tracked in less obvious ways, as when they are placed in English and math courses that have similar course titles as non-ELL classes, but that are taught separately. For example, Cogan, Schmidt, and Wiley (2001) find nearly 30% of U.S. eighth-grade mathematic course titles do not match the textbooks they employ. For this reason, the classroom composition must also be examined. This type of ELL student tracking is mainly researched at the elementary school level. Here, students tend to be "segregated by classroom," where 25% of the first-to-fourth grade California teachers have 50% of the English language learners (Gandara et al., 2003). Most prior research demonstrates ELL elementary and high school students do not have sufficient opportunities to learn. However, Robinson (2011) results show that language classification itself did not affect English, math, science, or other college preparatory course placement in high school. This is the only study that used an RD method where we can make causal inferences.

Middle school course placement research, as well as work that uses rigorous statistical methods, is lacking, which I hope to begin to remedy with the present work. Based on interviews with district administrators and documentation, English and math classes can be heterogeneous by language classification and theoretically these students can be placed in the same classrooms because teachers and counselors make final course placement decisions. I run RD models to account for the strong relationship between language classification and English assessment scores. When the three eighth grade cohorts were combined, RFEP students had about a 24 percentage-point ($p < .01$) jump in likelihood of being placed in one mainstream English course rather than two when compared with ELL students who had nearly equivalent CELDT and CST scores and thus, arguably, equivalent English proficiency levels. RFEP students also had about a 19 percentage-point ($p < .01$) jump in likelihood of being placed in an accelerated math course than a basic math course when compared with ELL students. However, the effect sizes became non-statistically significant when the RD bandwidths decreased. My investigation shows that language classification itself did not always determine student course placement for each cohort. For the first two cohorts, I found that a student's English and math course placement was not determined by student classification itself. Only in the third cohort did students classified RFEP have a 51 percentage-point ($p < .001$) increase of being placed into one English course. Likewise, Cohort 3's RFEP students had a 24 percentage-point ($p < .05$) increase of being placed into an accelerated math course instead of a basic math course. These RD effect sizes provide us with estimates that are less biased than ordinary least square (OLS) regressions that are predominantly used in the classification literature. Based on OLS regressions, we would have concluded the ELL classification negatively affects English and math course placement in all cohorts—a downward bias compared with the RD.

Based on the CDE, I expected ELL students to receive different English courses than RFEP students because the overriding purpose of classifying students is to provide them suitable and tailored courses to become English proficient. However, based on interviews with administrators I anticipated that ELL and RFEP students would have the same opportunities to enroll in mainstream English and accelerated math courses. For the first two cohorts, I found that language classification did not affect course placement. However, for the third cohort, the district had started providing English courses specifically for their ELL students, while, simultaneously, explicitly distinguishing between algebra and pre-algebra. This created a situation where, seemingly unintentionally, ELL students were more likely to be placed in the lower math track.

Thus, in the third cohort, the language classification itself could have implications on whether a student was able to access specialized English courses, but also advanced math courses. This, in turn, could affect ELL and RFEP achievement and behavioral outcomes.

**Language Classification and Achievement and Behavioral Outcomes**

For study three, a quantitative study, I focused on MUSD, and estimated the effects of language classification by the end of the eighth grade on high school achievement (i.e., their English and math scores, and the highest level of courses they completed), as well as on their behavioral outcomes (i.e., attendance and suspensions). Here, I addressed the following two research questions:

1) How does language classification (ELL and RFEP) by the end of middle school affect the high school students' English and math achievement outcomes (i.e., assessments and course placement)? Further, how does middle school course placement moderate the association between language classification and achievement?

2) How does language classification (ELL and RFEP) by the end of middle school affect high school students' behavioral outcomes (i.e., attendance and suspensions)?

Based on different assessments, researchers find that ELL students have lower math and reading scores than RFEP students, even after including several controls (Edwards et al., 2008; Gandara & Rumberger, 2009; Grissom, 2004; Mosqueda & Maldonado, 2013; Saunders & Marcelletti, 2012; Slama, 2014). However, achievement outcomes might vary based on who is included in the RFEP category. ELL students have the opportunity to become RFEP each school year; for this reason, RFEPs are former ELL students, and the composition of the student body changes each year. In fact, students who reclassify early in elementary grades tend to have stronger academic outcomes than students who reclassify in later elementary grades (Halle et al., 2012; Hill, Weston, et al., 2014). Moreover, generally speaking, ELL students who are not reclassified score lower on different behavioral outcomes (Halle et al., 2012; Losen & Martinez, 2013; Uriarte et al., 2009). However, these previous studies used non-randomized data and did not use rigorous methods to determine whether it is reclassification itself or other confounders that explain early classification and future achievement. Furthermore, from their work, it remains unknown whether student achievement and behavioral outcomes would differ if one were to compare middle school ELL students with students who reclassify in middle school. Being classified ELL in middle school may have unintended consequences if classified students are not given sufficient opportunities to learn rigorous educational content.

To my knowledge, there are only two studies that estimate the effects of language classification itself, both of which use observational data and RD models to make causal inferences. Robinson (2011) finds that, in high school, RFEP Latino and Asian American students score lower than ELL high school students on the CST ELA exam, but that there is no difference for elementary or middle school students. He rationalizes that RFEP students' lower English scores may be because they are removed from English language development courses and moved into mainstream English courses with no support. The same study shows language classification does not affect high school course placement in English, math, science, and other college preparatory courses, or attendance. In another work, Robinson-Cimpian and Thompson's (2015) RD model, based on Los Angeles Unified School District data, shows that making it harder to reclassify increases high school Latino students' CST ELA scores (0.18SD), but that there is no effect on middle school CST ELA scores. Additionally, their work demonstrates that graduation rates increase by 11% when students reclassified. However, they restrict their sample to students who have passed all other CELDT requirements (a frontier RD). Furthermore, they do not distinguish between students who reclassified in elementary versus those who reclassified in middle school.

My OLS estimates agree with the previous studies, as they demonstrate that students who RFEP in middle school have higher CST ELA (0.30σ, $p < .001$), CAHSEE ELA (0.25σ, $p < .001$), and math scores (0.39σ, $p < .0001$). Additionally RFEP students are more likely to be placed in more advanced math courses (0.14σ, $p < .001$; 0.12σ, $p < .01$) in the ninth and eleventh grades than ELL students.[2] Furthermore, RFEP students are less likely to be absent (-0.16σ, $p < .01$) or have as many on-campus suspensions (-0.10, $p < .05$) as ELL students. However, the RD models, which are “as good as random assignment” (Lee & Lemieux, 2010), show that most of these differences are not due to the language classification itself. Rather, I find in most instances the achievement and behavioral differences among the language minorities stem from unmeasured factors that are strongly associated with ELLs and RFEP students. One exception is RFEP students had about a 30 percentage-point ($p < .05$) decrease in likelihood of passing the CAHSEE ELA portion. RFEP students are also more likely to be suspended on-campus (0.59σ, $p <.05$) in tenth grade and off-campus (0.13σ, $p <.05$) in eleventh than ELL students. However, these effect sizes are not consistently significant, but demonstrate that in some circumstances RFEP classification can have negative effects on student outcomes.

**Conclusion**

This dissertation deepens our understanding of language classification practices in middle schools and the effect it has on language minorities’ opportunities to learn. In middle school, male, Hispanic, and low-income students are less likely to be reclassified regardless of the district’s language classification policies. The work also provides evidence regarding the circumstances in which language classification influences eighth-grade English and math course

[2] The CAHSEE has both an English and math component. Each section score ranges from 275 to 450, and students must achieve a minimum score of 350 on each section to pass and graduate with a high school diploma. All tenth graders are required to take the CAHSEE, and only those students who fail the exam are required to take it in later years.

placement. However, as we shall see, one of the most important conclusions to be drawn is that, in most cases, middle school language classification itself does not affect achievement or behavioral outcomes in high school. The few instances in which language classification does affect achievement and behavior, RFEP students perform worse than ELL students. Reclassification itself can have negative consequences.

We can use this information to improve how language classification and course placement policies are created, particularly for middle school language minorities. The ELL classification may not be detrimental, as most prior researchers have concluded. In some districts, language classification and course placement policies may be appropriately working, but these findings may only become evident using rigorous statistical methods. Any differences amongst ELL and RFEP students are due to some other unmeasured variable (e.g., motivation, and academic ability) and not language classification itself. For the most part, MUSD is an example where ELL and RFEP middle school students receive appropriate English and math curricula, and therefore there are minimal achievement and behavioral differences. However, it also shows, in a few instances, that RFEP classification can harm students' educational outcomes. Educators must be concerned with developing both ELL and RFEP language and academic skills with careful monitoring and appropriate academic support so students can participate and succeed in rigorous high school, and ultimately college, courses.

References

Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in US schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoro, S. (2005). The new demography of America's schools: Immigration and the No Child Left Behind Act. *Urban Institute (NJ1)*.

Cogan, L. S., Schmidt, W. H., & Wiley, D. E. (2001). Who takes what math and in which track? Using TIMSS to characterize US students' eighth-grade mathematics learning opportunities. *Educational Evaluation and Policy Analysis, 23*(4), 323-341.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Edwards, B., Leichty, J., & Wilson, K. (2008). English Learners in California: What the Numbers Say. *EdSource*.

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources, Unequal outcomes. *Education Policy Analysis Archives, 11*(36), 1-54.

Gandara, P., & Rumberger, R. W. (2009). Immigration, language, and education: How does language policy structure opportunity. *Teachers College Record, 111*(3), 750-782.

Grissom, J. B. (2004). Reclassification of English Learners. E*ducation Policy Analysis Archives, 12*(36), n36.

Hahnel, C., Wolf, L., Banks, A., & LaFors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. The Education Trust-West.

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.

Hampden-Thompson, G., Mulligan, G., Kinukawa, A., & Halle, T. (2008). Mathematics Achievement of Language-Minority Students During the Elementary Years.

Hill, L. E., Betts, J. R., Chavez, B., Zau, A. C., & Bachofer, K. V. (2014). Pathways to Fluency: Examining the Link between Language Reclassication Policies and Student Success. *Public Policy Institute of California*.

Hill, L. E., Weston, M., & Hayes, J. M. (2014). Reclassification of English Learner Students in California. *Public Policy Institute of California. Retrieved from www. ppic. org/main/publication. asp*

Kohler, A. D., & Lazarín, M. (2007). Hispanic education in the United States. *Statistical Brief, 8*.

Krashen, S., & Brown, C. L. (2005). The ameliorating effects of high socioeconomic status: A secondary analysis. *Bilingual Research Journal, 29*(1), 185-196.

Kucsera, J. V., Siegel-Hawley, G., & Orfield, G. (2014). Are We Segregated and Satisfied? Segregation and Inequality in Southern California Schools. *Urban Education,* 0042085914522499.

Lee, D. S., & Lemieux, T. (2010). *Regression discontinuity designs in economics*. Retrieved from http://www.jstor.org/stable/20778728

Linquanti, 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. *Council of Chief State School Officers*.

Losen, D. J., & Martinez, T. E. (2013). Out of school and off track: The overuse of suspensions in American middle and high schools. *K-12 Racial Disparities in School Discipline*.

Maxwell, L. (2014). US school enrollment hits majority-minority milestone. *Education Week*.

Mayer, A. (2008). Understanding How US Secondary Schools Sort Students for Instructional Purposes: Are All Students Being Served Equally? *American Secondary Education*, 7-25.

Menken, K., & Kleyn, T. (2010). The long-term impact of subtractive schooling in the educational experiences of secondary English language learners. *International Journal of Bilingual Education and Bilingualism, 13*(4), 399-417.

Meyer, J. W. (1983). Centralization of funding and control in educational governance. *Organizational Environments: Ritual and Rationality, 179,* 197.

Mosqueda, E., & Maldonado, S. I. (2013). The Effects of English Language Proficiency and Curricular Pathways: Latina/os' Mathematics Achievement in Secondary Schools. *Equity & Excellence in Education, 46*(2), 202-219.

Olsen, L. (2010). Reparable Harm Fulfilling the Unkept Promise of Educational Opportunity for California's Long Term English Learners. . *California Together (Research Report)*.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292.

Saunders, W. M., & Marcelletti, D. J. (2012). The Gap That Can't Go Away The Catch-22 of Reclassification in Monitoring the Progress of English Learners. *Educational Evaluation and Policy Analysis*, 0162373712461849.

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.

Uriarte, M., Lavan, N., Agusti, N., & Karp, F. (2009). English Learners in Boston Public Schools: Enrollment and Educational Outcomes of Native Spanish Speakers.

Walqui, A., Estrada, P., Koelsch, N., Hamburger, L., Gaarder, D., Insurralde, A., . . . Weiss, S. (2010). *What are We Doing to Middle School English Learners?: Findings and Recommendations for Change from a Study of California El Programs (Research Report)*.

# CHAPTER 1

A Mixed-Method Study: Districts' Language Classification Policies and the Implications for Male, Hispanic, and Low-Income Middle School Students


## Abstract

At minimum, California requires English Language Learners (ELL) to pass the California English Language Development Test and the California Standards Test in English Language Arts to be Reclassified Fluent English Proficient (RFEP). However, most districts set even higher reclassification requirements than the state, making it more difficult to reclassify. Currently, about 25% of California middle school students are classified ELL, and the majority of them are long-term English language learners. Often researchers attribute the classification to students' characteristics rather than being the result of districts' imprecise policies. Building on organizational models of education, in order for language classification to be a tightly coupled process: 1) Administrators must set fixed language classification criteria on which administrators and teachers agree. 2) Administrators must evaluate how classification polices are implemented and modify their process based on their evaluation. With this model I examine how school districts establish and implement their language classification policies.

In this study, I examine ELL reclassification policies in two Southern California school districts. In Manzanita Unified School District, administrators described their policies as more subjective. In addition to students' test scores, they also considered the teachers' and parents' opinions. Granada Unified School District administrators, on the other hand, approached the classification process more objectively. They based reclassification primarily on students' test scores, along with a district essay exam. Nevertheless, quantitative data show in both districts that male, Hispanic, and low-income students were less likely to have all the required assessment scores. Furthermore, male, Hispanic, and low-income students who passed all the district's or state's requirements were still less likely to be reclassified.

In both districts, there is a disconnect between administrators' intentions and how the language classification policies are implemented. Teachers may not be fully aware of or they may disagree with administrators' classification objectives. In most cases, the exams are gatekeepers to becoming RFEP. Teachers make exceptions, but only for students with certain characteristics. I recommend that administrators work together with teachers to establish and implement language classification policies that best meet the needs of their particular ELL students.



*Keywords:* Language Classification, English Language Learners, Inequality, School Districts, Middle Schools, and Mixed-Methods

In California, 25% of K-12 students are English language learners (ELLs) (Hill, Betts, Chavez, Zau, & Bachofer, 2014), compared with only 11% of K-12 students nationally (Kohler & Lazarín, 2007). ELLs are students who speak another language at home and have not yet reached full English proficiency (Hanhnel, Wolf, Banks, & Lafors, 2014). They are one of the fastest growing student groups in the country, yet are also among the lowest performers on a broad range of educational outcomes (Capps et al., 2005; Maxwell, 2014). ELL students, on average, score lower than non-ELL students in English reading, writing, and comprehension, as well as in less language-intensive subject areas, such as mathematics and science (Edwards, Leichty, & Wilson, 2008; Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hampden-Thompson, Mulligan, Kinukawa, & Halle, 2008). Yet, the ELL classification itself may have unintended consequences if classified students do not have the opportunity to learn rigorous educational content or are not integrated with their non-ELL peers. Either the group's language skills, the stigma associated with the ELL label, or a lack of access to rigorous courses drives the achievement gap between ELL and non-ELL students.

One of the main goals for ELLs is to become Reclassified Fluent English Proficient (RFEP). For students, a RFEP classification affects the instructional services they receive, the curriculum to which they have access, how they are assessed, and the academic standards to which they are held. California sets minimum requirements on how to determine students' English proficiency levels, but individual districts have the freedom to add more rigorous requirements (Linquanti & Cook, 2013; Wolf et al., 2008); therefore, a student can be considered ELL in one district and RFEP in another. Hill, Weston, and Hayes (2014) demonstrate that 90% of California's districts set higher reclassification requirements than the state, making it more difficult for students to be reclassified. Some districts require ELL students to pass a higher threshold on state assessments than the state requires, whereas some also consider other factors (e.g., benchmarks, course grades, attendance, and participation). More research is needed to understand how school districts establish their language classification policies, and why they choose certain assessment factors. Often, researchers attribute classification status to students' abilities rather than to the result of stringent and imprecise policies. Understanding how language minorities are classified and reclassified is critical, given evidence showing that male, Hispanic, and low-income students are more likely to be categorized as ELL than their Asian American and high-income peers in elementary school, even when accounting for their English proficiency (Grissom, 2004; Halle, Hair, Wandner, McNamara, & Chien, 2012; Kohler & Lazarín, 2007; Slama, 2014). What is unclear is the role student demographics play in the language classification of middle school students while accounting for prior achievement.

Administrators develop policies based on the state's minimum requirements and adapt them to meet the particular needs of their student body. While districts are establishing their own language classification policies, different individuals impact how these policies are implemented. For example, teachers can make the final decision as to whether a student should be reclassified based on various factors (Hill, Weston, et al., 2014). Teachers in some districts may be limited to basing their decisions strictly on assessments, while other districts may not provide teachers with any direction. This study explores how districts construct their policies and, ultimately, how these policies are implemented. The districts' language classification philosophies and teachers' beliefs about individual students may affect the extent to which student demographics affect reclassification rates.

My work focuses on middle school because it is a significant schooling stage that often determines the educational foundation of adolescents' high school and college years (Walqui et al., 2010). Middle school ELL students are also less frequently studied than ELL elementary students yet they are more likely to be Long-Term English Language Learners (LTELL) and foreign-born children (Hahnel, Wolf, Banks, & LaFors, 2014; Olsen, 2010). Of course, ELL middle school students are a diverse and complex group, stemming from differences in English proficiency, national origin, socioeconomic status, previous schooling, and the number of years in the U.S. system (Callahan, 2005; Krashen & Brown, 2005). For this reason, many aspects must be considered when discussing ELL students' academic needs and how educators can improve their academic and behavioral outcomes. Specifically, I address the following three research questions which focus particularly on middle school students:

1) How do school districts establish language classification policies? (Qualitative)
2) How do school districts implement the language classification policies? (Quantitative)
    a. What roles do student gender, race and socioeconomic status play in the language classification process?
3) How would school districts' reclassification rates change if the state minimum language classification policies are implemented? (Quantitative)

To investigate the specific situation of middle school learners, I use data from two Southern California districts.[3] For the first research question, I investigate the district administrators' rationale in implementing higher language classification requirements than the state requires, focusing on interviews with district administrators regarding language classification policies. For the second research question, I focus on ELL students who have not reclassified as of sixth grade, the beginning of middle school. I examine how districts' language classification policies are actually implemented, and if male, Hispanic, and low-income students

[3] In this study, I examine the years prior to California's recent implementation of the Common Core State Standards (CCSS), the Smarter Balanced Assessment Consortium (SBAC), and the increase in per-pupil funding for ELL students, each of which will almost certainly affect future ELL policies and practices (Hill, Weston, et al., 2014; Umansky et al., 2015). In the discussion section of this work, I will address the implications of my findings as they pertain to the new policy changes.

are less likely to be reclassified. In the third research question, I demonstrate how many more middle school students would be reclassified if the districts merely followed the state's minimum requirements. This study will deepen our understanding of how districts establish and implement their own language classification practices, particularly for middle school students.

## Literature Review

### California State and District Policies Regarding Proficiency Classification

During initial school registration, which usually occurs in kindergarten, California public schools administer the Home Language Survey (HLS), which asks parents whether a language other than English is spoken at home. If the answer is no, their children are classified English only (EO). If the answer is yes, their children must take the California English Language Development Test (CELDT),[4][5] which assesses their children's English proficiency. Students who pass the CELDT the first time are identified as "Initially Fluent English Proficient" (IFEP), and those who do not pass are identified as ELLs (Edwards et al., 2008). The California Department of Education (CDE) also has required that districts use the California Standards Test in English Language Arts (CST ELA) to determine the initial classification for students who arrive in California schools in grades three and above.[6][7] The CELDT and CST ELA scores have also been used to reclassify ELL students as RFEP. For a student to be reclassified, the CDE requires students to score "intermediate" or higher (at least 3 out of 5) in the listening, speaking, reading, and writing subcategories of the CELDT; "early advanced" or higher (at least 4 out of 5) overall

---

[4] In the near future the English Language Proficiency Assessment for California (ELPAC) will replace the CELDT (Umansky et al., 2015).
[5] The CDE requires that students with disabilities to take the CELDT but may receive accommodations based on their individualized education programs (IEPs) (http://www.cde.ca.gov/ta/tg/el/celdtfaq.asp).
[6] Starting in the 2014-2015 school year, the Smarter Balanced Assessment Consortium (SBAC) replaced the CST. This change has implications for future ELL students' reclassification process. This present study focuses on classification policies prior to this change, but the implications of my results on the new assessment will be addressed in the discussion section of the present study.
[7] The CDE only exempts students with significant cognitive disabilities from the CST (http://www.cde.ca.gov/ta/tg/sr/cefstar.asp).

on the CELDT; and “basic” or higher (at least 3 out of 5) on the CST ELA. Classification decision-makers can consider teacher and parent recommendations as well, but the extent to which these recommendations are incorporated depends on the district. Further, because individual districts can prioritize different components independently, about 90% of California’s districts set even higher reclassification requirements than those set by the state. Logically, those with more stringent criteria have lower reclassification rates (Hill, Betts, et al., 2014; Hill, Weston, et al., 2014). Different test threshold requirements and different recommendation considerations can lead to a student being categorized as an ELL in one district but not in another. For example, in the Los Angeles Unified School District (LAUSD), teachers’ evaluations are the most highly-considered component, but in the San Diego Unified School District (SDUSD), higher CST score standards are set, where students are required to score “mid-basic” on the CST and “early advanced” on at least three of the CELDT subcategories (Hill, Betts, et al., 2014). If LAUSD used SDUSD standards, 25% of their students would face delayed reclassification; inversely, if SDUSD used LAUSD standards, 70% of their students would face delayed reclassification.

**Implementation of Policies Varies Between and Within Districts**

Because of different implementation of district policies, sometimes more students are reclassified than are eligible, while, at other times, the reverse is true (Abedi, 2008; Estrada & Wang, 2013). Different reclassification rates can be attributed to subjective district reclassification criteria, such as teacher evaluations, parent recommendations, course grades, and district teachers’ graded exams. Furthermore, evidence shows that reclassification rates can vary based on specific student characteristics, for example, current grade, race, age, and immigration status, as well as district characteristics, like neighborhood poverty and school funding incentives.

**School level differences.** Umansky and Reardon (2014) demonstrate ELLs are more likely to be reclassified at the end of a school cycle, such as in the fifth, eighth, and eleventh grades, with reclassification rates of 39%, 62%, and 75%, respectively; however, their sample only includes Latino students who enrolled in the district in kindergarten. In early grades, they also find more students are eligible for reclassification than are actually reclassified, while the reverse is true for middle and high school students. Evidence suggests a districts' composition can influence the reclassification rates, but the effects may differ by grade level. Hahnel et al. (2014) estimate 15% of elementary students who attend high poverty and predominantly Spanish-speaking districts are reclassified, compared with 20% of students who attend low poverty and predominantly Spanish-speaking districts and 30% of students who attend districts where other non-English languages are predominant. Middle school reclassification rates are more wide-ranging, where 20-40% of students who attend high poverty and predominantly Spanish-speaking districts are reclassified. This percentage range can be compared with 25-30% of students who attend low poverty and predominantly Spanish-speaking districts and 30-35% of students who attend districts where other languages are predominant. Thus, the relationships between a district's composition and its reclassification rates are more variable in middle schools than elementary schools.

Different reclassification rates can also be influenced by conflicting accountability requirements and funding incentives. Most key provisions affecting limited English proficient and immigrant students are established in Title I and Title III of the Elementary and Secondary Education Act (ESEA). Schools and districts have an accountability-driven incentive to keep their top performing English language learners classified as ELLs in order to have higher achievement scores meet Title I requirements (Christopher & de Alth, 2005; Slama, 2014). Schools and districts failing to make adequate progress in this area are subjected to interventions, such as allowing parents to send their children to other schools, or offering supplemental after-school programs, or to more extreme consequences, such as restructuring, or even closing the school.[8] On the other hand, Title III incentivizes districts to reclassify students as quickly as possible to demonstrate that a greater number of their students have reached proficiency. The promise of financial gain in either keeping or reclassifying students can result in students being wrongfully classified, which, in turn, can lead to inappropriate services. Christopher and de Alth (2005, p. 50) assert "The size of Title I apportionments dwarfs those of Title III, so districts and schools face stronger incentives to hold back high-performing English learners rather than reclassify them." Ultimately, the ELL classification has many implications for the student opportunity to learn, and unspecified methods may result in some students being unfairly placed into or excluded from ELL classification.

**Student level differences.** Some evidence shows that Hispanic students are overrepresented in the ELL category (Halle et al., 2012; Kohler & Lazarín, 2007). For instance, Kohler and Lazarín (2007) report first- and second-generation Hispanic students make up 58% of the total population of children of immigrants,[9] yet they represent more than 75% of ELL students. In contrast, Asian students make up 22% of the total number of children of immigrants, but represent only 13% of ELL students (Kohler & Lazarín, 2007). Furthermore, ELL students who do not qualify for the National School Lunch Program as well as those who speak another language besides Spanish, and female students are more likely to be reclassified as RFEP (Abedi,

[8] Title I rules changed, the state provides more money per ELL pupil and districts now have the freedom to allocate those resources as they choose creating a greater incentive in keeping language minorities reclassified ELL.
[9] Tienda and Mitchell (2006) provide a thorough discussion on this topic and define "children of immigrants."

2008; Grissom, 2004; Slama, 2014),[10] but this might not always be the case when accounting for students' English proficiency (Grissom, 2004),[11] showing students' background characteristics may predict an ELL student's reclassification probability. Yet, these studies only focus on elementary school grades, and it is uncertain whether administrators make language classification decisions based on students' background characteristics, or if other, unmeasured factors explain the overrepresentation of Hispanic, male, and low-income students' in the ELL category.

## Expanding on the Language Classification Literature

Often researchers use socially constructed categories to compare students' academic achievements without questioning the categories themselves. Therefore, we need to investigate how educators construct their language classification categories, particularly when school districts have the freedom to establish their own language classification policies. Building on organizational models of education, school districts are "fragmented centralized" organizations, where procedures can be either tightly or loosely coupled (Fusarelli, 2002; Meyer, 1983, p. 181). In multi-layered educational systems, different actors play roles when constructing and implementing policies, and administrators decide if they should add requirements beyond the state's. If so, they choose which factors to add and to what extent these factors should matter. At the end of the language classification process, teachers and parents may be able to make final reclassification decisions. However, their recommendations vary based on the districts' policies and their awareness of those policies. The process can be tightly or loosely coupled, depending on the relationship between district administrators, teachers, and parents.

[10] Slama's (2014) sample is drawn from Massachusetts.
[11] Grissom (2004) uses 1998-2003 data, from when students took the statewide norm-referenced test (NRT) and the Stanford Achievement Test, version 9 (SAT-9) instead of the CST.

Educational processes that are tightly coupled share four characteristics: 1) specified rules, 2) agreement on what those rules are, 3) a system of inspection to see if compliance occurs, and 4) feedback designed to improve compliance (Weick, 1982). Administrators would set fixed language classification criteria on what administrators and teachers have agreed to prior. Further, administrators would evaluate how classification polices are implemented and modify their process based on their evaluation. In this study, I examine the language classification process in two Southern California school districts. Usually, the ELL classification is attributed to student characteristics (e.g., special needs) rather than because of restrictive and imprecise policies. I will compare administrators' reclassification policy descriptions (qualitative analyses) and the actual implementation of those policies (quantitative analyses) to understand the possible disconnect between policies and implementation. I hypothesize a disconnect in districts with a loosely coupled language classification process.

There are policy implications whether certain student background characteristics determine students' language classification, or if they only highly correlate. Evidence indicates that child and family characteristics (e.g., race, language, parent's education, income, and family size) predict English proficiency by kindergarten (Halle et al., 2012). However, the extent to which these background characteristics predict English proficiency in middle school remains unknown. If certain students struggle to meet proficiency standards, we could provide more resources for these students to increase their reclassification probability. If students are being unfairly classified, we would need to improve how educators determine classification within districts, though I do not expect middle school students' background characteristics with regard to English proficiency to be as relevant as those for elementary students' because they have more years of school exposure. In any event, because being classified ELL can have unintended consequences for students it is crucial to understand how educators make these decisions, and determine whether biases based on ethnicity or socioeconomic status exist.

The purpose of this study is threefold. First, I aim to understand how districts establish their reclassification policies. Second, I examine how district polices are actually implemented, and assess whether students' demographics play a role in determining their middle school language classification. Third, I contrast the school districts' language classification policies against the state's minimum requirements. Several studies aid me in these goals, such as work providing descriptive reclassification rates (Abedi, 2008; Grissom, 2004; Hahnel et al., 2014), others that use logistic regressions to estimate reclassification rates for subgroups of students (Grissom, 2004; Halle et al., 2012; Umansky & Reardon, 2014), and others that estimate the probability of being reclassified if students passed all the measurable reclassification near the cut-off (Robinson-Cimpian & Thompson, 2015; Robinson, 2011). I will expand on these prior language classification research studies, and include special education, foreign-born, and students who entered the school district in grades other than kindergarten to gain a more accurate representation of ELL middle school students. Prior research studies have excluded these students (Halle et al., 2012; Robinson, 2011; Umansky & Reardon, 2014).

## Data Source and Sample

I accessed district data through the Spencer-funded Evaluating the Quality of Universal Algebra Learning (EQUAL) project. Table 1.1 shows the demographic breakdown of both districts, which I will describe in more detail in the quantitative section. I focused on middle school students from two diverse Southern California school districts. I chose to focus my investigation on the Manzanita Unified School District (MUSD), and the Granada Unified School District (GUSD) because they have large ELL percentages, and, more specifically, large numbers of Hispanic and Asian students. Comparing these two districts also afforded insight into practices that prevail in relatively low-income communities. The project provided qualitative data through interviews and district documents related to reclassification policies. The quantitative data included student-level demographic, language classification, transcripts, and achievement data from district administration records.

In this mixed-methods study, I used the concurrent triangulation strategy (Creswell, 2013) where I collected both quantitative and qualitative data concurrently, and then compared the results to determine if there was convergence, differences, or some combination. First, I present a qualitative piece that describes the language classification policies based on interviews with district administrators. Second, I present a quantitative piece to determine how those language classification policies are implemented in actuality. I also compare reclassification rates based on school districts' policies and state minimum reclassification policies.

## Qualitative Section: Districts Establish Language Classification Policies

### Qualitative Methods

Between March, 2015 and November, 2015, I interviewed three MUSD and four GUSD district administrators (i.e., directors of the English Language Programs, Language Assessment Center supervisors, and administrators of K-12 Educational Services) because language classification policies are established at the district level. The purpose was to understand the district classification policies, and how the administrators viewed the implementation of those written policies. First, I formally interviewed each district administrator (see interview questions in Table 1.1A in the Appendix). Next, during the summer months (June, 2015-August, 2015), I worked for both districts, and had several informal conversations with these same district administrators, as well as a few others. Finally, between January, 2016 and July, 2016, I shared my qualitative and quantitative results with the district administrators, and, based on their feedback, conducted more analyses.

**Qualitative Analysis**

To address the first research question, I used interviews with district administrators. I conducted formal interviews, which typically lasted 30-45 minutes (see Table 1.1A in the Appendix for the interview questions). I wrote shorthand notes as I conducted the interviews, and after each interview—on the same day—I edited and typed each participant's complete responses. District administrators described their job responsibilities to provide a greater understanding of how they have been involved, directly or indirectly, in classifying language minorities. They also discussed whether they thought they had influenced the academics of language minorities in any other way. Additionally, the district administrators described the language classification policies of their district. When necessary, I probed further, asking which of the following components were considered, and to what extent: the CELDT, CST ELA, ELA course grade, teacher recommendations, and parent recommendations. I also asked if they thought the Smarter Balanced Assessment Consortium (SBAC) would provide the same information as the CST scores after it replaces the CST. This was for general information only, as my quantitative analysis did not include SBAC scores; the purpose of this question was to gain greater understanding regarding future language classification policies. I assigned each district administrator a number to keep his/her personal responses confidential and secure.

During the summer months, I had several informal conversations with these same administrators. In these more conversational talks, they provided me with different information that included explanations for inconsistencies with the student-level data, and the implementation of the language classification process in different school years. Once I had preliminary qualitative and quantitative results, I discussed the results with my participants as a group to get greater clarification on any inconsistencies.

I provide descriptive information. First, I show the months in which the students had to take the required reclassification exams, the months when the students were reclassified, and the scores used for reclassification. Second, I provide the districts' reclassification requirements compared with the state's minimum requirements, as well as a comparison with the average California district requirements based on the Hill, Weston, et al. (2014) study.

## Qualitative Results: Districts Establish Language Classification Policies

### District Administrators' Descriptions of Language Classification Policies

To address how school districts establish language classification policies particularly for middle school students (Research Question 1), I provide the district administrators' policy descriptions and rationales. Generally, MUSD administrators described a more subjective approach when they accounted for a student's test scores, but also noted they considered parent opinion. On the other hand, GUSD administrators approached the classification process more objectively, basing language classification primarily on students' test scores. Both districts conducted reclassification in the spring, but GUSD also allowed students to reclassify in the fall as well.

**MUSD policies.** The students who were reclassified in the spring were chosen based on that school year's CELDT scores and the prior school year's CST scores (see Table1.2). For example, for a given seventh-grade girl, her seventh-grade CELDT and sixth-grade CST scores were used. Prior CELDT and CST scores could be used if the current scores were unavailable. Once a student's classification was determined, the parents were informed about the recommended classification. Parents could then request their child's classification be changed, regardless of their child's test scores. As the Language Assessment Center's supervisor explained, "I meet with children's parents to see if there are other factors that can explain their English proficiency. I go on a case-by-case basis." The Center's supervisor believed some students can have an off day and the assessment score may not accurately reflect their English proficiency. He described his role as essential to the reclassification process. The district's reclassification documents also specified that teachers' recommendations should also be considered, and that those should be based on a student's cumulative grade point average (GPA); however, administrators also stated the GPA criterion was not followed in practice. Most teachers make recommendations based on unspecified criteria.

**GUSD policies.** The students were reclassified in the spring based on that current school year's CELDT, the prior school year's CST, and the essay scores (see Table 1.2). For example, for an eighth-grade girl, the school district used her eighth-grade CELDT score, seventh-grade CST score, and seventh-grade essay score. However, for an eighth grader reclassified in the fall, the district used her seventh-grade CELDT, CST, and essay scores. Thus, the CELDT was the main difference between the fall and spring. At GUSD, students were allowed to retake the CELDT throughout the year, but the CST and the essay—the more difficult tests to pass—could only be taken once each school year. The school district created the essay exam, and, at the end of each school year, all ELL and non-ELL students are required to take this exam. EO, IFEP, and RFEP students were also required to take the essay exam in order to determine English course placements. Teachers administered the essay exam at the end of the school year around the same time they administered the CST exam. English teachers read and scored each essay based on a predetermined 1 to 4 rubric scale. One administrator explained "administrating and scoring the district writing assessment can be labor intensive for teachers." For this reason, students were not allowed to take the essay exam on another day other than the assigned test date.

Granada's policies also permitted students' prior CELDT and CST scores be used if their current scores were missing. According to the Director of English Language Programs, "We want students to reclassify, so we use what we have." However, the director also stated ELL students without essay scores were considered a no pass. Administrators also recognized that the essay was the hardest requirement to pass and that most ELL and non-ELL students failed this exam. As a final step in the reclassification process, the district provided the recommended classification to the child's English teacher. According to administrators, most teachers followed the district's recommendation, especially when the student had passed all the district requirements (i.e., CELDT, CST, and written essay). Parents were then informed about their child's classification, but were not allowed to argue in favor of changing their child's classification status. At the same time, parents could choose to opt of ELL resources for their child, if that was their wish.

**Exceptions to the districts' policies.** Administrators in both districts explained that students with disabilities and students who had been in the U.S. for less than a year would be exempted from taking the CST. However, the districts did not collect information about when students entered the United States. As one of MSUD administrator explained, "Many of our students go back to their countries and then come back. It is complicated to determine how many years the student has actually been in this country. In some instances, students are born here, but then they move to Mexico for a few years, and later they come back." The state of California only exempts students who have been in the U.S. for less than one year, but the individual school districts have not collected this information. A GUSD administrator explained, "this type of information can be too sensitive and controversial to collect." Based on these comments, the majority of students without the required reclassification assessments should be students with disabilities.

**Comparison to other California districts.** Both districts set higher minimum requirements than the state requires (see Table 1.3). Both MUSD and GUSD required ELL students to score a 3.5 or higher on the CST, while the state only required a 3.0 or above. Hill, Weston, et al. (2014) show about 45% of California school districts required an ELL to score 3.5 or higher on the CST, and 27% of California school districts required a 4.0 or higher. Furthermore, MUSD set higher CELDT subcategory requirements, where ELL students had to score a 4.0 or higher in reading and writing. About 35% of California school districts set higher CELDT subcategory requirements (Hill, Weston, et al., 2014). As discussed above, GUSD also required ELL students to pass an essay exam created by the district, and only 9% of California school districts had a similar criterion (Hill, Weston, et al., 2014). MUSD teachers were instructed to base their evaluation on a student's GPA, but, as mentioned, according to interviews teachers did not follow this criterion. In comparison, GUSD did not formally instruct their teachers to base their recommendations on any specific criteria. Most California school districts (78%) required teacher recommendations as well, but what went into those recommendations varied by school district (e.g., grades, assessments, discipline, and attendance).

Similar to most California districts, MUSD and GUSD set higher language classification policies than the state requires. District administrators' language classification policy descriptions appear straightforward and noncontroversial. Administrators from both districts believed it essential that ELL students were not reclassified too early because, once reclassified, the student would lose language support. However, neither district provided evidence that reclassification can harm students. GUSD administrators in particular were concerned that the CELDT and CST were not rigorous enough to determine English proficiency, and they therefore added the additional written essay component. GUSD administrators did not express concern the essay was created to determine course placement for non-ELL students and not necessarily to measure English proficiency. They also reported only half of all GUSD students passed the essay exam.

These qualitative results informed me how district administrators establish their language classification policies. The remainder of the study focuses on student-level district data. In the discussion section I will provide concluding remarks regarding both analyses.

## Quantitative Section: Language Classification Policies Implemented

### Quantitative Methods

In my quantitative data work, I used student-level district data to examine how language classification policies were implemented. Table 1.1 shows the district and student demographic information I analyzed. For the Manzanita district, the student-level data included three cohorts of eighth graders from 2010-2013, whereas, for the Granada district, it comprised two cohorts of eighth graders from 2012-2014. I selected these years because language classification policies are fluid and implemented differently in individual school years. In MUSD, administrators were starting to consider different assessments in 2013-2014 in anticipation of the fact that the CST was not going to be available in the future. GUSD, on the other hand, had required ELL students to pass the CELDT, the CST, and an essay exam. However, only a third of students actually took the essay exam in 2009-2011 because it was still being established. It was not until 2012-2013 that the percentage of students who took the exam increased to 85%.

Table 1.1 illustrates the demographic breakdown of both districts. This table provides information on the full sample, the selected sample, and the students with complete data. As my analysis focused on ELL middle school students, I restricted the data to students who were classified ELL in middle school (26% in Manzanita, and 38% in Granada) and referred to this group as the "selected sample." I excluded EO, IFEP, and students who had been reclassified RFEP in elementary school. I also excluded both White and African American students, who were mostly EO students. The first half of my analysis focused on the selected sample. Therefore, the results can be generalized to all middle school students classified as ELL.

I further restricted the data to students who had both CELDT and CST scores, referring to them as the "complete sample." Students in the selected sample (n=4,231 in Manzanita, n=2,905 in Granada) were different from the students in the complete sample (n=2,969 in Manzanita, n=2,054 in Granada). As Table 1.2A in the Appendix shows, in MUSD, the final sample had fewer special education, foreign-born, non-FRL, Asian American, and female students than the selected sample. Similarly, as Table 1.3A in the Appendix demonstrates, in GUSD, the final sample also had fewer special education, foreign-born, and Asian American students.

**Quantitative Measures**

**Reclassification.** The dependent variable was coded 1 for RFEP, and 0 for ELL. Here, I first compared seventh graders who remained ELL versus those students who were reclassified RFEP in the seventh grade. Secondly, I compared eighth graders who remained ELL versus those students who were reclassified RFEP in the eighth grade.

**Reclassification assessments.** An ELL student had to pass all requirements with their respective different cutoff scores, depending on their district and grade level, to be reclassified. Failing to meet even one of the requirements would have been enough to prevent a student from being reclassified. The CELDT's overall raw scores and the pass cutoffs varied by school grade. For middle school students, the overall CELDT ranged from 248-741, but both districts required a score of at least 556.[12] The CST raw scores ranged between 150 and 600, and both districts required students score at least 325, considered "mid-basic." In MUSD, according to documents, eighth graders were also required to have a cumulative GPA higher than 2.0 in a range between 0.0-4.0. In GUSD, students were also required to pass a district-wide written essay exam, scored on a scale of 1-4. A passing score was at least a 3 or higher for seventh and eighth graders.

**Assessment outcomes categorized.** Options for Manzanita ELL students included passing the CELDT and CST (Passed the District's Criteria), failing the CELDT and/or CST (Failed One or All), no CELDT or CST scores (One Criterion Missing), or no CELDT and CST scores (All Criteria Missing). Options for Granada ELL students included passing the CELDT, CST, and Essay (Passed the District's Criteria), passing the CELDT and CST but failing the essay (Passed CELDT and CST only), passing the CELDT and Essay but failing the CST (Passed CELDT/Essay only), failing the CELDT, CST and Essay (Failed All), no CST or Essay (One Criterion Missing), or missing any of the required exams (All Criteria Missing). These categorical variables were used to determine which students should be included in the final reclassification rate models. These categories did not include the CELDT subcategories (e.g., reading or writing) because most students who passed the overall CELDT also passed the subcategories requirements, and thus the main results did not change without the subcategories.

**Control variables.** The models also included student-level covariates to explain

[12] The CELDT subcategories' raw scores and pass cutoffs also varied. The scores for English language listening ranged from 230-715, speaking 225-720, reading 320-750, and writing 220-780. Granada seventh graders were required to score at least a 495, 476, 529, and 508 in listening, speaking, reading, and writing, respectively. Eighth graders had to score at least a 508, 480, 543, and 511 in listening, speaking, reading, and writing, respectively. Manzanita had similar requirements, with the exceptions that they required a 572 in reading and a 554 in writing for seventh graders, and a 588 in reading and a 557 in writing for eighth graders.

differences in the students' language classification. These covariates included gender (1=female, 0=male), race/ethnicity, birth country, socioeconomic status (SES), and special education status. Racial/ethnic categories included Hispanic (reference group), Asian American, and an "other race" category, including American Indian, Alaskan Native, Filipino, Native Hawaiian, and Pacific Islanders. The "other" race category sometimes only included a few students, therefore, sometimes it was combined with Asian American students. The birth country was a binary outcome, where 1 was coded for those "born in the United States," and 0 was coded for those "born in another country." SES was based on students' "free or reduced lunch" (FRL) status, where students who qualified for FRL (reference group) were compared to students who did not qualify for FRL status. Special education status was binary, where 1 was coded "special education," and 0 was coded "no special education."

The models also included school-level covariates to explain differences in the students' language classification. The models included the student's current middle school (eight schools in Manzanita, and ten schools in Granada) as well as their seventh and eighth grade English teachers who were asked, to some extent, to make the final language classification decisions.

**Quantitative Analysis**

I conducted logistic regression models to specifically address how school districts implement their language classification policies, as well as to determine what roles student gender, race, and socioeconomic status play in the language classification process (Research Question 2). For each district, I conducted four logistic regression models to predict the odds that a student would be reclassified RFEP in either the seventh or eighth grade, based on that student's gender, race, and SES, while also accounting for whether they were born inside or outside of the U.S., had special education needs, and whether they passed the district's reclassification criteria, their cohort year, their eighth grade school, and English teacher fixed effects. This main model can be expressed as:

$$\ln\left[\frac{p\,(LC)}{1-p\,(LC)i}\right] = \beta_0 + \beta_1\, GEND_i + \beta_2\, RACE_i + \beta_3\, SES_i + \beta_4\, TESTS_i + \beta_5\, Controls_i + F\delta_{s(i)} + e_i \qquad \text{(F1)}$$

In (F1), $\ln\left[\frac{p\,(LC)}{1-p\,(LC)i}\right]$ is a variable representing student $i$'s log odds of being reclassified RFEP (reference group) in middle school. In the first two models, I compared seventh-grade RFEPs with seventh-grade ELLs for both districts separately. In the second two models, I compared eighth-grade RFEPs with eighth-grade ELLs for both districts separately. The models titled "District" [Models 1 and 3] include students who passed all the district's language classification requirements; for Manzanita, it includes CELDT, CST 325, and GPA (for 8$^{th}$ grade only), and for Granada it includes CELDT, CST 325, and essay. These models demonstrate the extent assessments and students' demographics explain reclassification for students who met the districts' standards. The models titled "State" [Models 2 and 4] include all students who passed the state's minimum requirement of 556 or higher on the CELDT and 300 or higher on the CST. These models show the extent assessments and students' demographics explain reclassification for students who met the state's minimum standards.

F1 model included $GEND_i$ a dummy variable, coded 1 for female and 0 for male. $RACE_i$ includes Hispanic (reference group), Asian, and "other race." $SES_i$ includes a student's FRL status (1=qualified for FRL, 0=did not qualify). $TESTS_i$ includes the dichotomous variable passing or failing the CST at the 325 district cutoff. In the Granada "State" models also include whether a student passed (1), failed (2), or did not have the essay scores (3). Other student-level covariates ($Controls_i$) included birth country and special education status. In addition, cohort year, current middle school, and teacher fixed effects were also included to control for annual changes and school factors (e.g., other unmeasured confounders) that might also have explained classification.

For Granada, I expect *TESTS*$_i$ to be the strongest indicator of reclassification because they rely heavily on assessments. Granada's administrators describe their policies as rigid, and they expect teachers to base reclassification on student performance. It is unlikely that Granada students who pass the state's minimum requirements, but not the district's requirements, will be reclassified. For Manzanita, I expect *TESTS*$_i$ to determine student reclassification to a lesser extent, because administrators, teachers, parents, and students can request exceptions. Teachers in particular make the final reclassification decision without any specific district guidelines. Here, gender, race, and SES may be strong indicators of student reclassification because teachers may be unconsciously biased against certain subgroups. Furthermore, parents and students can contact the Language Assessment Center's supervisor and request the language classification to be changed. Certain demographics may be more inclined to request reclassification exceptions.

To address how districts' reclassification rates would change if the state minimum language classification policies were implemented (Research Question 3), I only included students with complete data in the analyses. I demonstrate the percentage of students who reclassified RFEP by district and grade level. I also show the percentage of students who would have reclassified if the districts reclassified solely based on CELDT and CST at the 325 cutoff. The end of the table shows the percentage of students who would have reclassified if the districts reclassified all students who passed the state's minimum reclassification of 556 on the CELDT and 300 on the CST.

## Quantitative Results: Language Classification Policies Implemented

### Middle School English Language Learners

Many language minorities were classified ELL when they started middle school, but only a few reclassified by the end of middle school. Table 1.1 shows that one-quarter of the Manzanita sixth graders and two-fifths of Granada sixth graders were classified as ELL. Of those, 19% and 16%, respectively, were reclassified in the seventh grade. By the eighth grade, 28% of MUSD and 26% of GUSD students were reclassified RFEP. District administrators' descriptions of their language classification policies appear straightforward. However, the student-level data shows that several students did not have all the required assessment scores, and the extent to which the policies were followed depended on the individual student's current grade level. Further, there is evidence that reclassification rates varied by gender, race, and SES.

### Implementation of Language Classification Policies

To address Research Question 2, I calculated the reclassification rates and missing scores by gender, race, SES, special education, and grade level. Table 1.4 shows that males, Hispanics, students who qualify for FRL (henceforth referred to as "FRL students"), and special education students were less likely to reclassify than their counterparts. However, it is uncertain if these reclassification rates were lower because these students were less likely to pass the reclassification requirements, or if these students were unjustly not being reclassified. The next section of this work demonstrates that, in some instances, student demographics can be strongly associated with not having access to the district's required assessments.

**Missing reclassification requirements.** Not having the required assessments could preclude students from being reclassified. Table 1.5 shows ELL students were more likely to be missing reclassification criteria than RFEP students (see bold percentages). For example, in

MUSD, only 64% of ELL seventh graders had both CELDT and CST scores, compared with 95% of RFEP students. This rate was comparable at GUSD. For example, in GUSD, only 62% of ELL seventh graders had CELDT, CST, and essay scores, compared with only 92% of RFEP seventh graders who had all three requirements.[13] The percentage of GUSD students not having all the required scores did not greatly change if I only considered the CELDT and CST exams. Table 1.5 also shows the percentage of students with and without all the required assessments when excluding special education students (percentages in the first parentheses) and then also excluding foreign-born students (percentages in the second parentheses). ELL students were still more likely to be missing reclassification criteria than RFEP students. In both districts, only a few students reclassified with none, or only one, of the district's required assessments (see Figures 1.1 and 1.2).

A Pearson chi-square test showed Manzanita's male, Hispanic, and special education students were less likely to have all required scores (see Figure 1.3). In Granada, male, Hispanic, "other" race, and special education students were less likely to have all the required scores (see Figure 1.4). For both districts, these results were similar for eighth graders. Though the CDE has excused students with severe disabilities and students who have been in the country less than one year, the exemption of the required tests seemed to go beyond these specific exemptions. Certain groups of students were more likely to be missing required assessments in logistic regressions that accounted for students' special education status, country born, cohort, and school fixed effects. In Manzanita, seventh-grade females and Asian American were more likely than their male and Hispanic peers to be missing one or all reclassification criteria in the logistic models with control variables (see Table 1.2A in the Appendix). Furthermore, eighth graders who

[13] Missing essay scores by school and classroom were comparable. Specific schools or teachers were not opting out of offering the essay exam.

qualified for FRL were less likely to be missing all reclassification criteria. Asian American were more likely than Hispanic students to be missing one reclassification criteria. As seen in Table 1.3A in the Appendix, in Granada, in both grades, Asian American, foreign-born, and special education students were more likely than Hispanic, U.S. born, and non-special education students to be missing one or all of the reclassification criteria. In both the Manzanita and Granada districts, administrators stated that students without the required scores could be reclassified based on scores from the prior school year. However, students who were missing current reclassification criteria were also likely to be missing prior reclassification criteria. In both districts, students without the required scores were likely to remain classified ELL. Not having the required assessment scores prevents ELL students from reclassifying. It is important to note that there were gender, racial, and SES differences regarding those who did not have the required scores. However, the gender and race differences wavered when control variables were added, showing that the gender and race differences occur through different mechanisms.

**Passing district's reclassification requirements.** Moreover, passing the districts' required exams did not guarantee reclassification. Figure 1.1 shows only 70% of Manzanita's students who passed CELDT and CST at 325 were reclassified. A few exceptions were made where 10% of students who passed the CELDT and the CST at the 300 cutoff were also reclassified. In comparison, Figure 1.2 shows 94% of Granada students were reclassified who passed the CELDT, CST, and essay. Exceptions include the 10% of students who met the state's minimum requirements (CELDT 556, CST 300) who were also reclassified, but most of these students passed the essay. Only about 3% of students who passed the CELDT and CST at 325 but failed the essay were reclassified. In both districts, English teachers distrust tests and prevent students from reclassifying. Educators believe that some students need to remain classified ELL even when they passed their more stringent assessment requirements. However, a few exceptions are made for individuals who meet the state's minimum requirements. The following analyses will demonstrate whether students' demographics played a role when educators used their discretion in keeping students classified ELL when they passed the district's or state's minimum requirements.

In Manzanita, a Pearson chi-square test showed that female, non-Hispanic, and non-FRL students were more likely to pass all reclassification criteria (see Figure 1.3). Meanwhile, in Granada, female, Asian American and non-FRL students had the highest CELDT, CST, and essay pass rates (see Figure 1.4). However, male, Hispanic, and non-FRL students had the highest CELDT and CST pass rates among those who failed the essay. The following analysis addresses how school districts implement the language classification policies (Research Question 2), explaining the extent to which student demographics determine the odds of being reclassified when students passed the district's requirements, and, in separate analyses, how implementation determined the odds of being reclassified when students passed the state's minimum requirements.

**Odds ratios of reclassifying in MUSD.** As Table 1.6 Model 1 demonstrates, in Manzanita, seventh-grade females were more likely to be reclassified RFEP (OR 1.78, $p < .001$) when the model only included students who passed the district's requirements (n=1,043) and accounted for students' demographics, current middle school, and teacher fixed effects. Further, in seventh-grade, Hispanic students were less likely to be reclassified RFEP (OR .49, $p < .001$) than non-Hispanic students. FRL students were also less likely to be reclassified (OR 0.58, $p < .01$) than non-FRL students. Lastly, several schools were more likely to reclassify students than other schools. The results remained the same when Mode 2 accounted for students who passed the state's minimum requirements (n=1,612). Female, non-Hispanic, and non-FRL students were more likely to be reclassified if the model included all students who met the minimum state requirements. School differences also persisted. On the other hand, students' demographics did not increase eighth graders' odds to reclassify when the models included students who met the district's requirements (Model 3), or when it included students who met the state's requirements (Model 4). Only a few schools' differences persisted in eighth grade. Fewer exceptions were made for eighth graders who met the state's 300 CST and not the district's 325 CST requirement, which can explain why student's demographics did not associate with reclassification.

**Odds ratios of reclassifying in GUSD.** The Granada "district" models only included students who passed the CELDT, CST, and essay (n=347 for seventh graders, n=308 for eighth graders). As Table 1.7 in Model 1 demonstrate, in Granada, FRL students were less likely to RFEP in the seventh grade (OR 0.48, $p < .05$), although they met the district's requirements. In both grade levels, a few schools were unlikely to reclassify students even when the student passed the CELDT, CST, and essay. These schools may have requirements in addition to the district's, or very few students at these schools would pass all three requirements. Granada schools did make some exceptions for students who met the state's minimum requirements, or who passed the essay, but failed either the CELDT or CST. Thus further analyses were conducted for individuals that met the state's minimum requirements (n=1,263 for seventh graders, n=1,125 for eighth graders). The sample size more than tripled. Model 2 demonstrates, female students were more likely to be RFEP (OR 1.79, $p < .001$) and Hispanic students were less likely to be RFEP (OR 0.44, $p < .001$) in seventh grade. Several schools were also less likely to reclassify seventh graders who passed the state's minimum requirements. Model 4 demonstrates female and FRL students were more likely to RFEP (OR 1.43, $p < .01$; OR 1.50, $p$

< .05), and Hispanic students were less likely to RFEP (OR 0.63, $p < .05$) in eighth grade. Only one school was less likely to reclassify eighth graders who passed the state's minimum requirements. Models 2 and 4 included whether the student passed the CST at 325, and whether the student passed or failed the essay exam. Students without an essay score, or who failed the essay, had a near zero probability of being reclassified. To be reclassified, students had to pass the essay, even though they might have failed the CELDT or CST. Granada teachers used some discretion when students only met two of the three requirements, but this lead to different reclassification rates by gender, race, and FRL status.

**District Implementation of the State Minimum Requirements**

Research Question 3 asks what percentage of ELL middle school students would become RFEP if the school districts implemented the state's minimum requirements. I compared only ELL and RFEP students who had both CELDT and CST scores (i.e., students with complete data). I compared their achievement scores by district and language classification. I then compared the percentage of students who met the minimum state requirements by language classification and district.

**Achievement distribution.** Figures 1.5(1a) and 1.5(1b) show that RFEP students from both districts had comparable CELDT scores, demonstrated by the red and black dashed lines. However, the Granada ELL students had a higher CELDT distribution of scores than the Manzanita ELL students. Similarly, Figure 1.6(2a) shows that RFEP seventh graders had comparable CST scores. However, Figure 1.6(2b) illustrates that GUSD's eighth-grade RFEP students had higher CST scores than MUSD's RFEP students. Additionally, in Figures 1.6(2a) and 1.6(2b), we see that GUSD's ELL students had higher CST averages than MUSD's ELL students. These figures demonstrate that the Granada district tends to keep students with higher scores classified ELL compared with those in the Manzanita district. Table 1.4A in the Appendix presents the average CELDT and CST scores by grade and language classification. This table also illustrates that the average essay scores for ELL students were below the required passing cutoff.

**State minimum requirements.** Significantly, more students would be reclassified in both districts if the districts based reclassification solely on the CELDT and CST; the same is true if the districts followed the state's minimum language classification requirement. As it currently stands, Table 1.8 shows that when comparing students with complete data, only 26% of Manzanita ELL seventh graders reclassified, but that this percentage would increase to 35% if reclassification was based solely on assessments, and would jump to 54% if the CST score were set at 300. From the remaining ELL students, the 14% RFEP who reclassified in the eighth grade would increase to 20% if reclassification was based solely on the CELDT and CST assessments, and it would grow to 40% if the students needed only a score of 300 on the CST.

Additionally, Table 1.8 shows that reclassification percentages would increase for Granada students as well. Whereas only 21% of Granada ELL seventh graders reclassified, 43% would be reclassified if there were no essay exam, and if reclassification was based solely on CELDT and CST scores. Further, the reclassification rate would increase to 63% if Granada required only a score of 300 on the CST. From the remaining ELL students, the 17% RFEP who reclassified in the eighth grade would increase to 41% if there were no essay exam, and if reclassification was based solely on the CELDT and CST scores. Further, reclassification would swell to 62% if the district required a CST score of only 300.

Moreover, many more students from both districts would reclassify if they simply took all of the required assessment tests. As previously explained, about 30% of students classified

ELL at the beginning of middle school did not take the CELDT, CST, or essay exam (at Granada only), and thus did not have the opportunity to reclassify.

## Discussion

Manzanita and Granada are typical California districts that set higher reclassification requirements than the state's minimums for their students. In addition to passing the CELDT with a score of 556, both districts require students score a 325 instead of a 300 on the CST. Granada also requires an essay exam that is scored by their English teacher. Manzanita administrators also explained that there was some flexibility, and that occasionally exceptions were made despite a student's test scores whereas, on the other hand, Granada administrators described their language classification policies as primarily based on students' test scores. In both districts, however, teachers ultimately make the final decision in implementing the policies. Taken together, findings suggest that both district-chosen language classification policies dramatically reduce reclassification eligibility despite administrators setting different policies.

Often the long-term English language learner concerned is more often considered an outcome of the student's characteristics, rather than the result of restrictive and vague policies. With an organizational model framework, I examine how administrators establish the districts language classification policies and how their policies are implemented. In both districts, the imperfect language classification process prevents many ELL students from reclassification. Manzanita's administrators may have encouraged reclassification be based on teacher and parent recommendation instead of assessments, but, in most instances, ELL students had to meet the district's higher than state standards to be considered for reclassification. Granada's administrators may have encouraged reclassification be based mainly on assessments, but English teachers made exceptions for students who failed the CELDT or CST, but passed the essay exam (that they themselves scored). The educational language classification process in both districts is loosely coupled, where administrators set language classification rules, and teachers are not necessarily following those policies. Furthermore, district administrators did not put in place an evaluation system to determine teacher compliance with their policies, and, likewise, no discussion as to how to improve the implementation process occurred. These policies can have particularly far-reaching impacts on male, Hispanic, and low-income student opportunities to get ahead.

Both districts implemented more rigorous language classification policies to ensure their ELL students are not reclassified too early—ostensibly so their students do not stop receiving the language resources they need. At the same time, both districts want to equitably determine each student's language classification. Manzanita administrators claimed to make exceptions when teachers, parents, or students requested otherwise. The administrators believed that considering student classification on a case-by-case basis empowered parents and students to choose the child's rightful language classification and resources. In practice, however, these exceptions are rarely made. On the other hand, Granada administrators viewed their language classification policies as equitable because they supposedly determine classification objectively. Nevertheless, Granada ELL students are also required to pass an essay exam that, in fact, most students fail, including non-ELL students. The assessments should have strongly predicted language classification in Granada, but less so in Manzanita.

Furthermore, in both districts a student's demographics should not have predicted language classification, yet findings show that male, Hispanic, and low-income students were less likely to reclassify. At Manzanita, male, Hispanic, and low-income students were less likely to be reclassified when only considering students who passed the CELDT and the CST with a

325 or higher. Even when teachers made exceptions for students who passed the lower state requirements, the outcome was not favorable for male, Hispanic, and low-income students. Similarly, at Granada, "other" racial, and low-income students were less likely to be reclassified even when only considering students who passed the CELDT, CST at 325, and the essay test. The impact of student demographics became more pronounced when the students who met the state's minimum requirements were considered. Male and Hispanic students were less likely to be reclassified in both grades, and low-income students were unexpectedly more likely to be reclassified in the eighth grade.

Based on these findings, I recommend that language classification be a tightly coupled process where administrators and teachers work together to establish and implement language classification policies. Administrators also need to evaluate how the classification policies are implemented and modify the process if their objectives are not met. Furthermore, administrators need to establish language classification policies based on empirical data supporting the supposition that their assessments measure English proficiency accurately. Districts need to determine if these higher requirements are the best way to measure their students' English proficiency. California allows districts to determine their own classification policies so they can meet their particular student body needs. The state does not want to set specific exams and exact cutoffs because ELL students are a diverse and complex group of students with different needs. Going forward, administrators must make more research-based decisions when it comes to setting and evaluating their language classification policies.

ELL students must circumvent the obstacle of passing several required exams to be considered for reclassification. If reclassification were based solely on the state's minimums of 556 on the CELDT and 300 on the CST, the reclassification rates would significantly increase for both districts. Further, either not taking or not passing the essay exam led administrators to not reclassify most of these students. Regardless of administrators' claims to be either subjective or objective, teachers rely heavily on the exam assessments, thus making the exams the gatekeepers to become RFEP. In reality, Manzanita teachers prevented students who passed the required exams from being reclassified. Meanwhile, Granada teachers made a few exceptions for some students who were able to meet the state's minimum requirements, however, here, the exceptions were unlikely to benefit male, Hispanic, and low-income students. Thus, though the districts may have good intentions in making fair policies, both continue to not reclassify historically disadvantaged students. Teachers' interactions and expectations of certain students may influence final reclassification decisions, and, here, they may view male, Hispanic, and low-income students negatively, or worse—incapable of succeeding in mainstream classrooms. These results reflect issues in our greater society, where male, Hispanic, and low-income individuals are underestimated, and so the districts reinforce and reproduce these inequalities.

Both districts relied heavily on assessments, and the tests were the gatekeepers that students had to pass in order to be considered for reclassification. Very few students who did not have all the required assessments or those students who did not pass one of the required assessments were reclassified. Many of the students without the required scores were special education students. Their individualized education program determined whether they needed to take exams; however, a number of non-special education students without the required assessments also existed. Both the CST and the essay exam were administered once every school year; if a student was absent, he or she had to wait an entire year to take that test.

More stringent language classification policies lead to middle school students who would otherwise be reclassified to remain ELL. In addition, teachers may be underestimating the capabilities of some language minorities based on gender, race, and SES, whereas these factors should not be part of the basis for establishing language competency. Differences persisted, even though the models accounted for special education, country born, assessment results, cohort, and middle school and teacher fixed effects. The results coincide with the work of previous researchers who demonstrated elementary school reclassification rates vary by gender and race, even after accounting for required assessments scores (Grissom, 2004; Halle, Hair, Wandner, McNamara, & Chien, 2012; Kohler & Lazarín, 2007). Teachers make exceptions for certain students who meet the lower state requirements, but these exceptions only benefit certain student subgroups. Datnow, Park, and Kennedy-Lewis (2012) found that teachers derive meaning from the data in an eclectic manner, sometimes drawing on their intuition and past interactions with individual students while simultaneously being influenced by policy and school content.

Other issues could have further influenced the districts' language classification policies and the teachers' final reclassification decisions. The current study looked at data during the No Child Left Behind (NCLB) years of 2002-2015. With NCLB, schools and districts had an accountability-driven incentive to keep their top performing English language learners classified as ELLs in order to have higher achievement scores meet Title I requirements (Christopher & de Alth, 2005; Slama, 2014). Each school year, districts had to report their ELLs' average achievement outcomes. Keeping higher skilled students classified as ELL increased the districts' averages. Neither of the districts in this study discussed this incentive during interviews. Rather, I found that Manzanita explicitly kept qualified students who passed their own higher thresholds as ELL, while Granada implicitly kept students classified ELL because they did not pass (or did not have) the required essay exam. Granada administrators also did not show great concern that the essay exam was created to determine non-ELL English course placement: initially, the essay exam was not intended to measure English proficiency. They were aware that most students failed the exam, but they were more concerned with having more ELL students meet the requirement than questioning the exam's validity. Furthermore, neither district discussed why a score of 325 on the CST was a better indicator of English proficiency than the state's minimum requirement of 300. Additionally, for Granada, there was no discussion as to how the essay exam measured English proficiency better than the CELDT and CST exams. These omissions showed that there was no real priority to discover whether these assessments truly measured proficiency. Rather, the administrators were more concerned with protecting students from losing language resources by being placed into mainstream English classes.

For decades, researchers have called on California to improve the ways in which those students who need language support are classified and reclassified, and have asked it to improve alignment across the state's districts (Abedi, 2008; Umansky et al., 2015). However, at the end of the NCLB era, district policies and their implementation continued to vary greatly between—and even within—districts. In addition, as discussed, gender, race, and SES continued to be significant indicators of a middle school student's probability of being reclassified. In the end, we find that beliefs about how best to measure English proficiency, the desire to meet state accountability requirements, and hoped-for financial gains to acquire more funding per pupil can strongly affect administrators' decisions regarding the creation of their language classification policies.

In fact, now, more than ever, it is critical to set a standard state language classification policy that districts must follow. Currently, immense federal and state changes are underway. The federal government has adopted the Every Student Succeeds Act (ESSA), which requires all schools demonstrate that they are improving the English language proficiency of their ELL students. ESSA is intended to strengthen the accountability provisions and to increase funding targeted at ELL students. The government will provide more money per ELL pupil and the districts will now have the freedom to allocate those resources as they choose, creating a greater financial incentive to keep students classified ELL. Locally, California has also implemented the Common Core State Standards (CCSS) and the English language development (ELD) standards, both of which are aimed at improving academic rigor in all subject areas and increasing the English language requirements for both ELL and non-ELL students. Furthermore, central language classification assessments, such as the CELDT and CST, are being replaced by the ELPAC and SBAC (Umansky et al., 2015). These changes can create additional barriers for ELL students to become RFEP.

Districts need strong guidance on best practices to determine English proficiency, particularly for middle school students, who are generally LTELL. Administrators from both districts showed concern over reclassifying students too early and the loss of resources if they reclassified. Both MUSD and GUSD have to make difficult language classification policies without much research or understanding about best practices to determine English proficiency. In 2013-14, MUSD discontinued the CST, and, in its place, administrators stated that they would use several other measures (e.g., HLAT/HULT, scholastic test, and students' grades) to determine student language classification status. During the data collection period of the present work, alternative tests were being considered by administrators. In the Granada school district in 2014-2015, CST scores were no longer available. The GUSD began to require ELL students to pass a benchmark assessment created by the district, in addition to the previously used CELDT and essay exams. In fact, for at least the past five years, Granada has required all non-ELL students to take the benchmark assessment for English course placement purposes. However, in

2014-2015, the benchmark assessment was updated to align with the CCSS. During my interviews, the GUSD administrators did not discuss how accurately the new benchmark measured English proficiency, or how it compared to the written essay. What is clear is that the state must develop a mechanism for checking to make sure that language minority students are appropriately classified and reclassified.

**Limitations**

In interpreting this study's findings, I note the following empirical limitations. First, the CELDT and CST have been shown to have limitations in determining English proficiency. In particular, the CST was not originally intended to determine English proficiency, and it has been normed based on non-ELL student performance. This means that many non-ELL students do not pass the CST exam. These tests are used for this study because these are the exams required by California. However, when the state determines the new state classification policy to determine English proficiency, it must strongly assess the established exams and cutoffs for each grade level. Language classification policies are being determined without any research-based evidence that the selected exams measure English proficiency, or if the set cutoffs are the best indicators that students have reached proficiency. However, assessing the best exams to determine English proficiency is beyond the scope of this paper.

A second limitation of this study is that the interviews took place during a period when administrators were preparing to adopt the new state standards and state exams. Many were eager to discuss the implications of the new policies, but were disinclined to talk about the old policies. Yet, in order to research the impacts of the policies, one must first have data to test the implications. The next step here will be to evaluate the districts' reclassification rates using the SBAC scores along with the other requirements, such as school grades and benchmark scores.

A third limitation is the absence of teacher and student interviews. Teachers appear to have played an essential role in making the final language classification decisions. Future research should include teacher interviews to understand their perception of their district's language classification policies. Teachers should also be asked to describe their rationale when making language classification decisions. Future researchers should also interview students. Middle school ELL students may not be informed as to how they can become reclassified. Complicated and changing reclassification criteria can make the pathway out of the ELL extremely difficult, thus placing an undue burden on students, particularly given the biased methods of entry into the system. Students may also be able to provide an explanation for not having the required CST and essay scores.

A final limitation is the lack of Individualized Education Program (IEP) information for special education students. IEP vary greatly and the specification for special education students can influence their probability of reclassification. Special education students are normally dropped from analysis (e.g., Hill, Weston, et al., 2014), but remain in my analyses due to the high representation of special education students in middle school. Many special education students are also long-term English language learners and the overlap between must be further researched, particularly in middle school, with emphasis on addressing the intersection of special education and ELL classification, especially when several of these students cannot RFEP because many are exempted from taking the required reclassification exams.

**Conclusion**

America's existing stratification systems distribute resources and rewards to its population differently. All too often, societal institutions, and especially the educational system, can enforce and reproduce gender, class, and racial inequalities. These processes can have far-reaching impacts on males, and low-income, and racial/ethnic minorities, and their opportunities to get ahead. Unintentionally, not being allowed to reclassify can lead to unintended negative consequences. In fact, prior research demonstrates the ELL classification itself can have implications on whether a student is able to access advanced English and math courses, which, in turn, can affect their educational outcomes (e.g., achievement, graduation, and attendance). In the following two studies, I will examine the effects of language classification on course placement, achievement, and behavioral outcomes.

References

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice, 27*(3), 17-31.

Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoro, S. (2005). The new demography of America's schools: Immigration and the No Child Left Behind Act. *Urban Institute (NJ1)*.

Christopher, J., & de Alth, S. (2005). *English Learners in California Schools*. Public Policy Institute of California.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Edwards, B., Leichty, J., & Wilson, K. (2008). English Learners in California: What the Numbers Say. *EdSource*.

Estrada, P., & Wang, H. (2013). Reclassifying and Not Reclassifying English Learners to Fluent English Proficient, Year 1 Findings: Factors Impeding and Facilitating Reclassification and Access to the Core. *Grantee Submission*.

Fusarelli, L. D. (2002). Tightly coupled policy in loosely coupled systems: Institutional capacity and organizational change. *Journal of Educational Administration, 40*(6), 561-575.

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources,Unequal outcomes. *Education Policy Analysis Archives, 11*(36), 1-54.

Grissom, J. B. (2004). Reclassification of English Learners. *Education Policy Analysis Archives, 12*(36), n36.

Hahnel, C., Wolf, L., Banks, A., & LaFors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. The Education Trust-West.

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.

Hampden-Thompson, G., Mulligan, G., Kinukawa, A., & Halle, T. (2008). Mathematics Achievement of Language-Minority Students During the Elementary Years.

Hanhnel, C., Wolf, L., Banks, A., & Lafors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. Retrieved from The Education Trust-West:

Hill, L. E., Betts, J. R., Chavez, B., Zau, A. C., & Bachofer, K. V. (2014). Pathways to Fluency: Examining the Link between Language Reclassication Policies and Student Success. *Public Policy Institute of California*.

Hill, L. E., Weston, M., & Hayes, J. M. (2014). Reclassification of English Learner Students in California. *Public Policy Institute of California. Retrieved from www. ppic. org/main/publication. asp*.

Kohler, A. D., & Lazarín, M. (2007). Hispanic education in the United States. *Statistical Brief, 8*.

Krashen, S., & Brown, C. L. (2005). The ameliorating effects of high socioeconomic status: A secondary analysis. *Bilingual Research Journal, 29*(1), 185-196.

Linquanti, 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. *Council of Chief State School Officers*.

Maxwell, L. (2014). US school enrollment hits majority-minority milestone. *Education Week*.

Meyer, J. W. (1983). Centralization of funding and control in educational governance. *Organizational environments: Ritual and rationality, 179*, 197.

Olsen, L. (2010). Reparable Harm Fulfilling the Unkept Promise of Educational Opportunity for California's Long Term English Learners. *California Together (Research Report)*.

Robinson-Cimpian, J. P., & Thompson, K. D. (2015). The Effects of Changing Test-Based Policies for Reclassifying English Learners. *Journal of Policy Analysis and Management*.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292.

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.

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.

Umansky, I. M., Reardon, S. F., Hakuta, K., Thompson, K. D., Estrada, P., Hayes, K., . . . Goldenberg, C. (2015). Improving the Opportunities and Outcomes of California's Students Learning English: Findings from School District-University Collaborative Partnerships. Policy Brief 15-1. *Policy Analysis for California Education, PACE*.

Walqui, A., Estrada, P., Koelsch, N., Hamburger, L., Gaarder, D., Insurralde, A., . . . Weiss, S. (2010). *What are We Doing to Middle School English Learners?: Findings and*

*Recommendations for Change from a Study of California El Programs (Research Report)*.

Weick, K. E. (1982). Administering education in loosely coupled schools. *The Phi Delta Kappan, 63*(10), 673-676.

Wolf, M. K., Kao, J., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S. M., & Farnsworth, T. (2008). Issues in Assessing English Language Learners: English Language Proficiency Measures and Accommodation Uses. Practice Review (Part 2 of 3). CRESST Report 732. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Table 1.1

*Full Sample, Selected Sample, and Complete Data*

| | Full Sample | | Selected Sample | | Complete Sample | |
|---|---|---|---|---|---|---|
| | **MUSD** | **GUSD** | **MUSD** | **GUSD** | **MUSD** | **GUSD** |
| *District information* | | | | | | |
| Total 8th grade enrollment | 16,144 | 7,582 | 4,231 (26%) | 2,905 (38%) | 2,969 (18%) | 2,054 (27%) |
| Average 8th grade cohort | 5,381 | 3,791 | 1,410 | 1,452 | 989 | 1,027 |
| Total # of middle schools | 13 | 10 | 8 (TS) | 10 | 8 (TS) | 10 |
| Cohort Years | 2010-2013 (3 cohorts) | 2012-2014 (2cohorts) | 2010-2013 (3cohorts) | 2012-2014 (2 cohorts) | 2010-2013 (3 cohorts) | 2012-2014 (2 cohorts) |
| *Student demographics in 8th grade* | | | | | | |
| % Female | 49.2 | 49.3 | 45.8 | 44.3 | 47.3 | 45.0 |
| % Hispanic or Latino | 65.8 | 52.8 | 87.6 | 70.5 | 88.3 | 72.8 |
| % Asian | 12.3 | 34.2 | 9.3 | 28.5 | 8.8 | 26.4 |
| % White | 13.3 | 10.1 | --- | --- | --- | --- |
| % African American | 3.1 | 0.9 | --- | --- | --- | --- |
| % Other race | 5.2 | 1.8 | 2.9 | 0.9 | 2.8 | 0.6 |
| % Born in United States | 84.0 | 84.4 | 74.1 | 74.8 | 75.3 | 78.1 |
| % Free- and Reduced-Price Lunch | 71.7 | 72.3 | 90.1 | 88.1 | 90.2 | 88.7 |
| % Special Education | 10.1 | 9.8 | 19.7 | 15.4 | 4.6 | 4.2 |
| *Dependent Variable* | 8th | 8th | 8th | 8th | 8th | 8th |
| % English language learners (ELL) | 22 | 29 | 72 | 74 | 65 | 66 |
| % Reclassified Fluent English Speakers (RFEP) | 40 | 45 | 28 | 26 | 35 | 34 |
| % English Only (EO) and Initially Fluent English Speakers (IFEP) | 38 | 26 | --- | --- | --- | --- |

*Note*. The full sample represents averages over several school years for middle school students provided by school districts. The selected sample represents 6th graders classified as English language learners (26% in Manzanita and 38% in Granada). Of these students 28% at Manzanita and 26% at Granada become RFEP by the 8th grade. The sample excludes White, African American, EO, IFEP, and elementary RFEP, and also excludes non-traditional schools (TS). The "other race" category includes American Indian, Alaskan Native, Filipino, Native Hawaiian, and Pacific Islanders. The complete data includes students who have CELDT and CST scores, which are the two main California reclassification requirements.

Table 1.2

*Annual Reclassification Procedures and Exams Used*

| | Summer | | Fall | | | | Spring | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | July | Aug | Sept | Oct | Nov | Dec | Jan | Feb | March | April | May | June |
| Exams are taken in: | | | | | | | | | | | | |
| CELDT | | X | | | | | | | | | | |
| CST ELA | | | | | | | | | | | X | |
| Essays | | | | | | | | | | | | X |
| Reclassification occurs in: | | | | | | | | | | | | |
| Manzanita | | | | | | | **X** | | | | | |
| Granada | | | **X** | | | | **X** | | | | | |
| Scores used for reclassification: | | | | | | | | | | | | |
| CELDT | | | Prior (Fall) | | | | Current (Fall) | | | | | |
| CST ELA | | | Prior (Spring) | | | | Prior (Spring) | | | | | |
| Essay | | | Prior (Spring) | | | | Prior (Spring) | | | | | |

*Note.* In both districts students are reclassified in the spring but in Granada students can also be reclassified at the beginning of the school year. In the spring, students are reclassified based on the current school year's CELDT and last school year's CST ELA and essay scores. In the fall, Granada students are reclassified based on last school year's CELDT, CST ELA, and essay sores. For example, for a GUSD 8$^{th}$ grader reclassified in the fall, administrators would use her 7th grade CELDT, CST, and essay scores. In comparison, for an 8$^{th}$ grader reclassified in the spring, administrators would use her 8th grade CELDT and 7$^{th}$ grade CST and essay scores. Most students take the CELDT between July and October but this test can be taken anytime during the school year. As for CST and written essays, they are only given to middle school students once per school year and students cannot take the test another day.

Table 1.3

*California and District Reclassification Requirements*

| Assessments and Other Reclassification Criteria (*Scoring Scale)* | CA Minimum Requirements | Manzanita | Granada | % of CA Districts (Hill et al., 2014) |
|---|---|---|---|---|
| CELDT Overall *(1-5)* | 4 | 4 | 4 | 91% (4)<br>7% (5) |
| CELDT four subtests *(1-5)* | 3 | Listening/ Speaking (3) Reading/ Writing (4) | Listening, Speaking, Reading, Writing (3) | 65% (3)<br>35% (4) |
| CST ELA *(1-5)* | 3 | mid-Basic (3.5) | mid-Basic (3.5) | 27% (3)<br>45% (3.5)<br>27% (4) |
| Essays *(1-4)* | --- | --- | 3 | 9% |
| Teacher Recommendation | Unspecified Criteria | Yes, criteria unspecified | Yes, mainly on assessments | 22% Unspecified<br>78% Specified Criteria |
| Parent Recommendation | Unspecified Criteria | Parents' opinion considered | Parents are informed | 85% |

*Note*. California sets minimum requirements that English language learners must pass in order to become RFEP, including passing the CELDT and CST ELA, and consulting parents and teachers. The CELDT and CST scores can be presented as scale and raw scores. I provided scale scores because raw scores and passing cutoffs vary by grade and subtests. For example, overall CELDT score ranges from 248-741 but passing is 556 for 7$^{th}$ graders and 569 for 8$^{th}$ graders (both equivalent to 4 on the scale). In Manzanita, parents are allowed to request their child's language classification be changed regardless of test scores, but in Granada parents are only informed about their child's language classification but they are not allowed to determine their child's classification. Furthermore, Manzanita's reclassification includes English teachers' recommendations but evaluation should be based on students' GPA where they must score higher than a 2.0+. Documentation also states English teachers can make exceptions if they believe low grades were not based on students' English proficiency. Hill et al. (2014) shows that most California districts, similar to my two districts, tend to add more reclassification requirements than the state requires.

Table 1.4

*Chi-Squared Test: Reclassification Rates by Gender, Race, SES, and Special Education*

| Manzanita | 7th Grade (n=4,231) | 8th Grade (n=3,430) |
|---|---|---|
| | RFEP % | RFEP % |
| Male | 16*** | 11* |
| Female | 22*** | 13* |
| Hispanic | 17*** | 10*** |
| Asian | 31*** | 30*** |
| Other | 32*** | 23*** |
| FRL | 18*** | 11*** |
| non-FRL | 31*** | 21*** |
| Special Edu. | 01*** | 02*** |
| Non-Special Edu. | 23*** | 15*** |
| Total | 19 | 12 |
| **Granada** | **7th Grade (n=2,905)** | **8th Grade (n=2,455)** |
| | RFEP% | RFEP% |
| Male | 14*** | 12+ |
| Female | 18*** | 14+ |
| Hispanic | 12*** | 12 |
| Asian | 23*** | 14 |
| Other | 26*** | 10 |
| FRL | 15* | 12* |
| non-FRL | 20 | 16* |
| Special Edu. | 03*** | 02*** |
| Non-Special Edu. | 18*** | 15*** |
| Total | 16 | 13 |

*Note*. For each district, I ran four separate chi-square tests to determine if reclassification rates differed by student's demographics + $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$. In both districts, males, Hispanic, FRL, and special education students are less likely to be reclassified. The 8th grade data only includes students who did not reclassify in 7th grade.

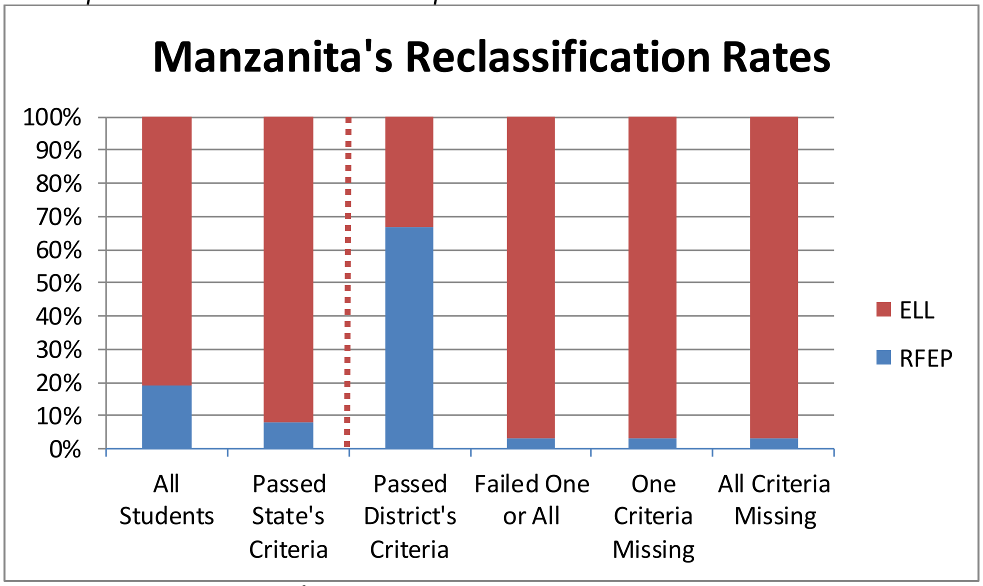
Table 1.5 *Missing Reclassification Criteria by Language Classification*

| | **Manzanita's Language Classification Criteria** | | | |
|---|---|---|---|---|
| | All criteria available % | One reclassification criteria missing % | All reclassification criteria missing % | CELDT and CST available % |
| ELLs 7th (CELDT/CST)<br>**N=3,430**<br>No SE (N=2,605)<br>No SE/F (N=1,862) | **64.2**<br>(79.6)<br>(84.0) | **28.2**<br>(13.0)<br>(10.3) | **7.4**<br>(7.2)<br>(5.5) | **64.2**<br>(79.6)<br>(84.0) |
| RFEPs 7th (CELDT/CST)<br>**N=801**<br>No SE (N=789)<br>No SE/F (N=576) | **95.3**<br>(95.3)<br>(96.1) | **3.5**<br>(3.5)<br>(2.4) | **1.1**<br>(1.1)<br>(1.3) | **95.3**<br>(95.3)<br>(96.1) |
| ELLs 8th (CELDT/CST/GPA)<br>**N=3,032**<br>No SE (N=2,224)<br>No SE/F (N=1,626) | **81.0**<br>(94.5)<br>(94.8) | **18.5**<br>(4.9)<br>(4.7) | **0.4**<br>(0.5)<br>(0.4) | **81.3**<br>(94.9)<br>(95.2) |
| RFEPs 8th (CELDT/CST/GPA)<br>**N=398**<br>No SE (N=381)<br>No SE/F (N=236) | **95.4**<br>(96.0)<br>(97.8) | **4.2**<br>(3.6)<br>(1.6) | **0.2**<br>(0.2)<br>(0.4) | **95.7**<br>(96.3)<br>(97.8) |
| | **Granada's Language Classification Criteria** | | | |
| | All criteria available % | One reclassification criteria missing % | All reclassification criteria missing % | CELDT and CST available % |
| ELL 7th (CELDT/CST/Essays)<br>**N=2,455**<br>No SE (N=2,021)<br>No SE/F (N=1,450) | **61.8**<br>(71.3)<br>(79.5) | **31.0**<br>(21.2)<br>(15.9) | **7.1**<br>(7.3)<br>(4.4) | **66.8**<br>(77.0)<br>(83.1) |
| RFEP 7th (CELDT/CST/Essays)<br>**N=450**<br>No SE (N=439)<br>No SE/F (N=344) | **91.5**<br>(92.9)<br>(93.9) | **5.7**<br>(4.5)<br>(4.6) | **2.6**<br>(2.5)<br>(1.4) | **92.0**<br>(93.3)<br>(94.4) |
| ELL 8th (CELDT/CST/Essays)<br>**N=2,145**<br>No SE (N=1,721)<br>No SE/F (N=1,206) | **65.5**<br>(77.4)<br>(83.3) | **33.7**<br>(22.2)<br>(16.3) | **0.79**<br>(0.2)<br>(0.3) | **68.8**<br>(81.1)<br>(85.1) |
| RFEP 8th (CELDT/CST/Essays)<br>**N=310**<br>No SE (N=300)<br>No SE/F (N=244) | **97.4**<br>(99.0)<br>(99.5) | **2.2**<br>(0.6)<br>(0.4) | **0.32**<br>(0.91)<br>(0.29) | **97.4**<br>(99.0)<br>(99.5) |

*Note*. The bold percentages represent all students in the selected sample. The percentages in the first parentheses exclude special education (SE) students. The percentages in the second parentheses exclude special education (SE) and foreign-born (F) students. Regardless of these two restrictions ELL students are still less likely than RFEP to not have all reclassification requirements. Columns 2-4 include all the districts' reclassification requirements. Manzanita requires CELDT and CST ELA scores, and for 8th graders GPA is also considered. Granada requires CELDT, CST ELA, and written essays. Students that have an overall CELDT score also have the four CELDT subtests due to the fact that the overall score is created from the subtests. Column 5 represents the percentage of students who have CELDT and CST scores. In Manzanita, column 5 is similar to column 2 but for Granada a few more students are missing essay scores who have CELDT and CST scores.

Figure 1.1

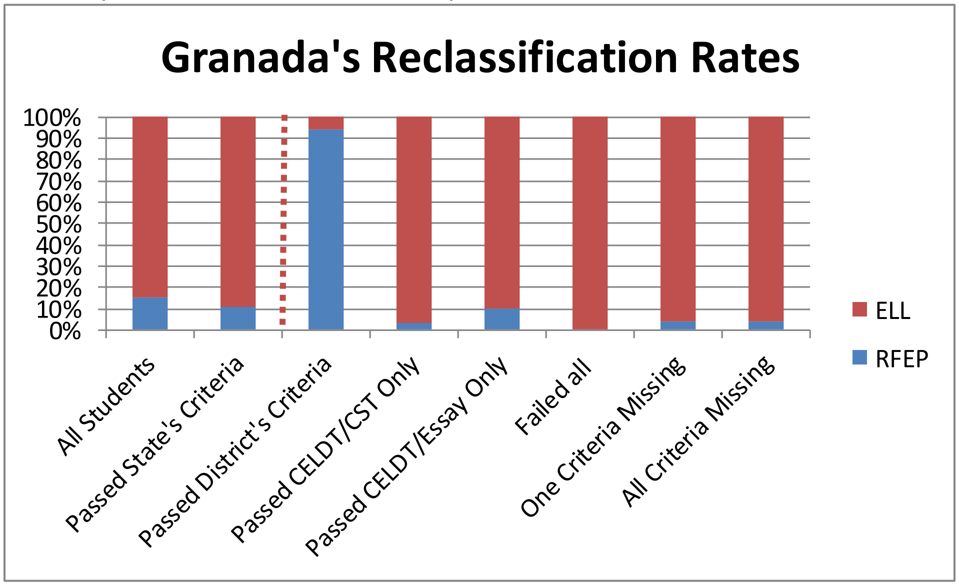
*Reclassification Criteria and Reclassification Rates in Manzanita District*




*Note*. This table represents 7th graders' (n=4,231) reclassification rates. Overall, about 19% of Manzanita's 7th graders become reclassified. About 8% of the students who pass the state's requirements but not the district's requirements become reclassified. From the students who pass the district's requirements about 70% become reclassified. Nearly no student becomes RFEP if they fail the CELDT or the CST, if they do not have CELDT or CST score or if they are missing both scores. At minimum students must pass the CELDT and the CST in order to be considered for reclassification.

Figure 1.2

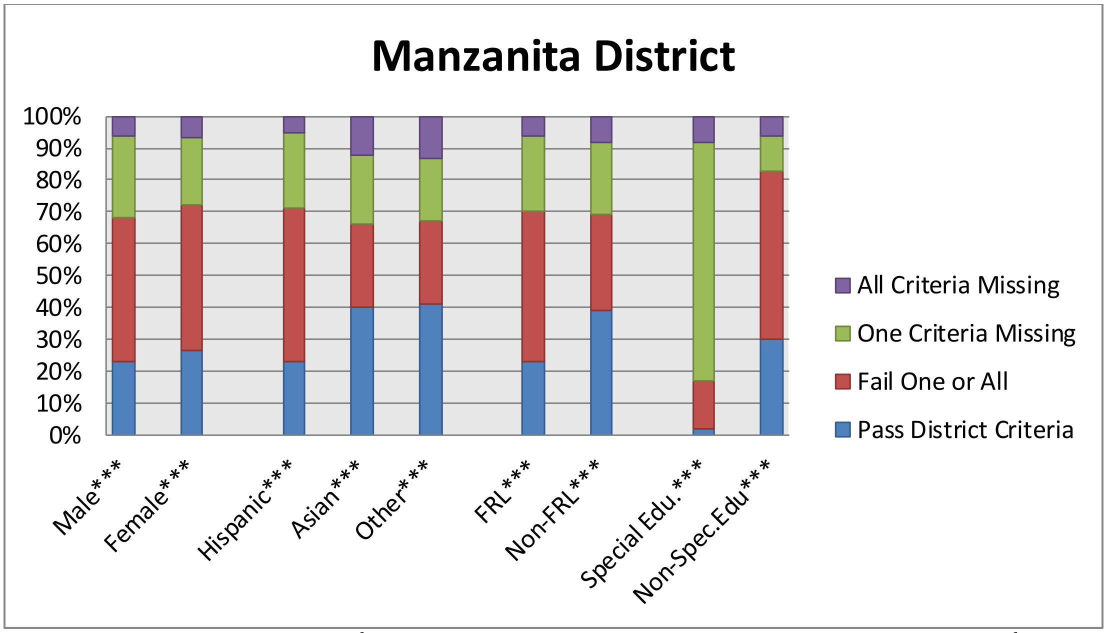
*Reclassification Criteria and Reclassification Rates in Granada District*




*Note*. This table represents 7th graders' (n=2,905) reclassification rates. In Granada, about 15% of 7th graders become reclassified. About 11% of the students who pass the state's requirements but not the district's requirements become reclassified. From the students who pass the district's requirements about 94% become reclassified. Nearly no student becomes RFEP if they only pass the CELDT and CST but not the essay or if they fail all of the requirements. However, about 10% of students who pass the CELDT and essay but fail the CST are reclassified. Less than 4% of students without one or any of the requirements are reclassified.

Figure 1.3

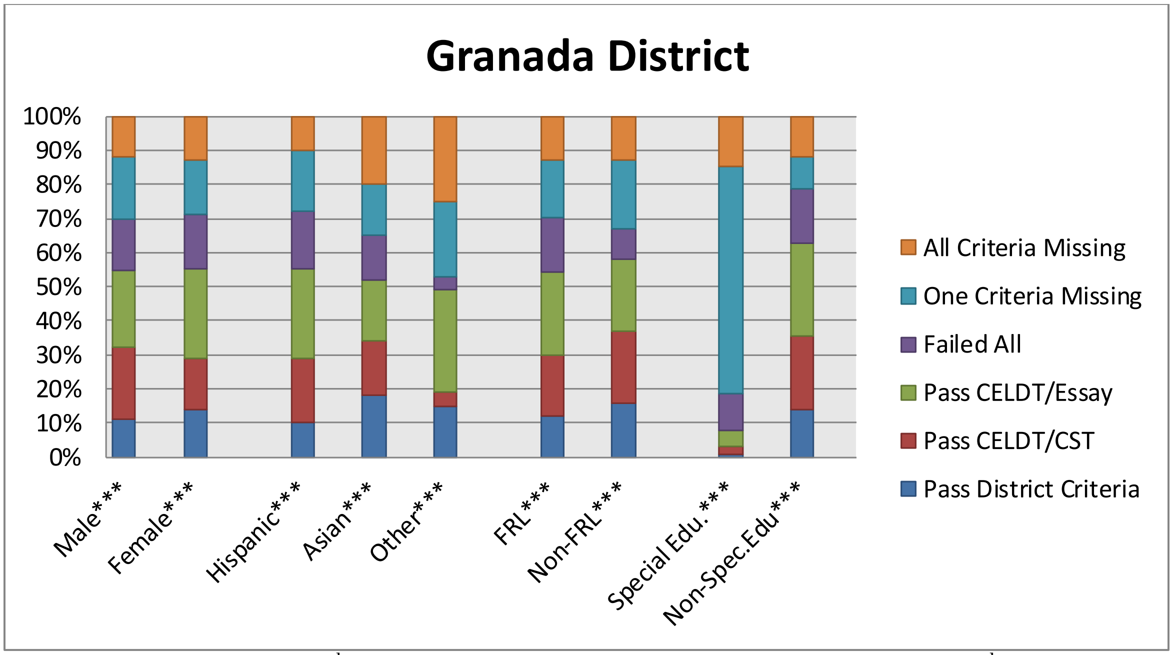
*Reclassification Criteria and Students' Demographics for Middle School ELL Students in Manzanita Unified School District (Chi-Square Test)*




*Note*. This table represents 7th graders (n=4,231) who are classified ELL as of 6th grade. Students either pass the district requirements at 556 on the CELDT and 325 on the CST (1/blue), fail the CELDT and/or the CST (2/red), do not have CST scores (3/green), or do not have any scores (4/purple). I ran four separate chi-square tests to determine if having and passing the reclassification criteria differed by student's demographics $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Male, Hispanic, FRL and special education students are less likely to pass both requirements (blue lines) and these differences are statistically significant. Male, Hispanic, "other" race, and special education are more likely to not have the required scores (green and purple lines). These results are similar for 8th graders.

Figure 1.4

*Reclassification Criteria and Students' Demographics for Middle School ELL Students in Granada Unified School District(Chi-Square Test)*




*Note*. This table represents 7th graders (n=2,905) who are classified ELL as of 6th grade. Students either pass the CEDLT, CST, and essay-the district's requirements (1/blue), pass the CELDT and CST but not the essay (2/red), pass the CELDT and essay but not the CST (3/green), fail all the requirements (4/purple), do not have a score for one of the scores (5/teal), do not have any of the scores (6/orange). I ran four separate chi-square tests to determine if having and passing the reclassification criteria differed by student's demographics $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Male, Hispanic, FRL and special education students are less likely to pass all three requirements (blue lines) and these differences are statistically significant. Hispanic, "other" race, and special education are more likely to not have all the required scores (teal and orange lines). These results are similar for 8th graders except for the gender difference was not statistically different.

Table 1.6

*Odd Ratios of Reclassifying for MUSD Students who passed the District's and State's Minimum Requirements*

| | 7th Graders | | 8th Graders | |
|---|---|---|---|---|
| | (1) District | (2) State | (3) District | (4) State |
| Female | **1.78*****[***] | **1.68*****[***] | 1.49 | 1.08 |
| | **(0.26)** | **(0.23)** | (0.31) | (0.21) |
| Hispanic | **0.49*****[***] | **0.55****[**] | 1.09 | 1.12 |
| | **(0.10)** | **(0.11)** | (0.26) | (0.23) |
| FRL | **0.58****[**] | **0.60****[**] | 0.77 | 0.82 |
| | **(0.11)** | **(0.10)** | (0.25) | (0.23) |
| Special Educ. | 0.67 | 0.81 | 0.57 | 0.51 |
| | (0.39) | (0.44) | (0.25) | (0.18) |
| Born US | 1.10 | 1.13 | 0.70* | 0.72* |
| | (0.14) | (0.16) | (0.12) | (0.11) |
| CST 325† | | 34.10*** | | 26.86*** |
| | | (7.50) | | (6.37) |
| GPA | | | 3.14*** | 2.65*** |
| | | | (0.91) | (0.71) |
| School 1(Ref.) | | | | |
| School 2 | 0.20** | 0.20** | 0.04*** | 0.04*** |
| | (0.11) | (0.11) | (0.02) | (0.02) |
| School 3 | 0.44** | 0.45** | 0.58 | 0.52 |
| | (0.13) | (0.14) | (0.25) | (0.20) |
| School 4 | 0.25** | 0.30** | 0.41 | 0.42 |
| | (0.11) | (0.13) | (0.20) | (0.21) |
| School 5 | 2.81* | 3.18* | 0.71 | 0.80 |
| | (1.43) | (1.58) | (0.35) | (0.45) |
| School 6 | 3.30** | 2.58* | 0.79 | 0.58 |
| | (1.49) | (1.10) | (0.34) | (0.21) |
| School 7 | 0.51 | 0.49* | 0.43 | 0.36* |
| | (0.18) | (0.16) | (0.23) | (0.15) |
| School 8 | 0.49* | 0.59 | 0.26 | 0.35 |
| | (0.15) | (0.22) | (0.22) | (0.33) |
| *N* | 1043 | 1612 | 556 | 1119 |
| df_m | 16.00 | 17.00 | 16.00 | 17.00 |
| chi2 | 146.59 | 531.18 | 115.78 | 348.91 |
| pr2 | .14 | .37 | .19 | .35 |

*Note.* All models include cohort and teachers fixed effects. The "other" race category only included a few students; therefore, it was combined with Asian American, and this group was referred as non-Hispanic. Models 1 and 3 include students who passed the district's minimum requirements of 556 on the CELDT and 325 on the CST. MUSD made some exceptions and reclassified some students who met the states but not the district's minimum requirements. Therefore, Models 2 and 4 include students who passed the state's minimum requirements of 556 on the CELDT and 300 on the CST. †CST 325 represents whether or not the student scores 325 or higher on the CST. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

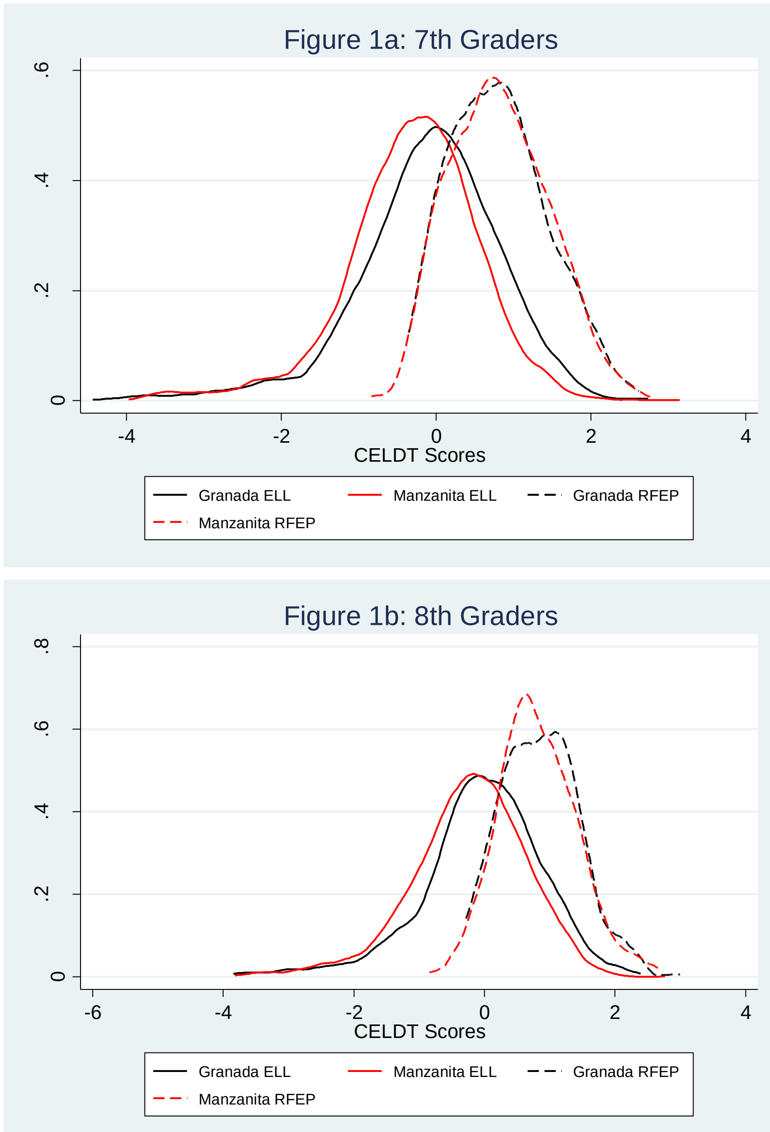
Table 1.7 *Odd Ratios of Reclassifying for GUSD Students who passed the District's and State's Minimum Requirements*

| | 7th Graders | | 8th Graders | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | District | State | District | State |
| Female | 0.57 | **1.79*** | 0.34 | **1.43** |
| | (0.20) | **(0.27)** | (0.23) | **(0.18)** |
| Hispanic | 1.21 | **0.44*** | 1.43 | **0.63*** |
| | (0.65) | **(0.08)** | (1.19) | **(0.12)** |
| FRL | **0.48*** | 0.97 | 2.02 | **1.50*** |
| | **(0.14)** | (0.18) | (2.42) | **(0.29)** |
| Special Educ. | 0.00*** | 0.45 | 0.00*** | 1.03 |
| | (0.00) | (0.39) | (0.00) | (0.65) |
| Born US | 0.53 | 1.17 | 0.52 | 0.86 |
| | (0.23) | (0.25) | (0.40) | (0.19) |
| CST 325† | | 4.07*** | | 568.21*** |
| | | (0.84) | | (725.00) |
| Passed Essay (Ref.) | | | | |
| Failed Essay | | 0.00*** | | 0.03*** |
| | | (0.00) | | (0.02) |
| No Essay | | 0.00*** | | 0.00*** |
| | | (0.00) | | (0.00) |
| School 1(Ref.) | | | | |
| School 2 | 0.00*** | 0.54** | 2.40 | 1.55 |
| | (0.00) | (0.11) | (3.86) | (1.75) |
| School 3 | 0.71 | 0.15*** | 0.59 | 0.57 |
| | (0.63) | (0.04) | (0.62) | (0.25) |
| School 4 | 0.95 | 0.63*** | 1.16 | 1.54 |
| | (0.80) | (0.08) | (1.10) | (0.70) |
| School 5 | 1.01 | 0.34*** | 0.31 | 1.10 |
| | (0.85) | (0.06) | (0.43) | (0.87) |
| School 6 | 0.33 | 0.33* | 0.00*** | 0.45 |
| | (0.33) | (0.18) | (0.00) | (0.22) |
| School 7 | 0.00*** | 0.31*** | 0.00*** | 0.58 |
| | (0.00) | (0.10) | (0.00) | (0.26) |
| School 8 | 0.23 | 0.41 | 0.00*** | 0.04*** |
| | (0.26) | (0.20) | (0.00) | (0.03) |
| School 9 | 0.00*** | 0.52* | 0.60 | 1.23 |
| | (0.00) | (0.13) | (0.92) | (0.48) |
| School 10 | 0.65 | 0.39*** | 0.00*** | 0.53 |
| | (0.65) | (0.08) | (0.00) | (0.18) |
| *N* | 347 | 1263 | 308 | 1125 |
| df_m | 16.00 | 19.00 | 16.00 | 20.00 |
| chi2 | 3767.22 | ---- | 2276.45 | 2616.21 |
| pr2 | .10 | .15 | .15 | .34 |

*Note*. All models include cohort and teachers fixed effects. The "other" race category only included a few students; therefore, it was combined with Asian American, and this group was referred as non-Hispanic. Models 1 and 3 include students who passed the district's minimum requirements of 556 on the CELDT, 325 on the CST, and 3 on the essay. GUSD reclassified some students who met the states but not the district's minimum requirements. Thus, Models 2 and 4 include students who passed the state's minimum requirements of 556 on the CELDT and 300 on the CST. †CST 325 represents if a student met the CST 325 cutoff. Students without an essay score, or who failed the essay, had a near zero probability of being reclassified. $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Figure 1.5

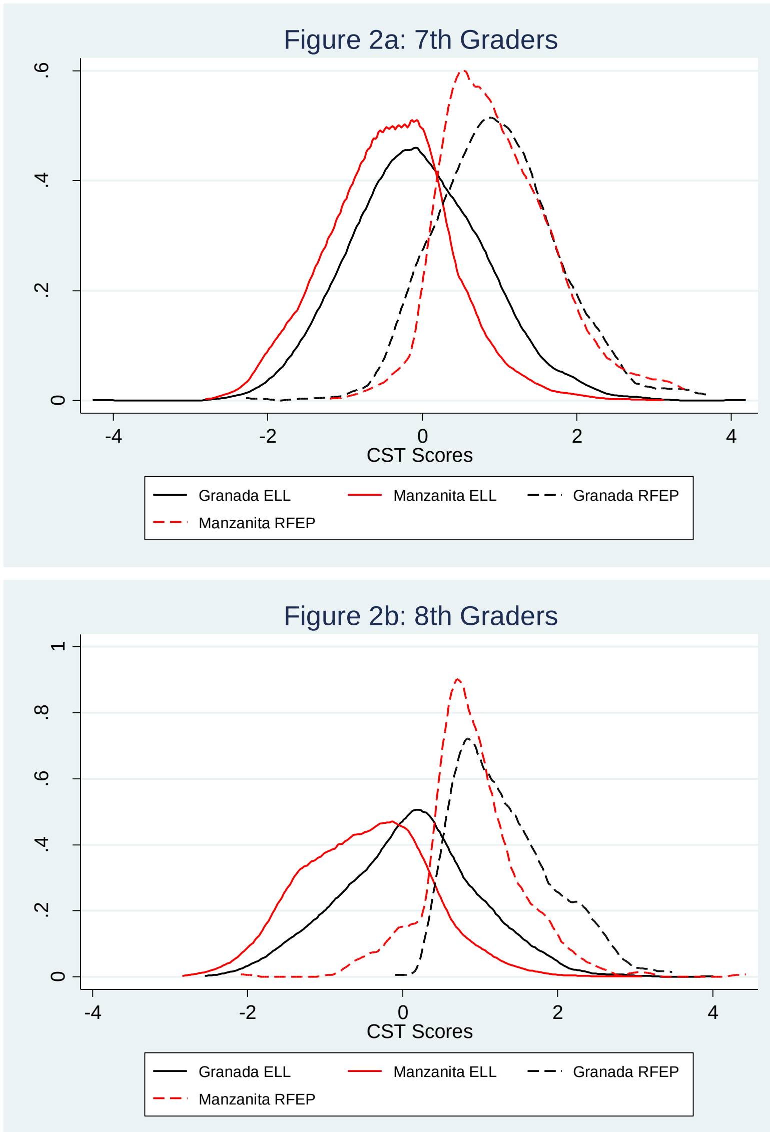
*CELDT Distribution Scores for ELL and RFEP Students*




*Note*. Seventh graders are reclassified based on their 7th grade CELDT scores. Eight graders are reclassified based on their 8th grade CELDT scores. Granada ELL students have higher CELDT averages than Manzanita students but RFEP students' scores are comparable. In 7th grade, about 1% (n=35) of Manzanita ELL students and Granada ELL students (n=32) score the lowest CELDT score of 248. In 8th grade, about 1% (n=15) of Manzanita ELL students and Granada ELL students (n=35) students score the lowest CELDT score of 248.

Figure 1.6

*CST ELA Distribution Scores for ELL and RFEP Students*




*Note*. Seventh graders are reclassified based on their 6th grade CST ELA scores. Eight graders are reclassified based on their 7th grade CST ELA scores. Granada ELL and RFEP students have higher CST averages than their counterparts in Manzanita district.

Table 1.8

*Reclassification Rates by Different Reclassification Policies*

| | % Reclassified Fluent English Proficient (RFEP) | | | |
|---|---|---|---|---|
| | Manzanita | | Granada | |
| | 7th Grade | 8th Grade | 7th Grade | 8th Grade |
| Actual Rates | 26% | 14% | 21% | 17% |
| CELDT 556 & CST 325 | 35% | 20% | 43% | 41% |
| CELDT 556 & CST 300 | 55% | 40% | 63% | 65% |

*Note*. Both Manzanita and Granada school districts reclassify fewer students than those who pass the minimum district's CELDT 556 and 325 requirements. Manzanita uses their own discretion on who should remain reclassified although they pass the CELDT and CST. In Granada, many students that pass the CELDT and CST do not pass or do not take the essay exam. The second set of percentages represents the number of students who would be reclassified if reclassification was solely based on the CELDT 556 and CST 325 exams. The third set of percentages represents the number of students who would be reclassified if reclassification was based on the state's minimum requirements of 556 and 300 on the CELDT and CST.

## Appendix

Appendix Table 1.1A

### Interview Questionnaire

**School and district administrators that work with ELL students particularly those that make language classification decisions.**

Code # of Interviewee:____________________ Date: ________________________________

1) Please describe your job responsibilities.
    a. Probe: What role do you play with ELL students?

**Study 1**

2) Please describe the language classification process at your district particularly for middle school students.
    a. Probe: Is the language classification processes decided at the district level? Can the process differ between schools? If so, what are those differences?
    b. Probe: Which of the following components are considered and to what extent: CELDT, ELA CST, ELA course grade, teacher recommendation, and parent recommendation.
    c. Probe: Will the Smarter Balanced Assessment Consortium (SBAC) be considered similarly to the CST when it comes to language classification? How? (Please provide details).
3) (If applicable) Based on district data it seems that Hispanic and low-income students are overrepresented in the ELL category? Can you describe why you think this may be occurring?

**Study 2**

4) What criteria are used to place ELL and RFEP middle school students in English courses?
    a. Probe: Is English proficiency considered when placing students into regular or advanced English courses? If so, how is it considered? In your experience, what are some reasons why English proficiency is considered?
    b. Probe: (If applicable) Based on district data it seems ELL students are not placed in regular or Honors English courses. Can you describe why you think this may be occurring?
5) What criteria are used to place ELL and RFEP middle school students in math courses?
    a. Probe: Is English proficiency considered when placing students into regular or advanced math courses? If so, how is it considered? In your experience, what are some reasons why English proficiency is considered?
    b. Probe: (If applicable) Based on district data it seems ELL students are not placed in advanced math courses. Can you describe why you think this may be occurring?
6) What types of support services do ELL students receive in middle school for language development and academic achievement?
    **a.** Probe: For example, are there any services such as 1-1 tutoring, extra ELD course after school reading program, certified ELL teacher/tutor, etc.?

Appendix Table 1.2A

*Missing Reclassification Criteria in Manzanita District (MLOGIT/ODD RATIOS)*

| | (1) 7th Graders | (2) 8th Graders | (3) 8th Graders | (1) 7th Graders | (2) 8th Graders | (3) 8th Graders |
|---|---|---|---|---|---|---|
| **All Criteria Available** | CELDT/CST | CELDT/CST/ GPA | CELDT/CST | CELDT/CST | CELDT/CST/ GPA | CELDT/CST |
| | **One Criterion is Missing** | | | **All Criteria is Missing** | | |
| Female | 1.20* | 0.96 | 0.93 | 1.29* | 0.97 | 1.47 |
| | (0.10) | (0.12) | (0.11) | (0.17) | (0.54) | (0.57) |
| Hispanic (Ref.) | | | | | | |
| Asian | 1.43* | 1.65 | 1.87* | 2.45*** | 0.35 | 0.57 |
| | (0.26) | (0.53) | (0.60) | (0.48) | (0.24) | (0.44) |
| Other | 1.20 | 0.93 | 1.00 | 2.37*** | 1.07 | 1.37 |
| | (0.22) | (0.42) | (0.44) | (0.56) | (0.81) | (0.84) |
| Born in the U.S. | 0.52*** | 0.93 | 0.87 | 0.59** | 0.67 | 0.88 |
| | (0.04) | (0.14) | (0.15) | (0.10) | (0.42) | (0.30) |
| FRL | 0.88 | 0.87 | 0.97 | 1.06 | 0.16*** | 0.24* |
| | (0.20) | (0.35) | (0.35) | (0.31) | (0.07) | (0.16) |
| Special Educ. | 41.59*** | 29.40*** | 31.89*** | 8.37*** | 1.50 | 22.38*** |
| | (9.33) | (5.70) | (6.00) | (2.64) | (0.66) | (9.13) |
| Cohort 1 (Ref.) | | | | | | |
| Cohort 2 | 1.48 | 1.60 | 1.59 | 0.52*** | 0.55 | 0.30* |
| | (0.31) | (0.45) | (0.45) | (0.10) | (0.24) | (0.15) |
| Cohort 3 | 1.74** | 4.47*** | 4.62*** | 0.18*** | 0.29 | 2.85** |
| | (0.29) | (1.46) | (1.45) | (0.05) | (0.21) | (1.15) |
| | | | *N* | 4231 | 3436 | 3430 |
| | | | df_m | 6.00 | 6.00 | 6.00 |
| | | | pr2 | .25 | .33 | .33 |

*Note*. The three separate models include school fixed effects. In Model 1, the outcome are three categories for 7th graders: 1) all criteria available (reference group) means they have CELDT/CST scores, 2) one criterion is missing means they do not have the CELDT or CST scores, and 3) all criteria is missing means they do not have both the CELDT and CST scores. In Model 2, the outcome are three categories for 8th graders: 1) all criteria available (reference group) means they have CELDT/CST scores and GPA, 2) one criterion is missing means they do not have the CELDT or CST scores or GPA, and 3) all criteria is missing means they do not have both the CELDT and CST scores. In Model 3, the outcome are three categories for 8th graders: 1) all criteria available (reference group) means they have CELDT/CST scores, 2) one criterion is missing means they do not have the CELDT or CST scores, and 3) all criteria is missing means they do not have both the CELDT and CST scores. Models 1 and 3 demonstrate that Female, Asian American, and special education students are more likely to be missing one or all reclassification criteria. In the 7th grade, students born in the United States are less likely to be missing one reclassification criteria. In the 8th grade, FRL are more likely to be missing all reclassification criteria.

Appendix Table 1.3A

*Missing Reclassification Criteria in Granada District (MLOGIT/ODD RATIOS)*

| | (1) 7th | (2) 7th | (3) 8th | (4) 8th | (1) 7th | (2) 7th | (3) 8th | (4) 8th |
|---|---|---|---|---|---|---|---|---|
| **All Criteria Available** | CELDT/ CST/ Essay | CELDT/ CST | CELDT/ CST/ Essay | CELDT/ CST | CELDT/ CST/ Essay | CELDT/ CST | CELDT/ CST/ Essay | CELDT/ CST |
| | **One Criterion is Missing** | | | | **All Criteria are Missing** | | | |
| Female | 1.14 | 1.07 | 1.14 | 1.10 | 1.03 | 1.13 | 0.56 | 0.65 |
| | (0.11) | (0.10) | (0.11) | (0.14) | (0.25) | (0.25) | (0.29) | (0.27) |
| Hispanic (Ref.) | | | | | | | | |
| Asian | 1.59* | 1.43 | 1.55** | 1.30 | 2.04** | 2.13*** | 3.29* | 3.09*** |
| | (0.34) | (0.32) | (0.21) | (0.21) | (0.47) | (0.48) | (1.71) | (1.01) |
| Other | 1.26 | 1.14 | 3.24 | 3.95 | 3.02 | 3.53* | 11.52 | 9.23 |
| | (1.02) | (0.96) | (2.39) | (2.87) | (2.32) | (2.09) | (16.79) | (12.97) |
| Born in the U.S. | 0.36*** | 0.40*** | 0.41*** | 0.45*** | 0.27*** | 0.40*** | 1.46 | 3.61 |
| | (0.05) | (0.04) | (0.03) | (0.04) | (0.03) | (0.04) | (1.33) | (2.50) |
| FRL | 1.08 | 0.98 | 0.91 | 0.90 | 0.89 | 0.74 | 0.36* | 0.52 |
| | (0.24) | (0.26) | (0.22) | (0.29) | (0.29) | (0.24) | (0.18) | (0.25) |
| Special Educ. | 23.39*** | 31.86*** | 25.55*** | 26.74*** | 6.03** | 5.52*** | 44.77*** | 15.12*** |
| | (5.54) | (7.21) | (6.51) | (6.23) | (3.32) | (2.70) | (34.27) | (8.24) |
| Cohort | 1.18 | 1.22 | 1.05 | 0.93 | 1.35*** | 1.28** | 4.48 | 1.92** |
| | (0.19) | (0.18) | (0.16) | (0.15) | (0.12) | (0.11) | (3.98) | (0.48) |
| *N* | | | | | 2905 | 2905 | 2455 | 2455 |
| df_m | | | | | 7.00 | 7.00 | 7.00 | 7.00 |
| pr2 | | | | | .17 | .19 | .23 | .23 |

*Note*. The four separate models include school fixed effects. Models 1 and 2 represents 7th graders, and Models 3 and 4 represents 8th graders. In Models 1 and 3, all criteria available means CELDT, CST, and essay scores; in Models 2 and 4, all criteria available means CELDT and CST. Asian American, foreign-born, and special education students are more likely to have missing one reclassification or all reclassification criteria in both grade levels.

Appendix Table 1.4A

*Average Reclassification Scores for Students with both CELDT & CST ELA scores*

| **Manzanita** | 7th Graders | | 8th Graders | |
|---|---|---|---|---|
| | ELL | RFEP | ELL | RFEP |
| N (%) | 2,205 (74%) | 764 (25%) | 2,456 (87%) | 380 (13%) |
| CELDT $\bar{X}$ (SD) **(248-741) 556 pass** | 556.9 (44.6) | 613.5 (31.2) | 570.0 (46.8) | 624.1 (28.6) |
| CST ELA $\bar{X}$ (SD) **(150-600) 325 pass** | 299.7 (35.5) | 354.5 (27.3) | 284.4 (38.0) | 342.2 (26.0) |
| GPA $\bar{X}$ (SD) **(0-4.0) 2.0 pass** | --- | --- | 2.1 (0.81) | 2.7 (0.84) |
| **Granada** | 7th Graders | | 8th Graders | |
| | ELL | RFEP | ELL | RFEP |
| N (%) | 1,640 (80%) | 414 (20%) | 1,477 (83%) | 302 (17%) |
| CELDT $\bar{X}$ (SD) **(248-741) 556 pass** | 568.5 (45.8) | 612.9 (30.7) | 580.4 (47.4) | 624.8 (28.8) |
| CST ELA $\bar{X}$ (SD) **(150-600) 325 pass** | 339.7 (53.0) | 384.5 (50.9) | 325.4 (49.5) | 368.9 (52.7) |
| N (%) | 1,519 | 412 | 1,405 | 302 |
| Essays % Passed **(1-4) 3 pass** | 1.9 (0.48) | 3.0 (0.27) | 2.0 (0.48) | 3.0 (0.21) |

N*ote*. This table only includes students who have both CELDT and CST scores. California requires a minimum score of 556 on the CELDT and 300 on the CST. Both school districts require a 556 or higher on the CELDT and a 325 on the CST. Manzanita district also requires that their 8th graders earn a 2.0 GPA or higher. Granada district also requires that their students pass an essay exam with a 3.0 or higher.

# CHAPTER 2

A Mixed-Method Study: Language Classification and English and Math Course Placement in Middle Schools


## Abstract

English Language Learners (ELLs) are students who are considered non-proficient in English and should be receiving academic support to become proficient. This mixed-method study shows that in one large Southern California school district language classification can sometimes have unintended consequences. According to district administrators, language minority students are placed in a variety of English and math courses based on their academic skills, but not their language classification. Based on ordinary least square regressions we would have concluded that ELL students are more often placed in two period remedial English courses and basic math courses than their RFEP peers. However, using regression discontinuity models I determined language classification itself affects both English and math course placement, net of skills, but more specifically for the third cohort. The types of courses offered and enrollment changed throughout the three cohort years. When ELL students began to be grouped and placed into appropriate English courses, unfortunately, this also led to them being grouped into lower-level math courses.



*Keywords:* English Language Learners, Middle Schools, English Course Placement, Math Course Placement, and Regression Discontinuity

English Language Learners (ELLs) are students considered non-proficient in English. According to the California Department of Education (CDE), ELL students must be provided with English language development instruction, targeted at their respective English proficiency levels. Districts can decide what they consider appropriate academic instruction to help their ELL students reach English proficiency. However, ELL students, on average, score lower than non-ELL students in English reading, writing, and comprehension, but also in less language-intensive subject areas, such as mathematics and science (Edwards, Leichty, & Wilson, 2008; Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hampden-Thompson, Mulligan, Kinukawa, & Halle, 2008). Lower achievement in these subjects may be due to a lack of English proficiency, or it may be the ELL classification itself, which could have unintended consequences if the classified students do not have the opportunity to learn rigorous educational content or are not appropriately integrated with their non-ELL peers.[14] Thus, it is unclear whether it is actually this group's language skills, the stigma associated with the ELL label, or a lack of access to rigorous courses that drives the achievement gap between ELL and non-ELL students.[15]

ELL students should be receiving academic support to become proficient in English. However, all too often, they are also receiving watered-down curricula in numerous subject areas. This study particularly focuses on English and math course placement—two important subject areas that measure academic readiness—with the idea that English placement should be influenced by language classification, but not math. Both elementary and high school ELL

[14] Zaragoza-Petty (2013) describes "opportunities to learn" as the school processes that shape and contribute to student learning. Here, the focus is on school practices that influence students' achievement. A thorough explanation of "opportunities to learn" is provided in the theoretical section of this study.

[15] Language classification and English proficiency can be different. A student is considered English proficient if they pass the CELDT 556+ and the CST 325+. However, a few students who did not pass these thresholds (not proficient) were reclassified, and some students who did pass these thresholds (proficient) were not reclassified. Therefore, it is possible to compare the effects of English proficiency and language classification separately.

students tend to be placed in classes that are separate from their non-ELL peers, and research has shown that these courses tend to be less rigorous (Callahan, 2005; Callahan, Wilkinson, & Muller, 2010; Gandara et al., 2003; Hahnel, Wolf, Banks, & LaFors, 2014; Mayer, 2008). However, we know less about middle school students' circumstances, where the potential inequalities in student learning opportunities may also compound certain disadvantages facing ELL students.

The CDE requires children who speak another language at home to take and pass the California English Language Development Test (CELDT) to assess their English proficiency, and the California Standards Test (CST) in English Language Arts to assess English skills.[16][17] English only (EO) monolingual students are not required to take these tests for language classification purposes. IFEP students are students who speak another language at home but have passed the language classification assessments on their first attempt. ELL students are students who failed at least one of the tests. Each school year ELLs can retake these tests to become Reclassified Fluent English Proficient (RFEP) and be considered English proficient.

To investigate the specific situation of middle school learners, I use data from one large Southern California district to examine whether language classification by the end of seventh grade influences how students are placed into their eighth grade English and math courses. My work focuses on English and math course placement in middle school, which often determines student course placement in high school (Walqui et al., 2010). Math courses typically follow a hierarchical sequence, beginning with pre-algebra and continuing through to calculus. English courses typically include mainstream English classes for non-ELL students, and separate English

[16] The CDE requires that students with disabilities to take the CELDT, but may receive accommodations based on their individualized education programs (IEPs) (http://www.cde.ca.gov/ta/tg/el/celdtfaq.asp).
[17] The CST and CELDT are being replaced by the ELPAC and SBAC (Umansky et al., 2015). This study takes place before these changes occurred.

classes for ELL students that are preparing them for mainstream courses. Advanced English and math courses prepare students academically for college preparatory courses in high school, and, ultimately, college.

Of course, ELL middle school students are a diverse and complex group, which, in part, stems from differences in English proficiency, national origin, socioeconomic status, previous schooling, and the number of years in the U.S. system (Callahan, 2005; Krashen & Brown, 2005). For this reason, there are many aspects to consider when discussing ELL students' academic needs and how educators might improve their academic opportunities. Specifically, I address the following research questions:

1) How is English and math course placement determined for ELL and RFEP middle school students? (Qualitative)
2) How does language classification (ELL and RFEP) affect middle school English and math course placement? (Quantitative)
    a. What is the peer achievement composition in ELL and RFEP students' English and math classrooms? (Quantitative)

In the first analysis, I focus on district administrators interviews regarding language classification course placement practices. In the second analysis, I focus on students who RFEP in seventh grade versus those students who remain ELL and make causal inferences based on regression discontinuity models. The quantitative analysis will only include Long-Term English Language Learners (LTELL) who have been classified ELL for six or more years and foreign-born children. Both analyses are meant to ascertain whether language classification itself determines student course placement, and will deepen our understanding of middle school course placement practice. The work will also provide evidence of how language classification in middle school influences students' academic opportunities to learn.

## Literature Review

### Classroom Course Placement and Tracking

ELL students should be receiving academic support to become proficient in English but are frequently placed in ineffective remedial and support English courses (Menken & Kleyn, 2010; Olsen, 2010). Furthermore, ELL students can be placed in an overarching track, such that students who are exposed to low-level instruction in one subject tend to be subjected to low-level instruction in all areas. For example, high school ELL students are tracked in this way, and are less likely to be enrolled in college preparatory coursework in math, science, or the social sciences (Callahan, 2005; Callahan et al., 2010; Gandara et al., 2003; Hahnel et al., 2014). ELL students are also less likely to be enrolled in advanced courses, such as honors and advanced placement courses (Callahan, 2005; Callahan et al., 2010), and more likely to be placed in special education (Artiles, Rueda, Salazar, & Higareda, 2005). ELL students can also be tracked in a less obvious way, such as when they are placed in English and math courses that have similar course titles as non-ELL classes but are taught separately. For example, Cogan, Schmidt, and Wiley (2001) find nearly 30% of U.S. eighth-grade mathematic course titles do not match the textbook employed. This type of ELL student tracking is mainly researched at the elementary school level. Here, students tend to be "segregated by classroom," where 25% of the first-to-fourth grade California teachers have 50% of the English language learners (Gandara et al., 2003). This means that elementary and high school ELL students may not have opportunities to be in the same classrooms as non-ELL students.

The literature on the reclassification effects on course placement focuses on high school students and compares ELL with various groups of students. Gandara et al. (2003) study show that 21% of the ELL high school students enrolled in college preparatory courses compared with

58% of English only students. Callahan (2005) finds 98% of ELL in her sample did not fulfill the college preparatory courses (e.g., English, math, science) required to apply for entry into California's 4-year state colleges and universities. However, based on propensity score matching, Callahan et al. (2010) concluded that language minorities placed in ESL courses are less likely than non-ESL language minority students to enroll in college preparatory science courses, but the likelihood of enrollment in college preparatory math or social science did not differ. Being placed in less rigorous courses and away from non-ELL peers has been shown to negatively affect ELL students' academic achievement in both English and math (Callahan, 2005; Menken & Kleyn, 2010; Mosqueda, 2010; Mosqueda & Maldonado, 2013).

To my knowledge, Robinson (2011) is the only study that has used a regression discontinuity (RD) model where we can make causal inferences. His results show that language classification itself did not affect English, math, science, or other college preparatory course placement in high school. However, this study combined RFEP students, regardless of when they reclassified. Some researchers demonstrate students who reclassify earlier may be academically differentiated from students who reclassify in later grades (Halle, Hair, Wandner, McNamara, & Chien, 2012; Olsen, 2010). Therefore, I focus on middle school students who are less frequently studied. Middle school teachers may not have different perspectives of a student's academic abilities if that student has been recently reclassified. Olsen (1997) demonstrates that acquiring a higher level of English proficiency does not guarantee promotion into higher tracked courses. Instead, these students may only be promoted horizontally in the tracking system (Valenzuela, 2010). Alternatively, students who recently RFEP may immediately be treated differently than ELL students and placed into mainstream courses.

Overarching tracking may occur for different reasons. One may be that educators are concerned that ELL students do not have sufficient English proficiency to succeed in advanced courses (Kanno & Kangas, 2014). Additionally, teachers may be reluctant to address ELL students' linguistic needs in their advanced math courses (Callahan et al., 2010), or it may be that they are unprepared to do so. Another reason is that ELL students can be required to take English as a second language (ESL) or English Language Development (ELD) courses scheduled at the same time as certain advanced courses (Callahan et al., 2010; Gandara & Orfield, 2010). It may also be that the parents of these students, and the ELL students themselves, are unlikely to request a higher math course placement than the one assigned (Kanno & Kangas, 2014), partially due to a lack of information. Regardless of the reason, overarching tracking clearly contributes to the ELL and non-ELL achievement gap (Gandara et al., 2003; Mayer, 2008; Schneider, Swanson, & Riegle-Crumb, 1997). Yet, ELL students may benefit from being exposed to advanced subject content, even as they develop their English proficiency (Callahan, 2005; Mosqueda & Maldonado, 2013). Likewise, these students can profit from interacting with their non-ELL peers (Gandara et al., 2003). In fact, language proficiency itself may not be the central issue inhibiting academic success; rather, it may be a lack of opportunity.

To my knowledge, no research has specifically studied language classification and course placement in middle schools, which is a crucial period, particularly for ELL students, that determines achievement gaps that can persist throughout high school. ELL middle schools require special attention because most of them are comprised of LTELL (Olsen, 2014) who need to acquire English proficiency and academic knowledge to prepare for high school and, ultimately, college.

**English Course Placement and Achievement**

According to the CDE, ELL students should be provided with English language development (ELD) instruction targeted to their English proficiency level and appropriate level of academic instruction (California Department of Education, 2015). The ELL academic setting can be structured English immersion (SEI), English language mainstream (ELM), or an alternative program (Alt). SEI classroom instruction is provided almost entirely in English, but with a curriculum designed for children who are acquiring English proficiency. ELM is a classroom setting for RFEP students who continue to receive appropriate educational services in order to "recoup any academic deficits that may have been incurred in other areas of the core curriculum as a result of language barriers" (California Department of Education, 2015). Alt classroom instruction is provided in the student's primary language. The state claims Alt instruction should be available only through a parental exception waiver. In theory, the instructional strategies in these classes are meant to scaffold content for English language learners.

Secondary school ELL students who have been in any form of specialized instruction are more likely to score at grade level and are less likely to drop out of high school than those who were placed directly into mainstream settings (Thomas & Collier, 2002). However, critics argue that most ELL students are placed in weak language development programs that are poorly implemented, or are placed in mainstream courses without any resources (Olsen, 2010). Researchers have shown that ELL students are taught at a slower pace and cover less language content than their mainstream counterparts (Walqui et al., 2010). Additionally, few districts have designated programs for their long-term English language learners in secondary schools (Olsen, 2010). Furthermore, the CDE does not provide advice on the types of math, science, or other non-English courses that ELL students should take. For those classes that are not as language intensive, such as math, ELL students should be able to qualify for the same regular and advanced courses as their non-ELL peers.

**Math Course Placement and Achievement**

ELL students are more likely to be placed lower in the math sequence than their RFEP peers, even though these math courses presumably do not require demanding English language skills. The math courses eighth grade students are enrolled in determine their math track; consequently, students who start higher up in the math sequence are positioned above others (Schneider et al., 1997). ELL students might benefit both academically and economically by accessing more advanced math courses. In general, researchers have shown that taking advanced math courses is strongly associated with students' standardized scores in math (Cool & Keith, 1991; Long, Conger, & Iatarola, 2012), a greater likelihood they will graduate from high school, and a stronger probability that the students will both enter and complete college (Attewell & Domina, 2008; Long et al., 2012; Schiller, Schmidt, Muller, & Houang, 2010; Schneider et al., 1997). Generally speaking, these factors equate to higher labor force earnings (Phipps, 1995; Rose & Betts, 2004). Yet, with the currently available research, we cannot make the assumption nor definitively claim that advanced math courses would be equally beneficial to all ELL students.

Several studies have demonstrated that it is math course completion—and not English proficiency—that best predicts students' math achievement (Callahan, 2005; Mosqueda & Maldonado, 2013). Using data from the state of California, as well as nationally-representative data, Callahan (2005) indicates that math course placement plays a greater role in predicting ELL students' SAT math, CAHSEE math, GPA, and the number of completed high school credits than student language proficiency. Mosqueda and Maldonado (2013) find that English proficiency has been a significant predictor of math achievement (standardized assessment), but that it did not have as long an effect as the courses taken. Both studies include controls, such as prior achievement, immigration status, and teacher preparation; however, neither study can make causal inferences. Nonetheless, this literature suggests that efforts to enroll more language-minority students in rigorous mathematics courses could substantially boost achievement for this specific group of students. Thus, simply stated, opportunity, and not proficiency, may be the reason why ELL students have lower math achievement outcomes.

**Contributions to the Current Literature**

First, most of the literature on language classification and course placement focuses on elementary and high school students (Callahan, 2005; Callahan et al., 2010; Gandara et al., 2003; Robinson, 2011) whereas I focus on middle school students. Second, both the literature on language reclassification and course placement has selection bias limitations. Students are not randomly reclassified RFEP, and they are not arbitrarily placed in advanced English and math courses. Students who are being reclassified in earlier grades may be students who are more academically inclined; therefore, the effect of their higher math achievement scores may be spurious, due to some other unmeasured variables (e.g., motivation, cognitive abilities, parental expectations, etc.) that capture academic abilities. ELL students who are placed in higher math courses may also be certain types of students teachers deem capable because of some unmeasured variable that teachers observe firsthand, but that tests do not capture. On the other hand, ELL students who are reclassified earlier than their peers, or who are placed in more advanced courses may be receiving these opportunities because of teacher bias against those who remain classified ELL. Additionally, certain groups, such as Asian American and high-income

ELL students, may be given the benefit of the doubt when it comes to reclassification and course placement, regardless of actual ability. Most current studies have not addressed these issues. More rigorous statistical and methodological steps are needed to determine the possible benefits of being reclassified, particularly for middle school students, and how this might influence their English and math course placement.

## Theoretical Framework

### Opportunities to Learn

To understand why ELL students underperform academically throughout their school years, I concentrate on how a student's language classification can influence that person's opportunity to learn (OTL), particularly in English and math. There is no unanimity among researchers and educators on how to define or measure OTL (Stein, 2000; Zaragoza-Petty & Zarate, 2014). Many researchers have extended the OTL term to include teacher quality, working conditions, curriculum quality, instructional processes, class size, and fiscal resources (Boscardin et al., 2005; Guiton & Oakes, 1995; Oakes, 1986; Porter, 1991; Stein, 2000), but, generally speaking, researchers focus on the schools' "inputs" that influence student achievement (Porter, 1991) instead of the individual student. In my work, I both contribute to and extend the understanding of the OTL term by examining the types of English and math courses that language minorities receive in middle school. In theory, ELL students are supposed to receive instruction that will bring low English proficient students up to par with their peers. However, often they are placed in remedial courses in numerous subject areas. Furthermore, course placement in middle school can determine course placement in high school (Walqui et al., 2010) and eventually in college. Ultimately, I want to understand how language classification influences ELLs' OTL, which I consider to be course placement and access to rigorous content knowledge.

## Hypotheses

To address my research questions I provide two hypotheses. First, I hypothesize that middle school ELL and recently RFEP students will be placed in different English classrooms. After all, the intention of language classification is to determine who needs additional language support to become proficient in English, with the expectation being that once students are reclassified RFEP, they will be enrolled in regular English courses with other non-ELL students. Second, I expect that ELL classification itself may have an effect on eighth-grade math course placement based on literature focusing on high school math course placement. Some research demonstrates language minorities with similar prior math scores are placed in lower math courses (Callahan, 2005; Mosqueda, 2010). On the other hand, results may coincide with research that used more rigorous statistical methods, where language classification did not affect enrollment in college preparatory courses, including mathematics (Callahan et al., 2010; Robinson, 2011). However, these studies do not necessarily compare ELL with RFEP students, and all of them focus on high school math course placement.

## Data Source and Sample

To study these hypotheses I use district data through the Spencer-funded Evaluating the Quality of Universal Algebra Learning (EQUAL) project. I focused on middle school students from one diverse southern California school districts with eight middle schools (see Table 2.1).[18] I chose to focus my investigation on the district I refer to pseudonymously as Manzanita Unified School District (MUSD) because it had a large percentage of ELLs, and more specifically, a

[18] This district had a total of 13 schools but the five schools that were removed were non-traditional schools. Four schools were for academically struggling students and one was for gifted students.

large number of Hispanic and Asian students. Furthermore, this district provided insight into practices that prevail in a relatively low-income community. The project provided qualitative data through interviews and district documents related to reclassification and to math course and English course placement policies. The quantitative data included student-level demographic, attendance, suspension, transcript, and achievement data (e.g., CELDT, CST) from district administration records from 2010-2013. This data makes it possible to measure a variety of school practices related to course placement without relying solely on student reports of their track locations.

In this mixed-methods study, I used the concurrent embedded strategy (Creswell, 2013), where I used qualitative data as a supporting role to the quantitative data. First I present a qualitative piece that describes the types of courses offered at MUSD and how language minorities are placed into those courses. MUSD offered different types of courses, and thus it is important to understand the types of courses offered before doing quantitative analyses. Then I will present a quantitative piece to determine the effects of language classification on middle school English and math course placement.

## Qualitative Section: English and Math Course Placement Policies

### Qualitative Methods

To investigate how eight graders were placed into English and math courses, I used district administrators' interviews and district documents. Between March 2015 and November 2015, I interviewed three MUSD district administrators (i.e., director of the English Language Program, Language Assessment Center supervisor, and curriculum specialist for English learner services) since language classification and course placement policies were established at the district level. The purpose was to understand the district classification and course placement policies, and how the administrators viewed the implementation of those written policies. I explicitly asked them questions regarding 2010-2013 policies to help me analyze and interpret . First, I formally interviewed each district administrator (see interview questions in Table 2.1A in the Appendix). Next, during the summer months (June 2015-August 2015), I worked in this district and had several informal conversations with these same district administrators, as well as a few others. Finally, between January 2016 and July 2016, I shared my qualitative and quantitative results with the district administrators, and based on their feedback, I conducted more analyses.

**Qualitative Analyses**

I conducted formal interviews, which typically lasted 30-45 minutes (see Appendix Table 2.1A for the interview questions). I wrote shorthand notes as I conducted the interviews, and after each interview—on the same day—I edited and typed each participant's complete responses to my questions. District administrators described their job responsibilities to provide a greater understanding of how they have been involved, directly or indirectly, in classifying language minorities. They also discussed whether they thought they had influenced the academics of language minorities in any other way. Additionally, the district administrators described the language classification policies of their district. Administrators also discussed how ELL and RFEP middle students were placed into English and math courses, and described other academic support services their students received. I assigned each district administrator a number to keep his/her personal responses confidential and secure.

During the summer months, I had several informal conversations with these same administrators. I also wrote shorthand notes after each conversation—on the same day—I edited and typed each exchange. In these more conversational talks, they provided me with different information that included explanations of course placement policies for ELL and RFEP students in different school years. Once I had preliminary qualitative and quantitative results, I discussed the results with my participants as a group to get greater clarification on any inconsistencies.

## Qualitative Results

### District Administrators' Descriptions of English and Math Course Placement

How is English and math course placement determined for ELL and RFEP middle school students (Research Question 1)? I provide district administrators' descriptions of and documentation on the types of courses offered, and how students were placed into those courses. Middle school eighth graders may be placed in a variety of English and math courses based on various assessments, teacher and counselor recommendations, and, particularly for English courses, the number of consecutive years a student has been in the U.S.

**English courses.** Table 2.2 shows MUSD eighth graders can be placed into ELD, Intensive Literacy, regular English, or honors English based on CELDT and CST assessments, teachers/counselors recommendations, and the number of years the student has been in the U.S. ELD courses were provided in a two-period time block, and these were for students who had been in the U.S. for less than five years, had CST ELA scores at 2 or lower, and had CELDT scores at 3 or lower. There were three ELD levels; however, in general, these courses were intended to prepare ELL students for regular English courses. Intensive Literacy courses were for those students who had been in the U.S. more than five consecutive years (i.e., LTELL or native born), but who had CST scores at 2.5 or lower, had CELDT scores that were at 4 or lower, and had independent reading levels at 3.5 or lower. As stated in the Manzanita's school handbook, Intensive Literacy courses were provided in a two-period block, and concentrated on preparing ELL and non-ELL students for mainstream English courses. However, for the third cohort of eight graders, English language mainstream (ELM) was offered along with an Intensive Literacy course, and was intended explicitly for LTELL. Students who placed in one regular English course were those who scored 3.0 or greater on the CST. Additionally, where applicable, the students' CELDT scores could also be considered, but, here, no specific passing score was provided. Those students placed into honors English were individuals identified as GATE, or those who were recommended by either a teacher or a counselor, and had a CST score at 4 or greater.

MUSD placement policies stated, "Mainstream English classes are heterogeneous (EL, RFEP, IFEP and EO students). Cohort groups of ELs may be considered for targeted instruction." In theory, ELL students may be placed in any of the English courses, however, in order for ELL students to reclassify, they must score a 4 or greater on the CELDT and a 3.5 or greater on the CST. Therefore, in actuality, ELL students may only be placed in two period English courses (e.g., ELD, ELM, and Intensive Literacy) unless a teacher or a counselor makes an exception. In the quantitative section, I will examine if language classification itself determines ELLs' placement in a two-period English course versus a one period English course. I will also run the analyses separately by cohort because the ELM two period courses were being offered particularly for LTELL in cohort three.

**Math courses.** The director of the ELL program stated that a student's language classification was never considered in math course placement decisions. The placement for eighth-grade math courses was based on previous CST math scores and teacher/counselor recommendations. However, the district did not set specific CST math scores to determine math course placement. District administrators stated that eighth graders could be placed into pre-algebra, basic algebra, regular algebra, honors algebra, honors algebra in Spanish, and geometry.

However, some of these math courses were only offered some school years and enrollment also varied by year. “Algebra HP in Spanish” was only offered in the last two cohorts, “Basic Algebra” was only offered in the first two cohorts, and “Pre-Algebra” was only offered in the third cohort. In 2010-2012, most Manzanita students were placed in an algebra course to provide them the opportunity to access advanced math courses in high school, however these students were placed into basic, regular, or honors algebra. The lowest math skilled students were placed in basic algebra, and most ended up retaking algebra in high school. In 2012-2013, the district moved away from algebra-for-all because too many students were failing algebra, and there was an increase in the number of students placed into pre-algebra. In the quantitative section I will examine if language classification itself determines ELLs’ placement in basic math courses (i.e., basic algebra and pre-algebra) versus accelerated math courses (i.e., regular algebra, honors algebra, and geometry). Eight graders who enroll in a basic math course are likely to enroll in an algebra course in ninth grade unlike students who enroll in an accelerated math course. I will also run the analyses separately by cohort because the math courses offered changed particularly for the third cohort.

The qualitative results informed me how to conceptualize the course placement for quantitative results. The remainder of the study focuses on student-level district data. In the discussion section, I will provide concluding remarks regarding both analyses.

## Quantitative Section: Classification Effects on Course Placement

### Quantitative Data

To investigate how language classification affects eight-grade English and math course placement, I use student-level district data. Here, I compare students who reclassified in seventh grade and ELL middle school students (Table 2.1, the selected sample). I excluded students who had been reclassified RFEP in elementary school. I also excluded both White and African American students, who were mostly EO students. About 26% of the district students were classified ELL in sixth grade and 19% of them reclassified RFEP in seventh grade (hereafter referred to RFEP). About 75% of them were LTELL and the remaining where foreign-born students.

I further restricted the data to students who had both CELDT and CST scores, referring to them as the final sample. The selected sample (n=4,231 in Manzanita) were different from the students in the final sample (n=2,969 in Manzanita). The final sample had fewer special education, foreign-born, and Asian American students than the selected sample (see Appendix Table 2.2A). Therefore, the final regression discontinuity results are generalized to middle school students who had both CELDT and CST scores—the main reclassification criteria.

**Measures**

**English course placement.** The first dependent variable was a student's English course placement in eighth grade (see Table 2.3). Based on interviews and transcript data English courses were grouped into five categories: 1) honors English; 2) regular English; 3) English courses provided in two periods (i.e., a combination of regular and/or intensive); 4) ELD courses provided in two periods; and 5) ELM course and one regular English course provided in two periods. Based on qualitative results, Intensive Literacy, ELM, and ELD courses were intended to get students ready for mainstream English courses. For the quantitative analyses English courses were further grouped into two categories: (1) students were enrolled in one English course (i.e., honors or regular English), or (0) students were enrolled in two English courses (i.e., a combination of ELM, ELD, Intensive, or regular courses) provided in two class periods.

**Math course placement.** The second dependent variable was a student's math course placement in eight-grade (see Table 2.3). Based on interviews and transcript data math courses were grouped into five categories: 1) geometry, 2) honors algebra, 3) regular algebra, 4) basic algebra, and 5) pre-algebra.[19] For the quantitative analyses math courses were further grouped into two categories: (1) "accelerated" math course (i.e., regular algebra, honor algebra, or geometry) or (0) "below basic" math course (i.e., basic algebra or pre-algebra). Here, I consider eighth grade regular algebra an accelerated course, as it puts students on track to complete calculus by the end of the twelfth grade (Domina, McEachin, Penner, & Penner, 2015). Basic algebra was placed in the "below basic" group because 99% of these students were enrolled in Algebra I once in the ninth grade. Whereas a portion of regular algebra and honors algebra students were enrolled in geometry in the ninth grade, this was more so for the students in honors algebra.

**English peer achievement.** The third dependent variable was the average CST ELA scores for an individual student's eighth grade English classroom minus that individual's English score. The purpose was to determine the probability a student would be placed in a classroom with high skilled English peers based on their language classification. For example, two students may be enrolled in a regular English course but one classroom has the lower English skilled students than the other classroom.

**Math peer achievement.** The fourth dependent variable was the average CST math scores for an individual student's eighth grade math classroom minus that individual's math score. The purpose was to determine the probability a student would be placed in a classroom with high skilled math peers based on their language classification. For example, two students

[19] Honors Algebra is equivalent to Algebra Spanish HP. Basic Algebra is equivalent to Algebra Core Concept I. Pre-Algebra is equivalent to Mathematics 8 and Algebra Readiness.

may be taking Algebra but one classroom has the lower math skilled students than the other classroom.

**Control variables.** The models also included student-level covariates to explain differences in the students' language classification. These covariates included gender (1=female, 0=male), race/ethnicity, birth country, socioeconomic status (SES), and special education status. Racial/ethnic categories included Hispanic (reference group), Asian American, and an "other race" category included American Indian, Alaskan Native, Filipino, Native Hawaiian, and Pacific Islanders. The birth country was a binary outcome, where 1 was coded for those "born in the United States," and 0 was coded for those "born in another country." SES was based on students' "free or reduced lunch" (FRL) status, where students who qualified for FRL (reference group) were compared to students who did not qualify for FRL status. Special education status was binary, where 1 was coded "special education," and 0 was coded "no special education."

**Quantitative Analyses**

I used student-level data to understand whether language classification predicts English and math course placement for eighth graders. I provided ordinary least squares (OLS) effect sizes to compare it to the regression discontinuity (RD) effect sizes. In the following sections, I explain both methods in great detail. I compared seventh graders who remained ELL versus those students who were reclassified RFEP in the seventh grade because these two student groups were the two that were most academically comparable. As language classification is determined at the end of the school year, the seventh-grade classification can influence students' eighth-grade course placement.

## Language Classification Association with Course Placement

To address the second research question, I ran two separate OLS models to estimate the effects of student's language classification by the end of seventh grade on their eight-grade English and math course, accounting for the student's race, gender, birth country, free or reduced lunch, special education, cohort, eighth grade school, and teacher fixed effects. These models can be expressed as:

$$Y_{it} = \beta_0 + \beta_1 LC_i + \beta_2 Demo_i + \beta_3 Cohort_i + \beta_4 School_i \ F\delta_{s(i)} + e_i , \quad \text{(F1)}$$

In (F1), $Y_{it}$ represents student $i$'s English or math course in the eighth grade; $LC_i$ represents whether a student reclassified RFEP in seventh grade (1) or remained ELL (0); $Demo_i$ stands for the student's characteristics, such as race, gender, country of birth, FRL, and special education status; and $\text{Cohort}_i$ is the school year in which the student was an eighth-grade student. In addition, current middle $School_i$, and teacher fixed effects $F\delta_{s(i)}$ were also included to control for annual changes and school factors (e.g., other unmeasured confounders) that might also have explained course placement using Stata's *xtlogit,fe* command (StataCorp, 2011). The betas (β) represent the estimated increase in the outcome per unit increase in the value for each given covariate variable. In the English model, $\beta_1$ represents the increase in the student's probability to be placed in "One English course" (1) versus "Two English Courses" (0) based on language classification. In the math model, $\beta_1$ represents the increase in the student's probability to be placed in "Accelerated Math" (1) versus "Basic Math" (0) based on language classification.

As a robustness check, I ran two other OLS models to estimate the effects of students' language classification by the end of seventh grade on their eight-grade classroom peer English and math achievement distribution, accounting for the same factors in F1 model. These OLS models will address the second research sub-question: what is the peer achievement composition in ELL and RFEP students English and math classrooms? The purpose was to determine if language classification affects the types of peers in the individual's English and math classroom regardless of the course titles. Course titles may be only explaining the "tip of the iceberg." Language minorities may be further divided in classrooms that share the same course title but their peers' English and math skills differ.

**Estimating the Effects of Language Classification on Course Placement**

The RD design can provide *causal inferences* that are "as good as random assignment" and has strong internal validity (Lee & Lemieux, 2010). Therefore, I ran two separate RD models to determine whether reclassifying in the seventh grade effects students' eighth-grade English and math course placement. As a robustness check, I also ran two other RD models to determine whether reclassifying in the seventh grade affects their eight-grade classroom peer English and math achievement distribution. I exploited the fact that the Manzanita district chooses a cutoff based on the CELDT and CST. Seventh grade language classification was based off student's seventh grade overall CELDT scores and sixth grade CST ELA scores. The CELDT ranged from 248-741, and to reclassify MUSD required a score of at least 556.[20] The CST raw scores ranged between 150 and 600, and MUSD required students to meet a score of at least 325, considered "mid-basic." Failing to meet even one of the requirements could have been enough to prevent a student from being reclassified.

For the RD approach to yield valid causal inferences, we must meet four key assumptions. First, the treatment must be endogenous. Here, language classification 0 for ELL and 1 for RFEP (the treatment) was determined by a set formula based on the CELDT and CST scores (the assignment variable). Second, the students and teachers must not be able to

[20] The CELDT subcategories' raw scores and pass cutoffs also varied. The scores for English language listening ranged from 230-715, speaking 225-720, reading 320-750, and writing 220-780. Manzanita seventh graders were required to score at least a 495, 476, 529, and 508 in listening, speaking, reading, and writing, respectively.

manipulate the assignment variable (Imbens & Lemieux, 2008). As the CELDT and CST tests are based on multiple questions, it would be difficult, if not impossible, to manipulate them to be either right above or right below the cutoff. This allowed me to use an RD to compare the outcomes of the students just above and just below the threshold. These two groups of students were nearly identical in all ways, with the exception that the former group was recommended to remain ELL while the latter was recommended to RFEP. Arguably the students near the cutoff also had similar English proficiency because the 556 CELDT and 325 CST are arbitrary cutoffs. Third, the CELDT and CST must be normally distributed and it may not have a jump at the threshold (see Appendix Figures 2.1A-2.3A). Fourth, there cannot be a discontinuity in covariates (e.g., race, gender). Therefore, I regressed each covariate on binary variables with the combined CELDT and CST assignment variable (see Appendix Figure 2.4A).

For the RD model, I created an assignment variable "CELDT/CST" that included a combination of the overall CELDT and CST score and to reduce the dimensionality to one composite score (Reardon & Robinson, 2012; Wong, Steiner, & Cook, 2013).[21] I centered and standardized the overall CELDT score at 556 and the CST score at 325—the requirements for MUSD seventh graders. The CELDT/CST represents the minimum score of the centered and standardized CELDT and CST scores, therefore a score lower than zero denoted the student had failed at least one exam. The assignment variable was used in RD models to determine language classification effects on student's probability of being placed in one English course (1) instead of two English courses (0), and being placed in accelerated math courses (1) instead of below basic math courses (0) when they were near the cutoff. The assignment variable was also used in two other RD models to determine if language classification affects the individual's students eighth

[21] As a robustness check, I created another assignment variable that also included the CELDT four subcategories (listening, speaking, reading, and writing) at their set required cutoffs. The results remained the same and for simplicity I only discuss the assignment variable with the overall CELDT and CST scores.

grade English and math classroom peer achievement distribution.

The RD design was implemented by estimating the equations of the following general form:

$$Y_{ist} = \alpha + \beta I(CC_{ist-1} < 0) + f(CC_{ist-1}) + \varepsilon_{ist} \ , \qquad \text{(F2)}$$

In (F2), $Y_{ist}$ is the outcome (e.g., course placement) for student *i* in school *s* in year *t*. The variable, $CC_{-1}$, is the "assignment variable" in this RD design, based on middle school students combined CELDT and CST scores. The parameter of interest, β, identifies the jump in outcomes when the middle school language minority student is above the assignment threshold, conditional on $f(CC_{ist-1})$ a function of the assignment variable which was estimated using local linear regressions. Those students arbitrarily close to the cutoff, on either side of the threshold, were observationally and non-observationally similar, and thus, they could be used as proxies for each other's missing counterfactual. I used Nichols (2007) RD Stata package, which required neither student-level covariates nor school factors. I found the discrepancies between eligibility and reclassification did not increase sharply from 0 to 1 at the threshold; as a result, it is considered a "fuzzy" RD (Papay, Murnane, & Willett, 2011). Therefore, I report the denominator of the local Wald estimator—the jump in the treatment (Stage 1)—and the local Wald estimator (numerator/denominator)—the jump in the outcome (Stage 2). Stage 1 results are alike throughout all the RD models. See Appendix Figure 2.5A as one example of these discontinuities. To conduct RD models the assignment variable (CELDT/CST) must predict the treatment (ELL and RFEP). On average, ELL middle school students had a 40% probability of being reclassified if they passed the CELDT at 556 and CST at 325. This probability ranged from 35% to 50% depending on the cohort. Administrators followed the recommended cutoffs more closely in the later school years. Furthermore, ELL students below the cutoff had about

10% probability of being reclassified. This fuzzy RD allows us to compare students near the cutoff threshold.

As a robustness check, I ran two other RD models to estimate the effects of students' language classification by the end of seventh grade on their eighth-grade classroom peer English and math achievement distribution, accounting for the same factors in F2 model.

The RD models are limited in two ways. First, it is only possible to estimate regression discontinuity analyses in districts that have implemented a formula-based placement system. These analyses tell us little about the extent to which various district language classification policies moderate the effects of course placement in districts that do not follow similar policies. Thus, these studies' results will be generalizable to school districts with similar language classification policies and policy-implementation procedures. Second, these analyses only estimate the effects of language classification near the classification threshold, providing limited evidence for those students who scored either very high or very low on the CELDT and CST exams. To address this issue, I provided several bandwidths as robustness checks for the RD models. Nichols (2007) RD Stata package automatically selected the three optimal bandwidths. The bandwidths varied for the English and math course placement models and they also varied by cohort year. I present the effect sizes based on the Nichols (2007) RD Stata package bandwidths. For the English and math models the optimal bandwidths are about 0.5, 1.0, and 1.5 standard deviations away from the required CELDT/CST cutoff. I rounded the bandwidths to the nearest tenth but the effect sizes presented are based on the exact bandwidths determined by the Stata program. Also, OLS regression coefficients were presented next to the RD coefficients to compare effect sizes.

## Quantitative Results

ELL students could be placed in an overarching track, such that students who were enrolled in low-level instruction in English tend to be subjected to low-level instruction in math too. A chi-square test of independence shows a strong relationship between English and math course placement ($\chi^2$ (2, $N$=15,332)=2.1, $p$ <.001). About 80% of geometry students were also enrolled in honors English and 20% were enrolled in one regular English course; 62% of the honors algebra students were enrolled in honors English and 36% were enrolled in one regular English course; and 18% of regular algebra students were enrolled in honors English, and 71% were enrolled in one regular English course. As for students in basic algebra, 56% were in one regular English course and 34% in two remedial English courses; and 43% of students in pre-algebra were enrolled in one regular English course and 36% in two remedial English courses, while 12% were enrolled in two ELD or ELM/regular English courses. Furthermore, ELL students were more likely to be placed into two English courses, and they were more likely to be placed into basic math courses. The following analyses will determine if language classification itself determines a student's eighth grade course placement.

### Course Placement for Middle School English Language Learners

Educators and researchers may argue that language classification itself does not influence course placement because students who become RFEP have higher English and math scores than ELL students (see Table 2.4). I conduct OLS and RD models to address how language classification affects middle school English and math course placement (Research Question 2). The OLS models provide estimates of the language classification effects for all middle school language minorities. The RD models provide us with a more precise and less biased estimate than the OLS models.

**English course placement.** Eighth grade minorities may be enrolled in one English course (1) or two English courses. Based on OLS models, Table 2.5 shows that RFEP students had approximately a 19% probability ($p < .001$) of being placed into one English course instead of two courses when compared with ELL students. The results increased from 18% to 25% ($p < .01$) and remained statistically significant when the OLS models were run separately by cohort. However, the OLS models do not account for the strong relationship between language classification and English assessment scores.

RD models account for spurious factors, showing that language classification itself did not always determine a student's English course placement (see Table 2.5). The probability ranged between 16% (NS) to 24% ($p < .01$) when the cohorts were combined. The RD effect size was comparable to the OLS estimate of 19%, but the results were only statistically significant at the largest bandwidth. The Stage 2 graph in Figure 2.1 visually shows the discontinuity when the cohort data was aggregated. However, for the first two cohorts, a student's course placement was not determined by the student's classification itself. In 2012-2013, students classified RFEP had a 48% to 51% ($p < .01$) probability of being placed into one English course instead of two English courses when compared with ELL students who had nearly equivalent CELDT and CST scores, even though arguably these students had equivalent English proficiency levels. For the third cohort, the RD effect size and standard errors was larger than OLS estimate of .25(.06). Figure 2.2 shows the discontinuity when the RD models were conducted separately for each cohort.

Cohort differences can be explained due to changes in courses offered and enrollment. In the third cohort, decreases in the number of students enrolled in honors English (from 4% to 2%) and in the number of students enrolled in double period remedial courses (from 44% to 24%)

occurred. Conversely, an increase in the number of students enrolled in one regular English course (from 40% to 44%), two double periods ELD or ELM courses (5% to 7% and 0% to 15% respectively), as well as an increase in the number of students not enrolled in any English course (2% to 5%) occurred. In the last cohort, a shift from two period block remedial courses to specialized two period ELD/ELM general courses, particularly for ELL students, emerged.

**Math course placement.** Eighth grade minorities may be enrolled in an accelerated math course (1) or a basic math course (0). Based on the OLS results, Table 2.6 shows that RFEP students had a 17% probability ($p < .001$) of being placed into an accelerated math course instead of a basic math course compared with ELL students. The results remain similar and statistically significant when run separately by cohort. The OLS results may be spurious where endogenous factors may explain why certain students RFEP in seventh grade, and why they are placed in an accelerated math course.

The RD models showed language classification itself determines student math course placement (see Table 2.6). The probability ranged between 18% (*NS*) to 25% ($p < .01$) when the cohorts were combined. The RD effect size was comparable to the OLS estimate of .17(.03), but the results were only statistically significant at the 0.5 and 1.5SD bandwidths. Figure 2.3, the Stage 2 graph visually show the discontinuity when the cohort data was aggregated. However, in Cohorts 1 and 2, RFEP students had a 10% probability (*NS*) of being placed into accelerated math courses, but it was not statistically significant. Cohort 3 RFEP students had a 24% ($p < .05$) to 29% (*NS)* probability of being placed into accelerated math courses. For the third cohort, the RD effect size and standard errors was larger than the OLS estimate of .17(.05). Figure 2.4 shows the discontinuity when the RD models were conducted separately for each cohort.

Generally, language classification in seventh grade determined eighth grade math course placement, but it was mainly driven by the third cohort. In 2012-2013, decreases in the number of students enrolled in geometry (from 1% to 0.5%), honors algebra (from 6% to 3%), regular algebra (from 31% to 18%), and basic algebra (from 54% to 25%) occurred while, at the same time, an increase in the number of students enrolled in pre-algebra (7% to 41%) occurred, causing a movement to place students in the "appropriate math courses" because too many students were failing algebra.

**Robustness Checks.** As robustness checks, I provide model estimates based on three different bandwidths comprised of students who are 1.5, 1.0, and 0.5 standard deviations away from the CELDT/CST combined required cutoff (i.e., 556 on the CELDT and 325 on the CST). Coefficients became non-significant for both English and math course placement when the bandwidths and standard errors increase. In another RD model, I made the English course placement dichotomous, where, for the reference group, 1 signified "One regular English course" and 0 signified "At least one remedial/intervention English course," while simultaneously excluding ELD courses, but the results remained the same. For the OLS regression models, I excluded special education students, but the results remained consistent.

As a final robustness check, I tested my hypothesis that language classification determined the types of peers ELL students would be exposed to in their English and math classrooms. Course titles may only explain some of the differences between ELL and RFEP students. This was particularly important to determine for the first two cohorts where there was no statistical significance between ELL and RFEP English and math course placement. Based on the OLS results, Table 2.7 shows that RFEP students were placed in classrooms with peers who scored on average 0.35σ higher ($p < .001$) on the CST ELA than ELL students' classroom peers.

However, the RD models demonstrate those results were spurious, and seventh grade language classification did not predict students' eighth grade English peer achievement (0.16σ, *NS*) classroom composition. OLS results in Table 2.8 show that RFEP students were placed in classrooms with peers who scored 0.22σ higher ($p < .001$) on the CST math than ELL students. However, seventh grade language classification did not predict students' eighth grade math peer achievement (0.16σ, *NS*) and classroom composition in the RD models, which suggest that OLS results were spurious. As expected, the English and math coefficients were positive, but unexpectedly not statistically significant. Disaggregating data by cohort produced similarly insignificant results.

## Discussion

At MUSD, language minority students are offered a variety of English and math courses based on their academic skills, teacher/counselor recommendations, and the number of years the student has been in the country. According to district administrators, language classification itself is not considered when making English and math course placement decisions. Theoretically the 81% of ELL middle school students have the same opportunities to learn as the 19% of RFEP students who RFEP in seventh grade. However, district documentation reveals that ELL students are likely to be placed into two period English courses (e.g., ELD, remedial) based on course placement policies. As for math course placement, district documentation supports administrators' accounts. Moreover, it is conceivable that language classification itself may not influence course placement because students who are RFEP have higher English and math scores than ELL students. Using RD models, I determined language classification itself affects both English and math course placement, net of skills, but these results are driven by the third cohort. Language classification did not predict course placement for the other cohorts. English courses offered, and the number of students enrolled in courses, can explain the differences throughout the three cohort years. Furthermore, classroom peer achievement did not show a greater difference between ELL and RFEP eighth grade course placement; therefore, course title explains the main difference amongst language minorities.

Based on the CDE policies, I expected ELL students to receive different English courses than RFEP students because the purpose of classifying students is to provide them suitable and tailored courses to become English proficient. OLS models show there is a strong association between language classification and English course placement for all three cohorts. However, RD results show that only for cohort three did language classification itself determine placement in two-period English courses compared with only one English course. At MUSD, only newly arrived ELL students had tailored two period ELD courses that particularly addressed their recent introduction to the English language. Olsen (2010) recommends that LTELL and newly arrived students take different English courses to become proficient in English. Until the third year of the study, MUSD aimed to provide LTELL more targeted language support. Simultaneously, an increase in students enrolled in two period ELM courses, intended specifically for ELL emerged. The effects of the one versus two period English courses are beyond the scope of this paper.

Based on the CDE policies and the interviews, I expected ELL and RFEP students to have the same opportunities to enroll in accelerated math courses. OLS models show that ELL students are less likely to be placed in accelerated math courses compared with RFEP students. One concern is that ELL students have lower CST math scores than students who RFEP in seventh grade. Students who are being reclassified can be students who are more mathematically inclined; because of this, their higher math achievement scores may be spurious, due to some other unmeasured variable that captures academic ability. Also, the CST math portion is in English; therefore, not surprisingly, ELL students have lower averages than RFEP students. However, ELL students who are reclassified earlier than their peers, or who are placed in more advanced courses, may be receiving these opportunities because of an administrator bias against those who have remained classified ELL. The RD models show the main results are driven by cohort three, where language classification itself determines placement in accelerated math courses versus basic math courses. In 2012-2013, RFEP students may be given benefit of the doubt when it comes to course placement, regardless of actual ability. This means academically comparable ELL and RFEP students are placed into different math courses based on their language classification. By the third year, Manzanita moved away from algebra-for-all. Although the district differentiated between honors algebra, regular algebra, and basic algebra language classification itself seemed to have more of an effect on math course placement than when they explicitly changed basic algebra to pre-algebra.

The main results of this study coincide with prior studies that show ELL high school students are disproportionately placed into low-level math classes, and, particularly for LTELL, in remedial English courses (Callahan, 2005; Callahan et al., 2010; Kanno & Kangas, 2014). However, the main results are driven by the 2012-2013 MUSD course placement policies. The non-statistically significant results based on the first two cohorts coincide with Robinson's (2011) RD findings where language classification itself did not predict English and math course placement for high school students. The third cohort findings of this study may be different from Robinson's (2011) RD estimates for several reasons. First, the district's course placement policies might be different regarding language classification policies. Unlike Robinson's district, Manzanita may be considering students' language classification when making English and math course placement decisions. Second, the effects of language classification on middle school course placement may be different than course placement in high school. Third, it is likely the districts offered different English and math courses because this study included middle school students, and Robinson's study included high school students. Fourth, it is unclear if Robinson considered the timing of reclassification, and whether he only included students who RFEP in high school, or all RFEP students. The estimates of this study are conservative given that I only compared ELL students with students who RFEP in seventh grade. Course placement differences can be much greater between current ELLs and students who RFEP in elementary school.

California districts may be moving in the right direction with the new state assessments and new state standards, but district administrators need to take several aspects into consideration. Starting in 2014-2015, central language classification assessments, such as the CELDT and CST, were replaced by the ELPAC and SBAC (Umansky et al., 2015). Furthermore, the state adopted the English language development (CA ELD) standards that are comparable in rigor and specificity to the California Core State Standards (CCSS), California Common Core State Standards for Mathematics (CA CCSSM), and the Next Generation Science Standards (CA NGSS) (Laguoff, Spycher, Linquanti, Carroll, & DiRanna, 2015). California now makes recommendations about what ELL students should be learning in mathematics and science classrooms. The content should be as rigorous as non-ELL curriculum, and student proficiency should also be developed in all subject areas. These changes can encourage district administrators to provide English and math courses that will develop ELL students' English proficiency, as well as provide them access to rigorous content knowledge. However, historically, "separate but equal" has not worked out. As this study demonstrates, in the third cohort year, ELL students were offered more tailored ELM courses, but at the same time they were more likely to be placed in basic math courses. Furthermore, these two-period English courses had a high representation of non-ELL students with low English skills. Prior research demonstrates that LTELL English proficiency may not improve in remedial courses with non-ELL students who have low English skills (Kanno & Kangas, 2014; Olsen, 2010). Furthermore, the SBAC exam will also be more language-intensive than the CST, and it is probable that ELL students will have lower English and math scores. ELL students may have greater access to rigorous content comparable with their non-ELL peers, however, these comparable peers may be students who have low English and math skills based on the more language-intensive SBAC exam. Lastly, district administrators need to decide if they are going to make course placement decisions based on assessments used for language classification purposes. MUSD used the CST to make language classification and course placement decisions. Districts may attempt to provide equal course placement opportunities to ELL and RFEP students, but their policies may prevent equal access. It remains to be seen if CCSS standards, CA ELD standards, and SBAC will improve ELL opportunities to learn in middle school.

**Limitations**

In interpreting this study's findings, I note the following empirical limitations. First, the quantitative analyses focuses on the 18% of students who have CELDT and CST scores—the required reclassification assessments. About 8% of sixth grade ELL students did not have the necessary scores to be included in the RD models. A second limitation is the lack of Individualize Education Program (IEP) information for special education students. IEP's vary greatly and the specification for special education students can influence there probability of course placement. Special education students are normally dropped from analysis (for example Hill, Weston, & Hayes, 2014) but they remained in these analyses due to the high representation of special education students in middle school. Many special education students are long-term English language learners and the intersection must be further research particularly in middle school. Future research should address the intersection of special education and ELL classification especially when several of these students cannot RFEP since many are exempted from taking the required reclassification exams. A final limitation of this study is that the interviews took place during a period when administrators were preparing to adopt the new state standards and state exams. Many were eager to discuss the implications of the new language classification policies and course placement policies, but were disinclined to talk about the old policies. Yet, in order to research the impacts of the policies, one must first have data to test the implications. The next step here will be to evaluate the districts' new language classification policies and English and math course placement policies.

**Conclusion**

District administrators need to deliberately decide the types of English and math courses ELL middle school students should be offered, as well as the types of courses RFEP students who were recently RFEP should take, and if they need to continue language support. Language minorities should be enrolled and have access to rigorous English and math content, along with the proper language assistance to help them address their particular needs. This study demonstrates that in some circumstances language classification itself can affect student English and math course placement, net of skills, when comparing academically equivalent ELL and recently RFEP peers. In the third study of this dissertation, I assessed whether reclassifying in middle school affects student achievement and behavioral outcome, mainly for the first MUSD cohorts.

References

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children, 71*(3), 283-300.

Attewell, P., & Domina, T. (2008). Raising the bar: Curricular intensity and academic performance. *Educational Evaluation and Policy Analysis, 30*(1), 51-71.

Boscardin, C. K., Aguirre-Munoz, Z., Stoker, G., Kim, J., Kim, M., & Lee, J. (2005). Relationship between opportunity to learn and student performance on English and algebra assessments. *Educational Assessment, 10*(4), 307-332.

California Department of Education. (2015). Facts about English Learners in California- CalEdFacts. http://www.cde.ca.gov/ds/sd/cb/cefelfacts.asp

Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in US schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Cogan, L. S., Schmidt, W. H., & Wiley, D. E. (2001). Who takes what math and in which track? Using TIMSS to characterize US students' eighth-grade mathematics learning opportunities. *Educational Evaluation and Policy Analysis, 23*(4), 323-341.

Cool, V. A., & Keith, T. Z. (1991). Testing a model of school learning: Direct and indirect effects on academic achievement. *Contemporary Educational Psychology, 16*(1), 28-44.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming High and Falling Short California's Eighth-Grade Algebra-for-All Effort. *Educational Evaluation and Policy Analysis, 37*(3), 275-295.

Edwards, B., Leichty, J., & Wilson, K. (2008). English Learners in California: What the Numbers Say. *EdSource*.

Gandara, P., & Orfield, G. (2010). A Return to the" Mexican Room": The Segregation of Arizona's English Learners. *Civil Rights Project/Proyecto Derechos Civiles*.

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources, Unequal outcomes. *Education Policy Analysis Archives, 11*(36), 1-54.

Guiton, G., & Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis, 17*(3), 323-336.

Hahnel, C., Wolf, L., Banks, A., & LaFors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. The Education Trust-West.

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.

Hampden-Thompson, G., Mulligan, G., Kinukawa, A., & Halle, T. (2008). Mathematics Achievement of Language-Minority Students During the Elementary Years.

Hill, L. E., Weston, M., & Hayes, J. M. (2014). Reclassification of English Learner Students in California. *Public Policy Institute of California. Retrieved from www. ppic. org/main/publication. asp.*

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615-635.

Kanno, Y., & Kangas, S. E. (2014). "I'm Not Going to Be, Like, for the AP" English Language Learners' Limited Access to Advanced College-Preparatory Courses in High School. *American Educational Research Journal*, 0002831214544716.

Krashen, S., & Brown, C. L. (2005). The ameliorating effects of high socioeconomic status: A secondary analysis. *Bilingual Research Journal, 29*(1), 185-196.

Laguoff, R., Spycher, P., Linquanti, R., Carroll, C., & DiRanna, K. (2015). *Integrating the CA ELD Standards into K-12 Mathematics and Science Teaching and Learning* Retrieved from WestEd: www.cde.ca.gov

Lee, D. S., & Lemieux, T. (2010). *Regression discontinuity designs in economics*. Retrieved from http://www.jstor.org/stable/20778728

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal, 49*(2), 285-322.

Mayer, A. (2008). Understanding How US Secondary Schools Sort Students for Instructional Purposes: Are All Students Being Served Equally? *American Secondary Education*, 7-25.

Menken, K., & Kleyn, T. (2010). The long-term impact of subtractive schooling in the educational experiences of secondary English language learners. *International Journal of Bilingual Education and Bilingualism, 13*(4), 399-417.

Mosqueda, E. (2010). Compounding inequalities: English proficiency and tracking and their relation to mathematics performance among Latina/o secondary school youth. *Journal of Urban Mathematics Education, 3*(1), 57-81.

Mosqueda, E., & Maldonado, S. I. (2013). The Effects of English Language Proficiency and Curricular Pathways: Latina/os' Mathematics Achievement in Secondary Schools. *Equity & Excellence in Education, 46*(2), 202-219.

Nichols, A. (2007). Causal Inference with Observational Data. *Stata Journal, 7*(4), 507-541.

Oakes, J. (1986). Keeping track, Part 1: The policy and practice of curriculum inequality. *Phi Delta Kappan, 68*(1), 12-17.

Olsen, L. (1997). *Made in America: Immigrant students in our public schools*: ERIC.

Olsen, L. (2010). Reparable Harm Fulfilling the Unkept Promise of Educational Opportunity for California's Long Term English Learners. *California Together (Research Report).*

Olsen, L. (2014). *Meeting the Unique Needs of Long Term English Language Learners: A Guide for Educators*. National Education Association.

Papay, J. P., Murnane, R. J., & Willett, J. B. (2011). *How performance information affects human-capital investment decisions: The impact of test-score labels on educational outcomes*. Retrieved from https://www.researchgate.net/profile/Richard_Murnane/publication/228303490_How_Performance_Information_Affects_Human-Capital_Investment_Decisions_The_Impact_of_Test-Score_Labels_on_Educational_Outcomes/links/0046352854781893380000000.pdf

Phipps, B. J. (1995). Career dreams of preadolescent students. *Journal of Career Development, 22*(1), 19-32.

Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis, 13*(1), 13-29.

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness, 5*(1), 83-104.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics, 86*(2), 497-513.

Schiller, K. S., Schmidt, W. H., Muller, C., & Houang, R. T. (2010). Hidden disparities: How courses and curricula shape opportunities in mathematics during high school. *Equity & Excellence in Education, 43*(4), 414-433.

Schneider, B., Swanson, C. B., & Riegle-Crumb, C. (1997). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2*(1), 25-53.

StataCorp. (2011). Stata Statistical Software: Release 12. StataCorp LP: College Station.

Stein, S. J. (2000). Opportunity to learn as a policy outcome measure. *Studies in Educational Evaluation, 26*(4), 289-314.

Thomas, W. P., & Collier, V. P. (2002). A national study of school effectiveness for language minority students' long-term academic achievement. *Center for Resesarch on Education, Diversity and Excellence. (Resesarch Report)*

Umansky, I. M., Reardon, S. F., Hakuta, K., Thompson, K. D., Estrada, P., Hayes, K., . . . Goldenberg, C. (2015). Improving the Opportunities and Outcomes of California's Students Learning English: Findings from School District-University Collaborative Partnerships. Policy Brief 15-1. *Policy Analysis for California Education, PACE.*

Valenzuela, A. (2010). *Subtractive schooling: US-Mexican youth and the politics of caring*: Suny Press.

Walqui, A., Estrada, P., Koelsch, N., Hamburger, L., Gaarder, D., Insurralde, A., . . . Weiss, S. (2010). *What are We Doing to Middle School English Learners?: Findings and Recommendations for Change from a Study of California El Programs (Research Report)*.

Wong, V. C., Steiner, P. M., & Cook, T. D. (2013). Analyzing regression-discontinuity designs with multiple assignment variables a comparative study of four estimation methods. *Journal of Educational and Behavioral Statistics, 38*(2), 107-141.

Zaragoza-Petty, A. L., & Zarate, M. E. (2014). College access factors of urban latina girls: The role of math ability perceptions *Journal of Urban Learning, Teaching, and Research, 10*, 64-72.

Table 2.1

*Manzanita Full Sample, Selected Sample, and Final Sample (2010-2013 Cohorts)*

| | Full Sample | Selected Sample | Final Sample |
|---|---|---|---|
| *District information* | | | |
| Total 8th grade enrollment | 15,417 | 4,231 (26%) | 2,969 (18%) |
| Average 8th grade cohort | 5,139 | 1,410 | 989 |
| Total # of middle schools | 8 | 8 | 8 |
| Total # of 8th grade English courses | 532† | --- | --- |
| Total # of 8th grade Math courses | 521† | --- | --- |
| *Student demographics in 8th grade* | | | |
| % Female | 49.1 | 45.8 | 47.3 |
| % Hispanic or Latino | 67.7 | 87.6 | 88.3 |
| % Asian | 10.4 | 9.3 | 8.8 |
| % White | 13.3 | --- | --- |
| % African American | 3.3 | --- | --- |
| % Other race0 | 5.0 | 2.9 | 2.8 |
| % Born in United States† | 84.1 | 74.1 | 75.3 |
| % Free- and Reduced-Price Lunch | 73.2 | 90.1 | 90.2 |
| % Special Education | 10.5 | 19.7 | 4.6 |
| *Language Classification* | 7th Grade | 7th Grade | 7th Grade |
| % English Language Learners (ELL) | 25 | 81 | 74 |
| % Reclassified Fluent English Speakers (RFEP) | 38 | 19 | 26 |
| % English Only (EO) and Initially Fluent English Speakers (IFEP) | 37 | --- | --- |
| *Dependent Variables* | 8th Grade | 8th Grade | 8th Grade |
| English Courses | % | % | % |
| Honors English | 20 | 3 | 4 |
| Regular English | 52 | 42 | 48 |
| ELM & Reg. English | 1 | 4 | 4 |
| Two English (Remedial/Regular) | 19 | 37 | 36 |
| Two ELD | 3 | 7 | 6 |
| No English Class | 4 | 6 | 2 |
| Math Courses | % | % | % |
| Geometry | 5 | 1 | 1 |
| Honor Algebra | 14 | 5 | 6 |
| Regular Algebra | 32 | 22 | 25 |
| Basic Algebra | 29 | 38 | 41 |
| Pre-Algebra | 17 | 30 | 25 |
| No Math Course | 3 | 4 | 2 |

*Note*. The full sample represents averages over three school years of complete cohort data for middle school students provided by the school districts. †Classroom counts only includes those with 10 or more students. The second column includes students who were classified ELL anytime between K-7th grade. RK-6 represents students who reclassified in elementary, R7 students who reclassified in 7th grade, and E7 students who remained classified ELL as of 7th grade. The third column represents students who were classified ELL in middle school (about 27% of Manzanita students). The last column represents middle school student's classified ELL, who had both CELDT and CST scores. About 80 ELL and 14 RFEP students did not take any math or English course in 8th grade.

Table 2.2

*Language Classification and English Course Placement Requirements*

| | CELDT (1-5) | CST (1-5) | Reading Level (1-8) | Years in Country |
|---|---|---|---|---|
| Reclassification | | | | |
| RFEP | 4 | 3.5 | --- | --- |
| English courses | | | | |
| Honors English | --- | 4.0 | --- | --- |
| Regular English | --- | 3.0 | --- | --- |
| Intensive Literacy[2] | 4 | 2.5 | 3.5 | >=5 |
| ELD[2] | 3 | 2.0 | --- | <5 |

Note. Students' English course placement was based on CELDT, CST and reading scores, as well as teachers/counselors' recommendations. ELD and intensive literacy courses were provided in a two-period block (denoted by the square). Intensive literacy courses included a combination of a regular English course or ELM course. Students placed in honors English were either identified as GATE or recommended by their teacher.

Table 2.3

*Eighth-Grade Math and English Course Placement by Language Classification*

| | **Language Classification in 7th Grade** | | |
|---|---|---|---|
| | RFEP N=764 | ELL N=2,205 | Chi-Square Test |
| **English Courses** | % | % | |
| Honors English | 13.4 | 0.7 | $X^2$(5,N=2,269)= |
| Regular English | 77.3 | 37.3 | 754.6, *p* <.001) |
| ELM & Reg. English | 0.0 | 5.9 | |
| Two English (Intensive) | 7.5 | 46.3 | |
| Two ELD | 0.0 | 7.4 | |
| No English Class | 1.5 | 2.1 | |
| **Math Courses** | % | % | |
| Geometry | 4.1 | 0.3 | $X^2$(5,N=2,269)= |
| Honor Algebra | 15.8 | 2.3 | 549.7, *p* <.001) |
| Regular Algebra | 43.9 | 18.3 | |
| Basic Algebra | 23.1 | 47.4 | |
| Pre-Algebra | 11.2 | 29.8 | |
| No Math Course | 1.5 | 1.5 | |

*Note*. This represents students who remained ELL versus students who reclassified RFEP in 7th grade. ELL students were overrepresented in "Pre-Algebra" and "Basic Algebra", and RFEP students were overrepresented in "Regular Algebra." About 60% of ELL students were taking either two English mainstream courses or two ELD courses. The rest of the ELL students were enrolled in one English course. In comparison, 91% of RFEP students were enrolled in one English course.

Table 2.4

*Average CELDT and CST scores by Language Classification*

| | **Language Classification in 7th Grade** | | |
|---|---|---|---|
| N (%) | RFEP N=764 | ELL N=2,205 | t-test |
| Prior English Scores | 764 | 2,205 | |
| 7th CELDT $\bar{X}$ (SD) **(248-741) 556 pass** | 613.5 (31.2) | 556.9 (44.6) | $t(2969) = 32.37$, $p < .001$ |
| Prior English Scores | 764 | 2,205 | |
| 6th grade CST ELA $\bar{X}$ (SD) **(150-600) 325 pass** | 354.5 (27.3) | 299.7 (35.5) | $t(2969) = 38.84$, $p < .001$ |
| Prior Math Scores | 763 | 2,191 | |
| 6th grade CST Math $\bar{X}$ (SD) **(150-600)** | 363.2 (58.4) | 298.2 (50.4) | $t(2952) = 29.36$, $p < .001$ |
| Prior Math Scores | 759 | 2,170 | |
| 7th grade CST Math $\bar{X}$ (SD) **(150-600)** | 346.9 (54.9) | 296.8 (49.0) | $t(2927) = 23.41$, $p < .001$ |

N*ote*. ELL students had overall lower CELDT and CST achievement scores than RFEP students.

Table 2.5

*Language Classification Effects Placement in One English Course versus Two English Courses in Eighth Grade*

| 8th Grade English Course 2010-2013 (N=2,910) | | | 2010-2011 | | 2011-2012 | | 2012-2013 | |
|---|---|---|---|---|---|---|---|---|
| OLS | --- | **.19*** (.03)** | --- | **.18** (.05)** | --- | **.15** (.04)** | --- | **.25** (.06)** |
| Bandwidths | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** |
| 1.5 | .39*** (.03) | **.24** (.08)** | .33*** (.04) | **.02 (.17)** | .42*** (.05) | **.19 (.14)** | .48*** (.06) | **.51*** (.13)** |
| 1.0 | .39*** (.03) | **.19† (.10)** | .32*** (.05) | **-.15 (.21)** | .43*** (.06) | **.22 (.16)** | .46*** (.06) | **.48** (.16)** |
| 0.5 | .43*** (.04) | **.16 (.13)** | .34*** (.07) | **-.11 (.29)** | .51*** (.08) | **.12 (.19)** | .46*** (.09) | **.50* (.22)** |

*Note*. Eighth graders (< 2%) who were not placed in an English course were removed from the analyses. The assignment variable was a composite variable that consisted of students' 7th grade CELDT scores and 6th grade CST ELA scores. The treatment (stage 1) was 0 "ELL" versus 1 "RFEP" in 7th grade. The outcome (stage 2) was 1 "One English Course" versus 0 "Two Period English Courses" in 8th grade. The RD bandwidths were based on Nichols (2007) RD Stata package. For the English models the optimal bandwidths were about 1.5, 1.0, and 0.5 standard deviations away from the required CELDT/CST cutoff. I rounded the bandwidths to the nearest tenth but the effect sizes presented were based on the exact bandwidths determined by the Stata program. ± $p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Figure 2.1

*Regression Discontinuity for One English Course versus Two English Courses*

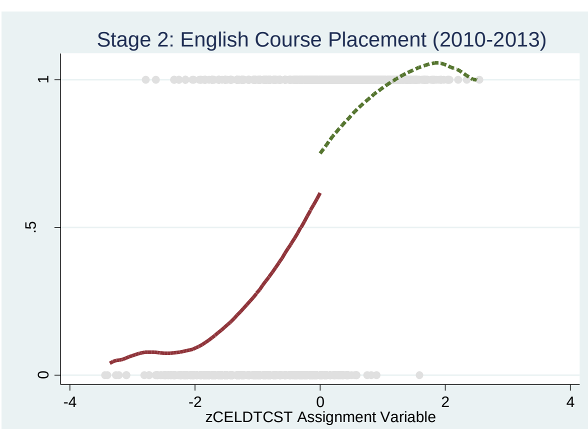


Figure 2.2 *RD for One English Course versus Two English Courses by Cohort*

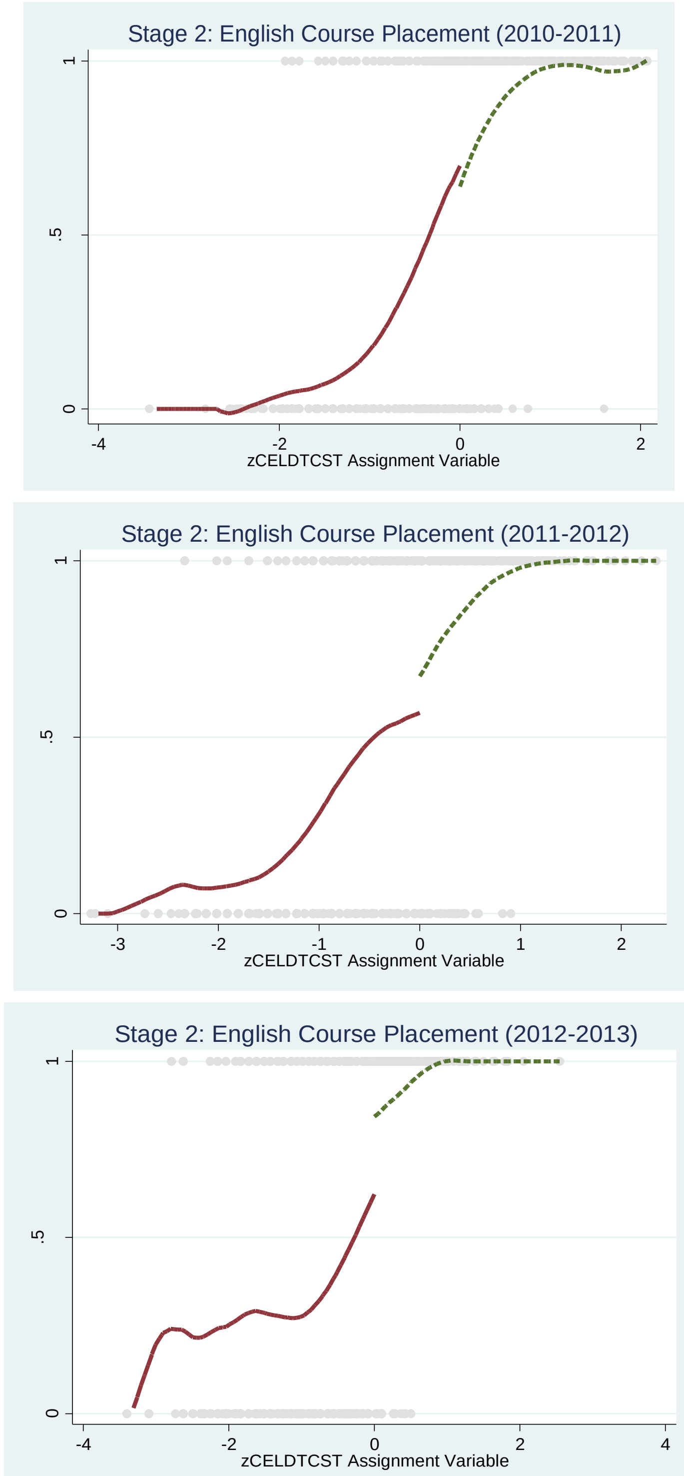


Table 2.6

*Language Classification Effects Placement in Accelerated Math versus Basic Math in Eighth Grade*

| 8$^{th}$ Grade Math Course 2010-2013 (n=2,922) | | | 2010-2011 | | 2011-2012 | | 2012-2013 | |
|---|---|---|---|---|---|---|---|---|
| OLS | --- | **.17$^{***}$ (.03)** | --- | **.15$^{***}$ (.03)** | --- | **.15$^{***}$ (.04)** | --- | **.17$^{***}$ (.05)** |
| Bandwidths | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** |
| 1.5 | .41$^{***}$ (.02) | **.19$^{**}$ (.07)** | .35$^{***}$ (.04) | **.17 (.15)** | .42$^{***}$ (.04) | **.14 (.12)** | .49$^{***}$ (.05) | **.24$^{*}$ (.12)** |
| 1.0 | .40$^{***}$ (.03) | **.18† (.10)** | .32$^{***}$ (.05) | **.05 (.21)** | .43$^{***}$ (.06) | **.10 (.06)** | .47$^{***}$ (.06) | **.26† (.16)** |
| 0.5 | .43$^{***}$ (.04) | **.25$^{*}$ (.13)** | .34$^{***}$ (.07) | **.09 (.29)** | .51$^{***}$ (.08) | **.10 (.17)** | .48$^{***}$ (.09) | **.29 (.22)** |

*Note*. Eighth graders (< 2%) who were not placed in a math course were removed from the analysis. The assignment variable was a composite variable that consist of students' 7$^{th}$ grade CELDT scores and 6$^{th}$ grade CST ELA scores. The treatment (stage 1) is 0 "ELL" versus 1 "RFEP" in 7$^{th}$ grade. The outcome (stage 2) was 1 "Accelerated Math" (i.e., regular algebra, honors algebra and geometry) versus 0 "Basic Math" (i.e. basic algebra and pre-algebra). The RD bandwidths were based on Nichols (2007) RD Stata package. For the English models the optimal bandwidths were about 1.5, 1.0, and 0.5 standard deviations away from the required CELDT/CST cutoff except. I rounded the bandwidths to the nearest tenth but the effect sizes presented were based on the exact bandwidths determined by the Stata program.† $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Figure 2.3

*Regression Discontinuity for Accelerated Math versus Basic Math*

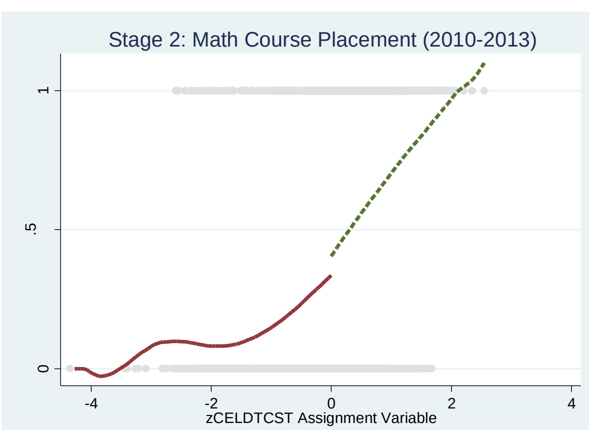


Figure 2.4 RD *for Accelerated Math versus Basic Math by Cohort*

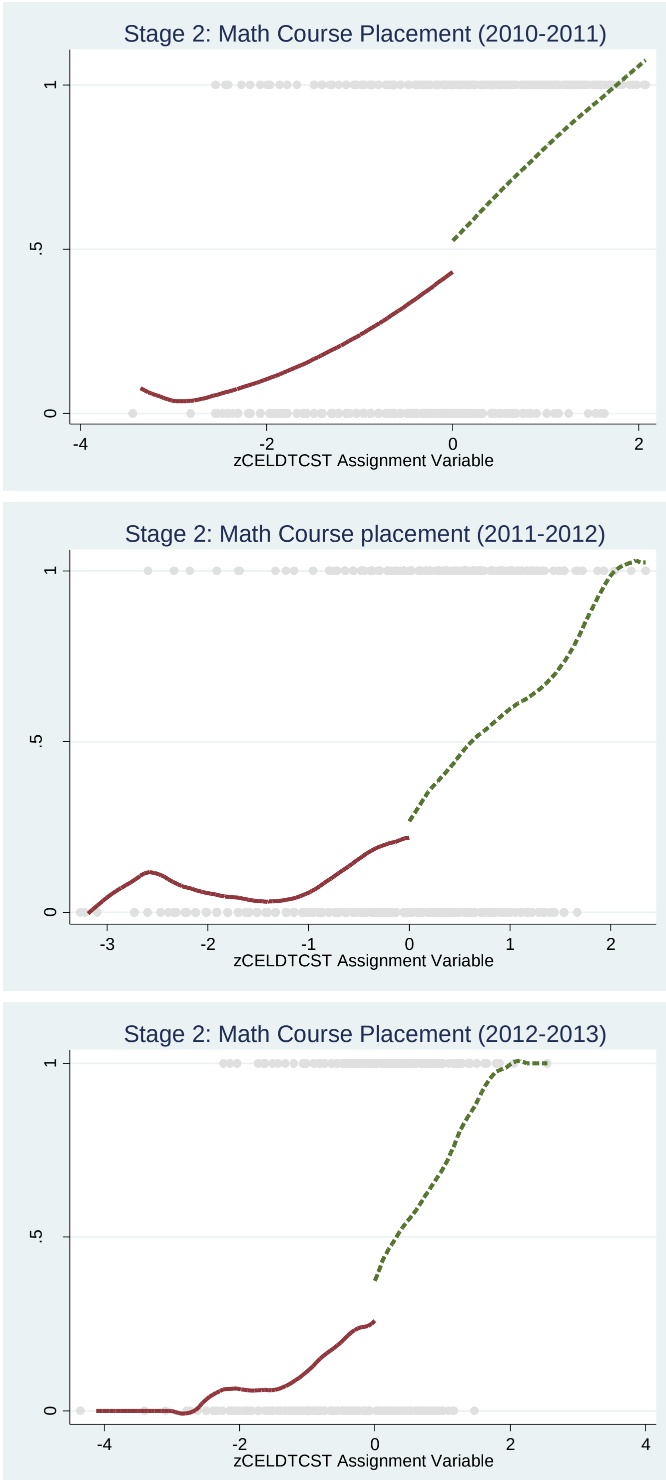


Table 2.7

*Language Classification and Average Peer Achievement in Eighth Grade English Course*

| 8th Grade ELA Course 2010-2013 (N=2,437) | | | 2010-2011 | | 2011-2012 | | 2012-2013 | |
|---|---|---|---|---|---|---|---|---|
| OLS | --- | **.35*** (.04)** | --- | **.24*** (.05)** | --- | **.46*** (.08)** | --- | **.30*** (.07)** |
| Bandwidths | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** |
| 1.5 | .41*** (.03) | **.16 (.14)** | .36*** (.05) | **.01 (.18)** | .43*** (.05) | **-.17 (.24)** | .46*** (.07) | **.47† (.35)** |
| 1.0 | .43*** (.04) | **.05 (.12)** | .36*** (.06) | **.02 (.26)** | .45*** (.06) | **.11 (.29)** | .46*** (.07) | **.25† (.23)** |
| 0.5 | .45*** (.06) | **.12 (.26)** | .35*** (.09) | **.02 (.42)** | .56*** (.08) | **-.04 (.32)** | .43*** (.11) | **.58 (.54)** |

*Note*. Eighth graders (< 2%) who were not placed in an English course were removed from the analyses. The assignment variable was a composite variable that consisted of students' 7th grade CELDT scores and 6th grade CST ELA scores. The treatment (stage 1) was 0 "ELL" versus 1 "RFEP" in 7th grade. The outcome (stage 2) was a continuous variable representing the individual's students average peer achievement score in their 8th grade English course. The RD bandwidths were based on Nichols (2007) RD Stata package. For the English models the optimal bandwidths were about 1.5, 1.0, and 0.5 standard deviations away from the required CELDT/CST cutoff. I rounded the bandwidths to the nearest tenth but the effect sizes presented were based on the exact bandwidths determined by the Stata program. † $p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 2.8

*Language Classification and Average Peer Achievement in Eighth Grade Math Course*

| 8th Grade Math Course 2010-2013 (N=2,449) | | | 2010-2011 | | 2011-2012 | | 2012-2013 | |
|---|---|---|---|---|---|---|---|---|
| OLS | --- | **.22*** (.04)** | --- | **.21*** (.04)** | --- | **.24** (.08)** | --- | **.19*** (.05)** |
| Bandwidths | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** | Stage 1 Coefficient (SE) | **Stage 2 Coefficient (SE)** |
| 1.5 | .41*** (.03) | **.16 (.14)** | .36*** (.06) | **.15 (.24)** | .44*** (.06) | **.25 (.34)** | .45*** (.05) | **.07 (.21)** |
| 1.0 | .44*** (.04) | **.26 (.16)** | .37*** (.08) | **.36 (.31)** | .53*** (.08) | **.36 (.28)** | .45*** (.07) | **.15 (.27)** |
| 0.5 | .45*** (.06) | **.18 (.22)** | .33*** (.11) | **.28 (.48)** | .69*** (.10) | **.21 (.28)** | .47*** (.10) | **.16 (.34)** |

*Note*. Eighth graders (< 2%) who were not placed in a math course were removed from the analysis. The assignment variable was a composite variable that consist of students' 7th grade CELDT scores and 6th grade CST ELA scores. The treatment (stage 1) is 0 "ELL" versus 1 "RFEP" in 7th grade. The outcome (stage 2) was a continuous variable representing the individual's students average peer achievement score in their 8th grade math course. The RD bandwidths were based on Nichols (2007) RD Stata package. For the English models the optimal bandwidths were about 1.5, 1.0, and 0.5 standard deviations away from the required CELDT/CST cutoff. I rounded the bandwidths to the nearest tenth but the effect sizes presented were based on the exact bandwidths determined by the Stata program. .† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Appendix

Appendix Table 2.1A

**Interview Questionnaire**

**Questionnaire for school and district administrators that worked with ELL students particularly those that made language classification decisions.**

Code # of Interviewee:____________________ Date: ________________________________

7) Please describe your job responsibilities.
    a. Probe: What role do you play with ELL students?

**Study 1**

8) Please describe the language classification process at your district particularly for middle school students.
    a. Probe: Is the language classification processes decided at the district level? Can the process differ between schools? If so, what are those differences?
    b. Probe: Which of the following components are considered and to what extent: CELDT, ELA CST, ELA course grade, teacher recommendation, and parent recommendation.
    c. Probe: Will the Smarter Balanced Assessment Consortium (SBAC) be considered similarly to the CST when it comes to language classification? How? (Please provide details).
9) (If applicable) Based on district data it seems that Hispanic and low-income students are overrepresented in the ELL category? Can you describe why you think this may be occurring?

**Study 2**

10) What criteria are used to place ELL and RFEP middle school students in English courses?
    a. Probe: Is English proficiency considered when placing students into regular or advanced English courses? If so, how is it considered? In your experience, what are some reasons why English proficiency is considered?
    b. Probe: (If applicable) Based on district data it seems ELL students are not placed in regular or Honors English courses. Can you describe why you think this may be occurring?
11) What criteria are used to place ELL and RFEP middle school students in math courses?
    a. Probe: Is English proficiency considered when placing students into regular or advanced math courses? If so, how is it considered? In your experience, what are some reasons why English proficiency is considered?
    b. Probe: (If applicable) Based on district data it seems ELL students are not placed in advanced math courses. Can you describe why you think this may be occurring?
12) What types of support services do ELL students receive in middle school for language development and academic achievement?
    a. Probe: For example, are there any services such as 1-1 tutoring, extra ELD course, after school reading program, certified ELL teacher/tutor, etc.?

Appendix Table 2.2A

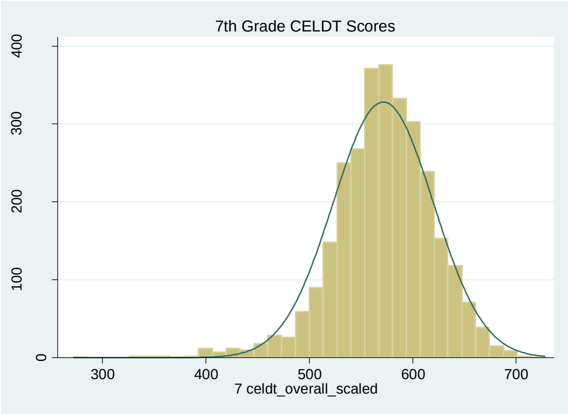
*Required Language Classification Criteria Data Missing*

| | (1) | (2) |
|---|---|---|
| All Criteria Available CELDT and CST | Missing One | Missing All |
| Female | 1.23*** | 1.29* |
| | (0.07) | (0.14) |
| Hispanic (Reference) | | |
| Asian | 0.83 | 1.01 |
| | (0.16) | (0.26) |
| Other | 0.84 | 1.39 |
| | (0.16) | (0.54) |
| Born in the United States | 0.71*** | 1.01 |
| | (0.07) | (0.16) |
| Free and Reduced Lunch | 0.71 | 0.57 |
| | (0.16) | (0.17) |
| Special Education | 46.41*** | 17.38*** |
| | (13.87) | (6.22) |
| Prior CELDT unavailable† | 3.58*** | 21.77*** |
| | (0.45) | (12.33) |
| Prior CST unavailable† | 15.70*** | 118.95*** |
| | (2.62) | (52.52) |
| *N* | | 4231 |
| df_m | | 6.00 |
| pr2 | | .47 |

*Note*. The logistic regression included students' cohort and school fixed effects. The outcome included three categories for 7th graders: 1) all criteria available (reference group) means they had CELDT/CST scores, 2) student had CELDT or CST scores, or 3) the student did not have any scores. † For 7th graders prior scores were 6th grade CELDT and 5th grade CST. Special education students and students who were missing the prior year's scores were more likely to be missing one or all reclassification criteria. Furthermore, students born in the United States were less likely to be missing one reclassification criteria.

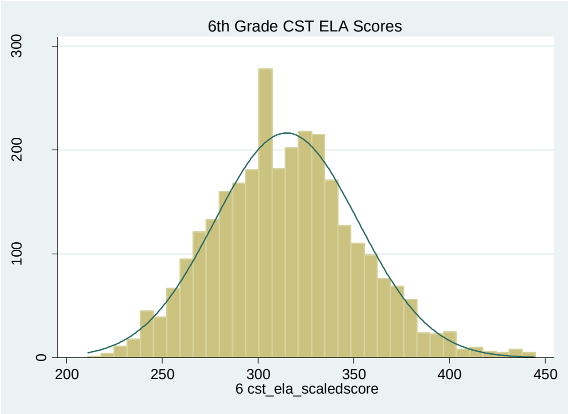
Appendix Figure 2.1A

*CELDT Scores were Normally Distributed (Passing 556 +)*



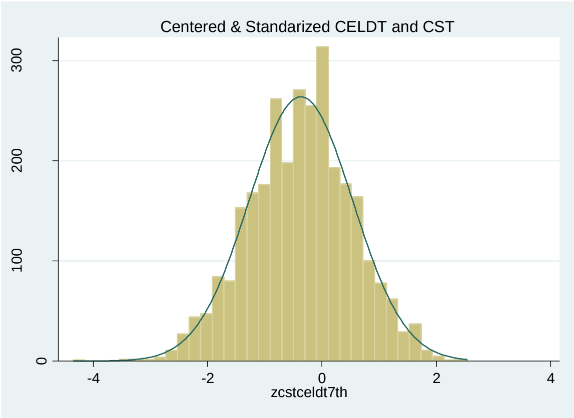

Appendix Figure 2.2A

*CST ELA Scores were Normally Distributed (Passing 325+)*




Appendix Figure 2.3A

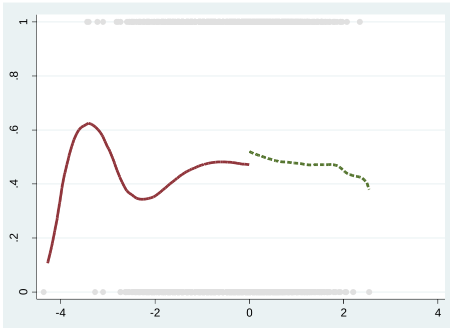
*CELDT and CST ELA (Centered at 556 and 325 and Standardized) Scores were Normally Distributed*



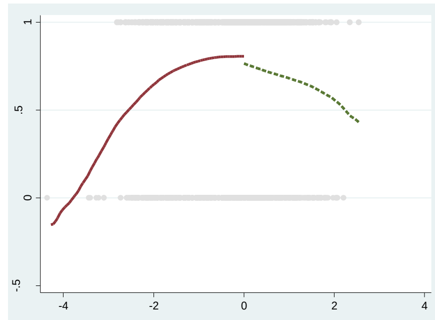
Appendix Figure 2.4A

*Student Demographics by RD Assignment Variable*

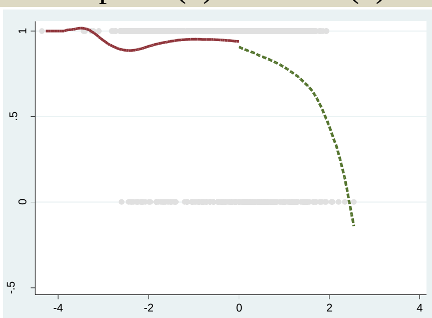
Female (1) vs. Male (0)



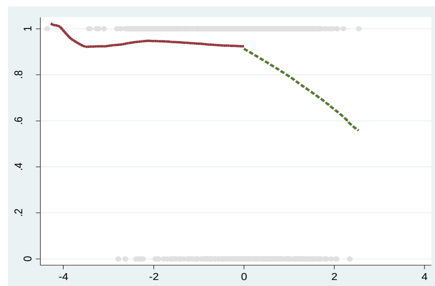

US (1) vs. Foreign Born (0)




Hispanic (1) vs. Asian (0)

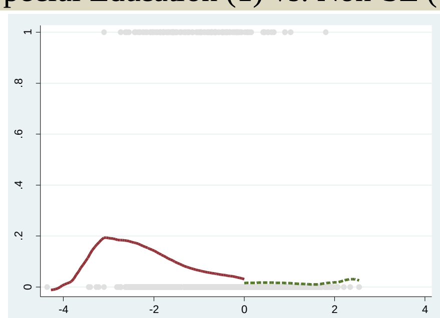



FRL (1) vs. Non-FRL (0)

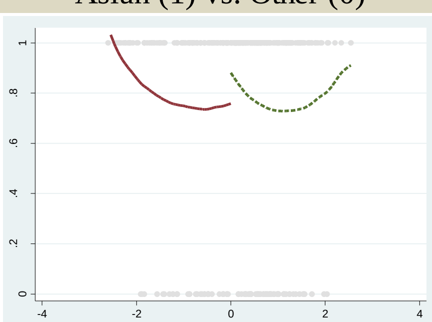



Hispanic (1) vs. Other (0)


Special Education (1) vs. Non-SE (0)



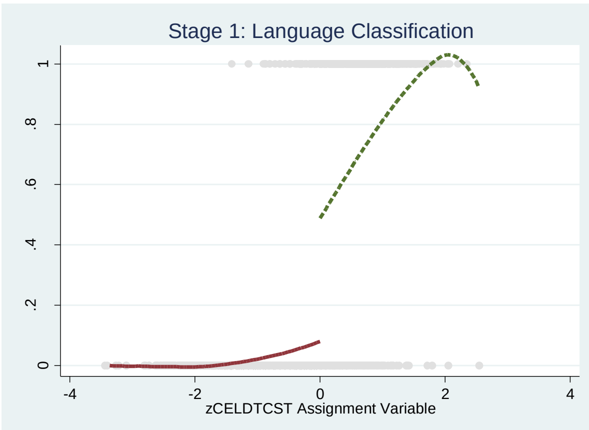

Asian (1) vs. Other (0)




*Note*. The X-axis represents the assignment variable based on students' 7th grade CELDT scores and 6th grade CST ELA scores. The Y-axis represents the binary outcome for each of the students' demographics described in the heading (e.g., Female (1) versus Male (0)). There were no jumps at the assignment variable cutoff based on student demographics.

Appendix Table 2.5A

*RD Stage 1 Language Classification (Treatment)*



# CHAPTER 3

Middle School Language Classification Effects on High School Achievement and Behavioral Outcomes


## Abstract

English Language Learners (ELLs) are students who speak another language at home and who have not yet reached full English proficiency. They are among the lowest performers on a broad range of educational outcomes even when they are compared with Reclassified Fluent English Proficient (RFEP), former ELL students. I use data from one Southern California district to examine how classification can influence students' high school achievement (i.e., their English and math scores and the highest levels of courses they complete) and behavioral outcomes (i.e., attendance and suspensions). The OLS estimates coincide with previous studies, demonstrating that students who become RFEP in middle school have higher California Standards Test (CST) in English Language Arts (ELA), and California High School Exit Examination (CAHSEE) ELA and math scores. Additionally, RFEP students are more likely to be placed in more advanced math courses in high school than the ELL students. Furthermore, RFEP students are less likely to be absent and have less on-campus suspensions than ELL students. However, the regression discontinuity models show, in most instances, academic and behavioral differences between ELL and students who RFEP in middle school are spurious and not due to language classification itself. Only in a few instances do differences exist, and RFEP students are less likely to pass the CHASEE ELA portion and more likely to be suspended compared with ELL students.



*Keywords:* English Language Learners, Secondary School, CAHSEE, Math Course Placement, Behavioral Outcome, and Regression Discontinuity

In California, 25% of students in K-12 schools are English Language Learners (ELLs) (Hill, Betts, Chavez, Zau, & Bachofer, 2014), compared with only 11% of K-12 students nationally (Kohler & Lazarín, 2007). ELLs are students who speak another language at home and who have not yet reached full English proficiency (Hahnel, Wolf, Banks, & LaFors, 2014). They are one of the fastest growing student groups in the country, yet they are also among the lowest performers on a broad range of educational outcomes (Capps et al., 2005; Maxwell, 2014). ELL students, on average, score lower than non-ELL students in English reading, writing, and comprehension, as well as in less language-intensive subject areas, such as mathematics (Edwards, Leichty, & Wilson, 2008; Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hampden-Thompson, Mulligan, Kinukawa, & Halle, 2008).The non-ELL category includes Reclassified Fluent English Proficient (RFEP) students who are former ELL students that, after passing the language classification assessments, are now considered English proficient. The ELL classification itself may have unintended consequences if classified students do not have the opportunity to learn rigorous educational content, or if they are not integrated with their non-ELL peers.[22] This can decrease ELL student engagement, and can affect their attendance and suspension rates (Losen & Martinez, 2013; Uriarte, Lavan, Agusti, & Karp, 2009). It is unclear whether this group's language skills, the stigma associated with the ELL label, or a lack of access to rigorous courses drives the achievement gap between ELL and non-ELL students.

This study specifically examines how language classification by the end of middle school affects high school student achievement and behavioral outcomes. My work focuses on middle school because it is a significant schooling stage that often determines the educational foundation of an adolescent's high school experience (Walqui et al., 2010). Middle school ELL students are

[22] Zaragoza-Petty and Zarate (2014) describes "opportunities to learn" as the school processes that shape and contribute to students' learning. The focus is on schools' practices that influence students' achievement. A thorough explanation of "opportunities to learn" is provided in the theoretical section.

also less frequently studied than ELL elementary students, and yet middle school students are more likely to be Long-Term English Language Learners (LTELL), and foreign-born children (Hahnel, Wolf, Banks, & LaFors, 2014; Olsen, 2010). Middle school children have less time to acquire both English proficiency and the rigorous academic skills needed to get ready for high school and post-secondary education. Of course, ELL middle school students are a diverse and complex group in terms of English proficiency, national origin, socioeconomic status, previous schooling, and the number of years in the U.S. system (Callahan, 2005; Krashen & Brown, 2005). For this reason, there are many aspects to consider when discussing ELL students' academic needs and how educators can improve their academic and behavioral outcomes. Specifically, I address two research questions:

1) How does language classification (ELL and RFEP) by the end of middle school affect the high school students' English and math achievement outcomes (i.e., assessments and course placement)? Further, how does middle school course placement moderate the association between language classification and achievement?
2) How does language classification (ELL and RFEP) by the end of middle school affect high school students' behavioral outcomes (i.e., attendance and suspensions)?

To investigate the specific situation of middle school learners, I use data from one Southern California district to examine how classification can influence students' high school achievement (i.e., their English and math scores and the highest levels of courses they complete), and behavioral outcomes (i.e., attendance and suspensions). I examine the years prior to California's recent implementation of the Common Core State Standards (CCSS), the Smarter Balanced Assessment Consortium (SBAC), and the increase in per-pupil funding for ELL students, each of which will almost certainly affect future ELL policies and practices (Hill, Weston, & Hayes, 2014; Umansky et al., 2015). In the discussion sections of this work, I will address the implications of my findings as they pertain to the new policy changes.

## Literature Review

### Language Minorities and Language Classification Policies

During initial school registration, which usually occurs in kindergarten, California public schools administer the Home Language Survey, which asks parents whether a language other than English is spoken at home. If the answer is no, their children are classified as English only (EO). If the answer is yes, their children must take the California English Language Development Test (CELDT),[23] which assesses their children's English proficiency. Students who pass the CELDT the first time are identified as Initially Fluent English Proficient (IFEP), and those who do not pass are identified as ELLs (Edwards et al., 2008). The California Department of Education (CDE) also has required that districts use the California Standards Test in English Language Arts (CST ELA) to determine the initial classification for students who arrive in California schools in grades three and above.[24] The CELDT and CST ELA scores have also been used to reclassify ELL students as RFEP. For a student to be reclassified, the CDE requires students to score "intermediate" or higher (at least 3 out of 5) in the listening, speaking, reading, and writing subcategories of the CELDT; "early advanced" or higher (at least 4 out of 5) overall on the CELDT; and "basic" or higher (at least 3 out of 5) on the CST ELA.[25] Classification decision-makers can consider teacher and parent recommendations as well, but the extent to which these recommendations are incorporated depends on the district. About 90% of California's districts set even higher reclassification requirements than those set by the state (Hill, Betts, et al., 2014; Hill, Weston, et al., 2014). To reclassify, ELL students must meet the

[23] In the near future, English Language Proficiency Assessment for California (ELPAC) will replace the CELDT (Umansky et al., 2015).
[24] Starting in the 2014-2015 school year, the Smarter Balanced Assessment Consortium (SBAC) replaced the CST. This change has implications for future ELL students' reclassification process.
[25] The CDE requires that students with disabilities to take the CELDT but may receive accommodations based on their individualized education programs (http://www.cde.ca.gov/ta/tg/el/celdtfaq.asp). The CDE only exempts students with significant cognitive disabilities from the CST (http://www.cde.ca.gov/ta/tg/sr/cefstar.asp).

district's higher requirements. Additionally, ELL students have one opportunity, each school year, to become RFEP.

**Language Minorities, English Proficiency, and Achievement Outcomes**

Using different measures, ELL K-12th graders have lower math and reading scores than their non-ELL peers (i.e., EO, IFEP, and RFEP) even when including several controls (Edwards et al., 2008; Gandara & Rumberger, 2009; Grissom, 2004; Mosqueda & Maldonado, 2013; Saunders & Marcelletti, 2012; Slama, 2014). For example, ELL students have lower CST, California High School Exit Examination (CAHSEE) and SAT 9 Reading scores than their non-ELL peers (Gandara et al., 2003; Hill, Betts, et al., 2014). IFEP, on the other hand, outperform their EO and RFEP peers on many achievement measures (i.e., CST ELA, SAT 9) (Edwards et al., 2008; Saunders & Marcelletti, 2012). RFEP student performance relative to EO peers is much more complex. Based on the SAT 9 test, RFEP students have comparable English proficiency as EO speakers, but start to fall behind by fifth grade and never catch up (Gandara et al., 2003). However, other studies demonstrate RFEP students outperform EO students in high school achievement tests (CST, CAHSEE), and have better on-time graduation progress (Hill, Betts, et al., 2014). Slama's (2014) longitudinal study also shows that 63% of RFEP fifth graders scored proficient on the math test compared with 54% of EO students; however the achievement gap reverses by seventh grade (55% EO and 46% RFEP). Achievement outcomes between RFEP and EO students can vary based on who is included in the RFEP category. ELL students have the opportunity to become RFEP each school year; therefore, RFEPs are former ELL students, and the student body composition changes each year.

Researchers are interested in whether earlier reclassification leads to better outcomes for ELL students. Students who reclassify early in elementary school have very strong academic outcomes throughout middle and high school compared with EO and students who remained classified ELL (Halle, Hair, Wandner, McNamara, & Chien, 2012; Hill, Betts, et al., 2014). Based on ECLS-K Item Response Theory scaled scores, ELLs who are reclassified at the beginning of first grade initially scored the same in reading ($\beta$ = -0.31, *NS*) and math ($\beta$ = -0.03, *NS*) compared with EO, and they have moderately steep growth trajectory from kindergarten to eighth grade in reading ($\beta$ = 0.90, $p$ <.001) and math ($\beta$ = 0.75, $p$ <.001) (Halle et al., 2012). ELLs who are not reclassified by first grade have a large initial gap in reading ($\beta$ = -45.71, $p$ <.001) and a smaller initial gap in math ($\beta$ = -1.96, $p$ <.001) compared with EO, and substantially did worse in math, but not in reading, overtime. The math rate growth between kindergarten and eighth grade is worse over time ($\beta$ = -1.30, $p$ <.001), but the growth in reading is significantly steeper ($\beta$ = 10.44, $p$ <.001) than EO students' growth. Furthermore, students who are proficient before they start kindergarten had similar behavior outcomes (e.g., externalizing, self-control, approaches to learning) as EO students; however, ELL students who are not proficient by first grade scored lower on approaches to learning (Halle et al., 2012). Hill, Betts, et al. (2014) conducted a similar comparison with a smaller dataset using district data from the Los Angeles Unified School District (LAUSD) and the San Diego Unified School District (SDUSD). They concluded that students who are reclassified in second or third grade have better English and math achievement scores and graduation rates than those reclassified in fifth grade. Both of these studies show that earlier reclassification leads to better achievement outcomes for students, however, neither used rigorous methods to determine whether it is reclassification itself or other confounders that explain early classification and future achievement. The researchers used non-randomized data, and there can be unmeasured cognitive factors that can explain why a student who reclassified early has higher achievement scores. Most current research has not disentangled the effects of English proficiency, opportunities to learn, and other unmeasured factors (e.g., motivation, cognitive abilities, and parent expectations) that may be driving student achievement and behavioral outcomes.

To my knowledge, there are only two studies that attempt to estimate the effects of language classification. Both studies use observational data and regression discontinuity (RD) models to make causal inferences. Robinson (2011) finds that in high school RFEP Latino and Asian American students score lower than ELL high school students on the CST ELA exam, but there is no difference for elementary or middle school students [Data from 2001-02 and 2006-07]. He explains that RFEP students' lower English scores may be caused by their removal from English language development courses and their move back into mainstream English courses with no support. The same study shows language classification did not affect high school course placement in English, math, science, and other college preparatory courses, nor attendance. This study used a frontier RD approach where their analyses only included students who have passed the CELDT, and they used the CST ELA as the assignment variable. The Robinson-Cimpian and Thompson (2015) RD model, based on LAUSD data, shows that making it more difficult to reclassify increases high school Latino students CST ELA scores (.18SD), but there is no effect on middle school students' CST-ELA scores [Data from 2004-06 and 2007-09]. Also, high school graduation rates increased by 11% percentage points. This study used "difference-in-regression discontinuities" approach that included regression discontinuity design, instrumental variables, difference-in-difference, and inverse probability weighting.

**Contributions to the Current Literature**

RFEP students outperforming ELL students on every measure examined can suggest that the criteria used to determine when an ELL student no longer needs support to learn English separates those who are stronger academic performers from those who are less able. At the same time, not all RFEPs are equal: those who are reclassified in earlier grades are more likely to progress on time, have higher test scores, and have other positive outcomes in their final years of high school (Halle et al., 2012; Hill, Betts, et al., 2014). These results suggest that cross-sectional views of RFEP progress over time are complicated by the fact that the grade a student is reclassified in matters. The work of Halle et al. (2012), and Hill, Betts, et al. (2014) focuses on the effects of reclassifying in different elementary school grade levels. It is for this reason I examine the effects of reclassification particularly for middle school students. I compare current ELL students with students who had only been recently reclassified in middle school, either in seventh or eighth grade. Middle school ELL students are more academically similar to students who RFEP in middle school than to EO, IFEP, or students who RFEP in elementary school.

Furthermore, language classification policies, and the implementation of those policies can vary by district. The studies by Robinson (2011), and Robinson-Cimpian and Thompson (2015) focused on two Southern California districts that followed the minimum state requirements of 556 for the CELDT and 300 for the CST. In their districts, the reclassification rates for middle school students was 70% above the cutoff, but with an almost zero probability below the set cutoff. The current study will thus shed light on the effects of language classification in a district where the cutoffs are higher than the state requirements (325 on the CST), and where administrators consider other factors that cannot be quantified when determining a student's classification (e.g., parent and teacher recommendations). When estimating effects, it is important to specify the district's language classification, as using stricter reclassification criteria than those suggested by the state guidelines is also associated with slightly better outcomes for RFEP students (Hill, Weston, et al., 2014). For example, the work of

Hill, Weston, et al. (2014) shows stricter criteria are associated with a greater likelihood of on-time grade progress among students reclassified in the eighth grade. Robinson-Cimpian and Thompson (2015) RD models also demonstrate that more difficult reclassification requirements increases high school students CST ELA scores, but it does not make a difference for elementary or middle school students. I expect that students who become RFEP in middle school will have better high school achievement outcomes than students who remained classified ELL, especially when these students attend school in a district that sets higher reclassification requirements than the state mandates.

Only a few studies have examined language minorities' behavioral outcomes (Halle et al., 2012; Hill, Weston, et al., 2014; Losen & Martinez, 2013; Robinson, 2011; Uriarte et al., 2009). I examine whether language classification by the end of middle school affects student attendance and suspension rates once they reach high school. ELL students are more likely to be placed in less rigorous courses separate from RFEP peers (Callahan, 2005; Callahan, Wilkinson, & Muller, 2010; Gandara et al., 2003; Hahnel et al., 2014; Mayer, 2008), which might lead students to disengage from school. Studies have demonstrated that ELL students are more likely to be absent and suspended (Losen & Martinez, 2013; Uriarte et al., 2009). Furthermore, high school teachers and school administrators may treat students who are reclassified differently than those classified ELL. Currently only one study shows language classification did not predict attendance rates for ELL high school students (Robinson, 2011). Unfortunately, it is unclear whether the study considered the grades in which the students were reclassified and whether RFEP students were combined. As previously explained, I compare current ELL students with students who reclassified in seventh and eighth grade only. I expect that students who become RFEP in middle school will have greater high school attendance and suspension outcomes than students who remained classified ELL.

## Theoretical Framework

Unfortunately, even when the intention is to provide greater opportunities, societal institutions—including educational institutions—can reinforce and reproduce inequalities. In fact, prior research shows that student background characteristics determine other types of academic sorting (e.g., course placement and acceptance into the Gifted and Talented Education Program [GATE]) despite students' prior achievement (Schneider, Swanson, & Riegle-Crumb, 1997; Stein, Hetzel, & Beck, 2011; Wang & Goldschmidt, 2003). Language classification can become a self-fulfilling prophecy, resulting from the label itself, and it can determine many students' academic trajectories and behavioral outcomes. Although language classification is supposed to be used to distinguish students' needs and to provide them adequate support, classifications, categories, and labels can have unintended consequences when they become indicators of the students' "abilities" instead of their actual needs (Lovaglia, Lucas, & Thye, 1998). ELL students may become discouraged to attend school and teachers may perceive and treat ELL students differently and in turn, this can lead to behavioral problems. Unfortunately, some groups become known as high-achievers while others are termed low-achievers, thus creating hierarchies and student sorting in schools. It is important to determine whether language classification creates unequal access to opportunities to learn and therefore RFEP students have higher achievement and behavioral outcomes.

In this study, I estimate the effects of language classification itself on students' high school achievement (i.e., assessments and course placement), and behavior (i.e., attendance and suspensions). I take into consideration the fact that elementary and high school ELL students tend to be placed in classes separate from their non-ELL peers, and that these classes tend to be less rigorous (Callahan, 2005; Callahan et al., 2010; Gandara et al., 2003; Hahnel et al., 2014; Mayer, 2008). Although we know less about middle school student course placement, the potential inequalities in student opportunity to learn may compound certain disadvantages facing ELL students once they reach high school. Students begin to be tracked in middle school, which determines their high school outcomes (Walqui et al., 2010). It is unclear whether ELL students' academic skills or a lack of access to rigorous courses drives the achievement gap between ELL and RFEP students. In my work, I consider the types of English and math courses these students received in middle school, and student language classification as a way to account for opportunity to learn. I use the regression discontinuity design to determine if language classification itself determines students' achievement and behavior instead of other unmeasured factors. Ultimately, my interest is to understand how middle school language classification influences high school English and math achievement levels while taking into account course placement, as well as the effects on student attendance and suspensions.

I hypothesize that middle school ELL students will have lower English and math achievement scores than comparable peers who reclassified RFEP by eighth grade, net of skills. Furthermore, these RFEP students are more likely to be placed in advanced math courses than their ELL counterparts. Lastly, ELL students are more likely to be disengaged and treated differently by educators; therefore, they are more prone to absences and suspensions.

## Methods

### Data Source and Sample

I had access to district data through the Spencer-funded Evaluating the Quality of Universal Algebra Learning (EQUAL) project. I focused on middle school students from one diverse southern California school districts with eight separate middle schools.[26] I followed three $8^{th}$ grade cohorts until they reached the $11^{th}$ grade. I chose to focus my investigation on the district I refer to pseudonymously as Manzanita Unified School District (MUSD) because it had a large percentage of ELLs, and more specifically, a large number of Hispanic and Asian students. Furthermore, this district provided insight into practices that prevail in a relatively low-income community. The project provided quantitative data that included student-level demographic, attendance, suspension, middle school transcript, high school transcript and achievement data (e.g., CELDT, CST, CAHSEE) from district administration records. This data make it possible to measure a variety of school practices related to achievement and behavioral outcomes.

The full sample (n=16,144) included three cohorts of eight graders from 2010-11, 2011-12 and 2012-13 (see Table 3.1). The selected sample excluded both White and African American students, who were mostly, English Only (EO) students. And it excluded IFEP students and students who reclassified RFEP in elementary school. My main analyses included the 26% of the students who were classified ELL when they began middle school. About 75% of those students were Long-Term English Language Learners (LTELL) born in the United States. The remaining 25% of students were born in another country and can either be LTELL or recently arrived immigrants. I further restricted the data to students who had both CELDT and CST scores requirements to reclassify, referring to them as the final sample. The selected sample (n=4,231) were different from the students in final sample (n=2,969). The complete sample had fewer special education, foreign-born, and Asian American students than the final sample (see Table 3.1 in the Appendix). Therefore, the ordinary least squares and final regression discontinuity

[26] This district had a total of 13 schools but the five schools that were removed were non-traditional schools. Four schools were for academically struggling students and one was for gifted students.

results are generalized to middle school language minorities who had both CELDT and CST scores—the main reclassification criteria.

**Measures**

Students' language classification by the end of middle school may influence their high school CST, CAHSEE, math course placement, attendance and suspensions.

**CST ELA.** The CSTs measures students' performance in relation to the California content standards. The present study focuses on the CST English portion for students in ninth and tenth grade. Students' scores were categorized into advanced, proficient, basic, below basic, or far below basic on a scale from 150 to 600, with proficient being a score greater than or equal to 350. In this study, only the first two cohorts had available CST ELA scores for ninth graders and only the first cohort had available scores for tenth graders.

**CAHSEE.** The CAHSEE has both an English and math component. Each section score ranges from 275 to 450, and students must achieve a minimum score of 350 on each section to pass and graduate with a high school diploma. All tenth graders are required to take the CAHSEE, and only those students who fail the exam are required to take it in later years. In this study, only the first two cohorts had available CAHSEE scores for tenth graders.

**Math course placement.** Ninth graders were placed into geometry, honors algebra, regular algebra, or basic algebra. These math courses were grouped into two categories: "accelerated" math course (i.e., honors algebra, geometry) and "below basic" math course (i.e., regular algebra, pre-algebra). Here, I consider ninth grade regular algebra a below basic course, because on track students complete algebra in eight grade (Domina, McEachin, Penner, & Penner, 2015).

Eleventh graders were placed into trigonometry, algebra II, geometry, or algebra I. Again, math courses were grouped into two categories: “advanced” math course (i.e., algebra II, trigonometry) and “non-advanced” math course (i.e., geometry, algebra I). Only the first cohort had eleventh-grade math course placement. According to Long (2009), algebra II is the minimum that students must complete to be college ready.

**Attendance.** High school students may be absent from 0-179 days in a given school year. The number of absences was counted separately for ninth, tenth, and eleventh graders. Ninth grade analyses included the three cohorts, tenth grade analyses included the first two cohorts, and eleventh grade analyses only included the first cohort data.

**Suspensions.** Students may be suspended on-campus or off-campus. Student’s suspended on-campus was required to come to school but could not enter their regular classroom for 1-32 days. Student’s suspended off-campus was required to stay home for 1-26 days. The number of suspensions was counted separately for ninth, tenth, and eleventh graders. Ninth grade analyses included the three cohorts, tenth grade analyses included the first two cohorts, and eleventh grade analyses only included the first cohort data.

**Language classification.** Student’s language classification by eighth grade may influence their high school achievement, course placement, and behavioral outcomes. Therefore, each student’s language classification was coded 0 for ELL and 1 for RFEP. I compared eight graders who remained ELL versus those students who were reclassified RFEP in the seventh or eighth grade. Seventh grade language classification was based on students seventh grade CELDT scores and sixth grade CST ELA scores. Eight grade language classification was based off students eight grade CELDT scores and seventh grade CST ELA scores. The CELDT ranged from 248-741, and to reclassify MUSD required a score of at least 556.[27] The CST raw scores ranged between 150 and 600, and MUSD required students to meet a score of at least 325, considered “mid-basic.”

**Control variables.** The models also included student-level covariates to explain differences in the students’ language classification. These covariates included gender (1=female, 0=male), race/ethnicity, birth country, socioeconomic status (SES), and special education status. Racial/ethnic categories included Hispanic (reference group), Asian American, and an “other race” category included American Indian, Alaskan Native, Filipino, Native Hawaiian, and Pacific Islanders. The birth country was a binary outcome, where 1 was coded for those “born in the United States,” and 0 was coded for those “born in another country.” SES was based on students’ “free or reduced lunch” (FRL) status, where students who qualified for FRL (reference group) were compared to students who did not qualify for FRL status. Special education status was binary, where 1 was coded “special education,” and 0 was coded “no special education.”

## Analyses

I used student-level data to understand whether language classification by the end of middle school predicts students’ achievement and behavioral outcomes. I provided ordinary least squares (OLS) effect sizes to compare it to the regression discontinuity (RD) effect sizes. I explain both methods in the following two sections.

## Language Classification Association with Achievement and Behavioral Outcomes

To address how middle school language classification (i.e., ELL and RFEP) affects high school English and math achievement outcomes (Research Question 1), I conducted OLS and

[27] The CELDT subcategories’ raw scores and pass cutoffs also varied. The scores for English language listening ranged from 230-715, speaking 225-720, reading 320-750, and writing 220-780. Manzanita seventh graders were required to score at least a 495, 476, 529, and 508 in listening, speaking, reading, and writing, respectively.[27] Eighth graders had to score at least a 508, 480, 543, and 511 in listening, speaking, reading, and writing, respectively.

multinomial logit models (MLOGIT). First, I ran separate OLS models to estimate the effects of student's language classification by the end of middle school on their CST ELA, CAHSEE ELA and math scores, accounting for the student's race, gender, birth country, free or reduced lunch, special education, cohort, prior achievement, and teacher fixed effects. These models can be expressed as:

$$Y_{it} = \beta_0 + \beta_1\, LC_i + \beta_2\, Demo_i + \beta_3\, Cohort_i + \beta_4\, Prior_i + \mathrm{F}\delta_{s(i)} + e_i\,, \quad \text{(F1)}$$

In (F1), $Y_{it}$ represents student *i*'s CST ELA in the spring of ninth or tenth grade, or student *i*'s CAHSEE ELA or math score in the spring of tenth grade; $LC_i$ represents whether a student reclassified RFEP by the eighth grade or remained ELL; $Demo_i$ stands for the student's characteristics, such as race, gender, country of birth, FRL, and special education status; and $\text{Cohort}_i$ is the school year in which the student was an eighth-grade student. $Prior_i$ is students eighth-grade English and math course, and students eighth grade CST scores. Additionally school fixed effects $\mathrm{F}\delta_{s(i)}$ were also included to control for annual changes and school factors (e.g., other unmeasured confounders) that might also have explained achievement or behavioral outcomes using Stata's *xtlogit,fe* command (StataCorp, 2011). The betas (β) represent the estimated increase in the outcome per unit increase in the value for each given covariate variable. In particular, $\beta_1$ represents the increase in the student's CST or CAHSEE score by the student's language classification.

Furthermore, to address the first research question sub-question, I also ran the main OLS models and included eighth-grade English and math course placement as moderators in determining the extent to which the relationship between the student's language classification and their achievement varies by course placement. A moderator provides information regarding under which conditions an interaction occur (Baron & Kenney, 1986). I expect the association between language classification and achievement to vary by course placement because different English and math courses provide different levels of rigor.

Second, I ran two multi-nominal logistic regression models to estimate the odds ratio of a high school student being placed in a higher-level math course based on their language classification and math course placement in middle school. Ninth graders were enrolled in basic algebra, honors algebra, or geometry, versus regular algebra (reference group). Eleventh graders were placed into algebra I, algebra II, or trigonometry, versus geometry (reference group). These models can be expressed as:

$$\ln\left[\frac{p\,(Acc)}{1-p(Acc)i}\right] = \beta_0 + \beta_1\, LC_i + \beta_2\, Demo_i + \beta_3\, Cohort_i + \beta_4 Prior_i + \mathrm{F}\delta_{s(i)} + e_i \quad \text{(F2)}$$

In (F2), $\ln\left[\frac{p\,(Acc)}{1-p\,(Acc)i}\right]$ is a variable that represents student *i*'s log odds of taking "geometry" (1), "honor algebra" (2), "regular algebra" (3), or basic algebra (4) in the ninth grade. For eleventh graders, this variable represents "trigonometry" (1), "algebra II" (2), geometry" (3), or "algebra I" (4). The other components of the model (e.g., $LC_i$, Cohort$_{i,}$) are similar to those described in the (F1) model. The betas (β) represent the estimated increase in the log odds of the outcome per unit increase in the value for each given covariate variable. In particular, $\beta_1$ represents the increase in log odds by the student's language classification. I also ran these two models with eighth-grade English and math course placement as a moderator to determine the extent to which the relationship between language classification and high school course placement varied.

Third, I ran three separate OLS models to reveal how a student's language classification in middle school is associated with their number of absences (Research Question 2) in the ninth, tenth, and eleventh grades while also accounting for the student's race, gender, birth country, free or reduced lunch, special education, cohort, and school fixed effects. These models were similar to (F1), except for the outcome variable $Y_{it}$ represents student $i$'s attendance.

Fourth, I ran six separate OLS models to reveal how a student's language classification in middle school is associated with their number of on- and off-campus suspensions (Research Question 2) in the ninth, tenth, and eleventh grades while also accounting for the student's race, gender, birth country, free or reduced lunch, special education, cohort, and school fixed effects. These models were similar to (F1), except for the outcome variable $Y_{it}$ represents student $i$'s on- or off-campus suspensions.

**Estimating the Effects of Language Classification Assignment**

I conducted RD models similar to the prior OLS and multi-nominal logistic regression models to obtain less biased effect sizes estimates. RD design can provide *causal inferences* that are "as good as random assignment" and it has strong internal validity (Lee & Lemieux, 2010). For the RD approach to yield valid causal inferences, we must meet four key assumptions. First, the treatment must be endogenous. I exploited the fact that the Manzanita district chooses a cutoff based on the CELDT and CST. Here, language classification 0 for ELL and 1 for RFEP (the treatment) was determined by a set formula based on the CELDT and CST scores (the assignment variable). Second, the students and teachers must not be able to manipulate the assignment variable (Imbens & Lemieux, 2008). As the CELDT and CST tests are based on multiple questions, it would be difficult, if not impossible, to manipulate them to be either right above or right below the cutoff. This allowed me to use an RD to compare the outcomes of the students just above and just below the threshold. These two groups of students were nearly identical in all ways, with the exception that the former group was recommended to remain ELL while the latter was recommended to RFEP. Arguably the students near the cutoff also had similar English proficiency because the 556 CELDT and 325 CST are arbitrary cutoffs. Third, the CELDT and CST must be normally distributed and it may not have a jump at the threshold (see Appendix Figures 3.1A-3.3A). Fourth, there cannot be a discontinuity in covariates (e.g., race, gender). Therefore, I regressed each covariate on binary variables with the combined CELDT and CST assignment variable (see Appendix Figure 3.4A).

For middle school ELL students to be reclassified, they were required to pass the CELDT and CST with their respective different cutoff scores. Failing to meet even one of the requirements could have been enough to prevent a student from being reclassified. For the RD model, I created one assignment variable CELDT/CST that included a combination of the two different requirements for seventh and eighth graders. This reduced the dimensionality to one composite score (Reardon & Robinson, 2012; Wong, Steiner, & Cook, 2013) for middle school students. I centered and standardized the CELDT score at 556 and the CST score at 325—the requirements for MUSD for middle school students. The assignment variable represents the minimum score of the centered and standardized CELDT and CST scores, therefore a score lower than zero denoted the student had failed at least one exam. The assignment variable was used in RD models to make causal inferences and determine the effects of reclassifying in middle school on their CST ELA, CAHSEE ELA/math, and math course placement; as well as, the effects on student's attendance and suspension.

The RD design was implemented by estimating the equations of the following general form:

$$Y_{ist} = \alpha + \beta I(CC_{ist-1} < 0) + f(CC_{ist-1}) + \varepsilon_{ist} \ , \qquad \text{(F3)}$$

In (F3), $Y_{ist}$ is the outcome (i.e., achievement and behavior) for student $i$ in school $s$ in year $t$. The variable, $CC_{-1}$, is the "assignment variable" in this RD design, based on middle school students combined CELDT and CST scores. The parameter of interest, β, identifies the jump in outcomes when the middle school language minority student is above the assignment threshold, conditional on $f(CC_{ist-1})$ a function of the assignment variable which was estimated using local linear regressions. Those students close to the cutoff, on either side of the threshold, were observationally and non-observationally similar, and thus, they could be used as proxies for each other's missing counterfactual. I used Nichols (2007) RD Stata package, which required neither student-level covariates nor school factors. I found the discrepancies between eligibility and reclassification did not increase sharply from 0 to 1 at the threshold; as a result, it is considered a "fuzzy" RD (Papay, Murnane, & Willett, 2011). Therefore, I report the denominator of the local Wald estimator—the jump in the treatment (Stage 1)—and the local Wald estimator (numerator/denominator)—the jump in the outcome (Stage 2).

Stage 1 results are alike throughout all the RD models. See Appendix Figure 3.5A as one example of these discontinuities. To conduct RD models the assignment variable (the combined centered standardized CELDT and CST) must predict the treatment (ELL and RFEP). On average, ELL middle school students had a 40% probability of being reclassified if they passed the CELDT at 556 and CST at 325. This probability ranged from 35% to 50% depending on the cohort and sample size. Administrators followed the recommended cutoffs more closely in the later school years. Furthermore, ELL students below the cutoff had about 10% probability of being reclassified. This fuzzy RD allows us to compare nearly similar students near the cutoff threshold.

The RD models are limited in two ways. First, since it is only possible to estimate regression discontinuity analyses in districts that have implemented a formula-based placement system, they tell us little about the extent to which various district language classification policies moderate the effects of achievement and behavioral outcomes in districts that do not follow similar policies. Most California school districts have adopted more rigorous reclassification requirements than those required by the state (Hill, Weston, et al., 2014). Thus, these studies results will be generalizable to school districts with similar language classification policies and policy-implementation procedures. Second, these analyses only estimate the effects of language classification near the classification threshold, providing limited evidence for those students who scored either very high or very low on the CELDT and CST exams. To address this issue, I provided estimates using several bandwidths as robustness checks for the RD models. Nichols (2007) RD Stata package automatically selected the three optimal bandwidths. The bandwidths varied for each model and they also varied by cohort year. I present the effect sizes based on the Nichols (2007) RD Stata package bandwidths. Also OLS regression coefficients were presented next to the RD coefficients to compare effect sizes.

## Results

### Descriptive Differences Amongst ELL and RFEP Students

Before I provide the language classification effects, I provide descriptive information. Students who became RFEP by the end of middle school (henceforth simply referred to as RFEP) had better achievement and behavioral outcomes than students who remained ELL. Table 3.2 shows RFEP CST ELA, CAHSEE ELA and CAHSEE math averages were higher than for ELL students. Furthermore, RFEP students were more likely to be enrolled in more advanced math courses in high school. For instance, 30% and 35% of RFEP ninth graders were enrolled in geometry and honors algebra, respectively, in comparison with ELL ninth graders, where only 7% and 15% were enrolled in geometry and honors algebra, respectively.[28] However, students

[28] I do not compare English course placement in high school because approximately 80% or more of ELL and RFEP students are enrolled in regular English courses; there is not enough variation to compare English course placement.

who reclassified may be those who were more academically inclined; for this reason, their higher English and math achievement scores may be spurious because of some other unmeasured variable (e.g., motivation, cognitive abilities, parental expectations, etc.) that capture academic ability. Students who were placed in higher math courses may also be certain individuals teachers deem more capable because of some unmeasured variables that teachers observe firsthand, but that tests do not capture. Also, ELL students, on average, had higher absence and suspensions rates than RFEP students. For example, RFEP ninth graders had a mean of 5.5 absences and 0.28 on-campus suspensions compared with ELL who had a mean of 8.1 and 0.52, respectively. In short, although language classification itself should not influence achievement and behavior outcomes because students should be getting appropriate courses, prior research demonstrates that misconceptions of students by language classification do indeed occur.

Furthermore, high school student achievement outcomes can be influenced by their middle school language course placement. Table 3.2A in the Appendix provides the percentage of RFEP seventh graders and ELL seventh graders who were placed into eighth-grade English and math courses that varied in their rigor. As language classification was determined at the end of the school year, seventh grade classification can influence students' eighth grade course placement. Appendix Table 3.2A shows a student's seventh grade language classification was associated with that student's eighth-grade English and math course placements, which, in turn, can determine the student's high school outcomes. A Pearson's chi-square test showed RFEP students were more likely to be placed into one mainstream English course instead of two remedial English courses in the eighth grade $X^2(5,N=2,269)=754.6$, $p<.001$. The mainstream courses included regular and honors English, and the remedial courses included ELM, ELD, and Intensive Literacy. Furthermore, RFEP students were also more likely to be placed into accelerated math courses (e.g., regular algebra, honors algebra) than basic math courses (e.g., basic algebra, pre-algebra) $X^2$ (5,N=2,269)=549.7, $p$ <.001. Language minorities' high school achievement can be influenced by the opportunities they received in middle school given their language classification. The main analyses of achievement include the effects of language classification and course placement as an interaction term.

**Middle School Language Classification and High School Achievement Outcomes**

**CST ELA.** To address how language classification affects achievement (Research Question 1) I run four logistic regression models that estimate the effects of students' language classification and prior course placement in middle school on ninth and tenth grade CST ELA scores. Models 1 and 3 on Table 3.3 demonstrate that RFEP classification was associated with a 0.30 standard deviation (hereafter σ) increase in the CST ELA ninth grade scores ($p$ <.001), and a 0.16σ increase in the CST ELA scores in tenth grade ($p$ <.001), respectively. Model 1 also shows that being placed into one mainstream English course (versus two remedial English courses) was associated with a 0.17σ increase in the CST ELA exam score ($p$ <.001) for ninth graders. Similarly, Model 3 shows that being placed into one mainstream English course (versus two remedial English courses) was associated with a 0.21σ increase in the CST ELA exam score ($p$ <.001) for tenth graders.

Next, to determine if mainstream English courses were equally beneficial to language minorities, I ran models with interactions. Model 2 shows that ninth grade RFEP students placed into one mainstream English course (0.47σ, $p$ <.001) or into two remedial English courses (0.32σ, $p$ <.001) outperformed ELL students placed into two remedial courses. Ninth grade ELL students placed into one mainstream English course (0.17σ, $p$ <.01) also outperformed ELL students placed into two remedial English courses. Similarly, Model 4 shows that tenth grade

RFEP students placed into one mainstream English course (0.37σ, $p$ <.001) and ELL students placed into one mainstream English course (0.17σ, $p$ <.01) outperformed ELL students placed into two remedial English courses. RFEP students placed into two mainstream English courses have equivalent CST ELA scores (0.18σ, *NS*) as ELL students placed into two remedial courses. Thus, ELL students appear to benefit from taking one English course instead of two period English courses. As the next step, I then sought to determine if language classification itself, net of skills, affects students CST scores.

Numerous studies show ELL students underperform on various measures relative to RFEP (e.g., Edwards et al., 2008; Gandara & Rumberger, 2009; Grissom, 2004; Mosqueda & Maldonado, 2013; Saunders & Marcelletti, 2012; Slama, 2014). My conclusion would have been the same if I only ran OLS models. The OLS coefficient with different bandwidths ranged between 0.17 ($p$ <.01) and 0.30 ($p$ <.001) for the CST ELA scores for ninth graders, and 0.18 (*NS*) and 0.16 ($p$ <.001) for the CST ELA scores for tenth graders. However, RD models provide more precise and less biased effect sizes than OLS models. The language classification itself is not significant for the marginal student. Table 3.4, RD Stage 2 results demonstrates that language classification does not directly affect a student's CST ELA scores. Figures 3.1 and 3.2 visually demonstrate there is no discontinuity at the combined CELDT/CST (assignment variable) cutoff that determines CST ELA scores.

**CAHSEE.** I also run four logistic regression models that estimate the effects of student language classification and prior course placement in middle school on tenth grade CAHSEE scores. Models 1 and 3 on Table 3.5 demonstrate that RFEP classification was associated with a 0.25σ increase in the CAHSEE ELA scores ($p$ <.001), and a 0.39σ increase in the CAHSEE math scores ($p$ <.001), respectively. Model 1 also shows that being placed into one mainstream

English course (versus two remedial English courses) was associated with a 0.17σ increase in the CAHSEE ELA exam score ($p$ <.001). Similarly, Model 3 shows that being placed into an accelerated math course (versus a basic math course) was associated with a 0.39σ increase in the CAHSEE math exam score ($p$ <.001).

Next, to determine if mainstream English courses and accelerated math courses were equally beneficial to language minorities, I ran models with interactions. Model 2 shows that RFEP students placed into one mainstream English course (0.42σ, $p$ <.001) or into two remedial English courses (0.17σ, $p$ <.001) outperformed ELL students placed into two remedial courses. ELL students placed into one mainstream English course (0.26σ, $p$ <.001) also outperformed ELL students placed into two remedial English courses. Similarly, Model 4 shows that RFEP students placed into accelerated math courses (0.76σ, $p$ <.001) or basic math courses (0.43σ, p <.001) outperformed ELL students placed into basic math courses. Additionally, ELL students placed into accelerated math courses (0.45σ, $p$ <.001) outperformed ELL students in basic math courses. These results show that middle school language classification and course placement are strongly associated with a student's CAHSEE scores. Thus, ELL students appear to benefit from taking one English course instead of two period English courses, and from taking an accelerated math course instead of a basic math course. As the next step, I then looked to determine if language classification itself, net of skills, affects students CAHSEE scores.

As previously stated, numerous studies show ELL students underperform on various measures relative to RFEP, and my conclusion would have been the same if I only ran OLS models. However, RD models provide more precise and less biased effect sizes than OLS models. The language classification itself is not significant for the marginal student. Table 3.6, RD Stage 2 results demonstrates that language classification does not directly affect a student's

CAHSEE ELA or math scores. Figures 3.3 and 3.4 visually demonstrates that there is no discontinuity at the combined CELDT/CST (assignment variable) cutoff that determines CAHSEE ELA and math scores.

As robustness checks, I ran two other RD models where students passed (1) or failed (0) the CAHSEE ELA and math portions. Appendix Table 3.3A shows that RFEP students are less likely to pass the CAHSEE ELA portion than ELL students -0.30 ($p < .05$), but are equally likely to pass the CAHSEE math -0.08 (*NS.*) Robinson (2011) RD models also show high school RFEP students score lower than ELL students on the CST ELA.

**Math course placement in high school.** High school math courses are typically hierarchical, meaning prerequisite classes position some students above others, which provides some students the advantage of starting higher in the sequence. In the majority of cases, a student's placement in the hierarchy of course-taking that begins in middle school limits how far they will reach in the math series by the end of high school. Prior studies show ELL students are more likely to be placed in less rigorous courses (Callahan, 2005; Callahan et al., 2010; Gandara et al., 2003; Hahnel et al., 2014; Mayer, 2008). In Table 3.7, Model 1 shows that RFEP ninth graders had greater odds of being placed into geometry (OR 3.95, $p < .001$) or honors algebra (OR 3.22, $p < .001$) than into regular algebra, and that they were less likely to be placed into basic math (OR 0.24, $p < .001$). However, Model 2 also illustrates that math course placement in middle school moderates the association between language classification and high school math course placement. For example, ELL students placed into accelerated math courses in middle school had greater odds of being placed into geometry (OR 5.06, $p < .001$) or honors algebra (OR 2.42, $p < .001$) compared with ELL students who had been put into regular algebra courses.

Similar analyses were conducted with the first cohort regarding eleventh grade math course placement. At the point at which the data was collected, neither Cohort 2 nor Cohort 3 had reached the eleventh grade. In Table 3.8, Model 1 shows RFEP eleventh graders had greater odds of being placed into trigonometry (OR 2.59, $p$ <.001) or algebra II (OR 1.94, $p$ <.001) than geometry, but they were equally likely to be placed into algebra I (OR 0.91, *NS*). Model 2 shows that math course placement in middle school moderates the association between language classification and math course placement in the eleventh grade. For example, ELL students placed into accelerated math courses in middle school had greater odds of being placed into trigonometry (OR 17.46, $p$ <.001) or algebra II (OR 1.62 , $p$ <.001) when compared with ELL students who were placed in basic math courses in middle school. Thus, we see that language classification and course placement in middle school is strongly associated with math course placement in high school.[29]

ELL students may have the math skills to do well in accelerated math courses, but, due to their language classification, may be placed into basic math courses. For example, some math teachers may believe that their ELL students do not have substantial English proficiency to master rigorous math material, and, because of this, their ELL students were enrolled in less stringent math courses. Table 3.9 shows that the student's language classification itself does not affect the student's high school math course placement. Figures 3.5 and 3.6 show a linear relationship between middle school language classification and high school math course placement, but no discontinuity at the cutoff. This table highlights the comparable OLS and RD effect sizes for students who were about 0.5, 1.5, and 2.5 standard deviations away from the assignment variable. Thus, we find that comparable RFEP and ELL students were neither more

[29] English course placement in high school was not included because most students were placed in regular English courses. For instance, 75% of ELL and 83% of RFEP ninth graders were placed in regular English. Further, 90% of ELL and 82% of RFEP eleventh graders were placed in regular English courses.

nor less likely to be placed into accelerated math courses due to their language classification. The overall effect of the marginal student is not significant.

**Middle School Language Classification and High School Behavioral Outcomes**

I conduct several OLS and RD models to address (Research Question 2), how does language classification by the end of middle school affect high school students' behavioral outcomes (i.e., attendance and suspensions)?

**Absences.** A student's behavioral outcome may also be affected by language classification due to how others perceive them and/or the student's disengagement. For example, ELL students may be more likely to be absent because they feel that their classes are not engaging. The OLS results in Table 3.10 illustrate that ELL students had more absences (0.16σ, $p$ <.01) than RFEP students in the ninth through eleventh grades. However, when I restricted the OLS models to include students closer to the assignment variable cutoff, the absence coefficient decreased, becoming insignificant. Furthermore, the RD models, Stage 2 coefficients were not significant, regardless of cohort, bandwidth, or grade level.

**On-and off-campus suspension.** Students classified ELL by the end of middle school are not always more likely to be suspended. The OLS results in Table 3.11 illustrate that ELL students had approximately one more on-campus suspension (0.10σ, $p$ <.05) than RFEP students only in the ninth grade. Table 3.12 illustrates that the number of off-campus suspensions was only statistically different amongst tenth grade ELL and RFEP students. However, for both on- and off-campus suspensions when I restricted the OLS models to include students closer to the assignment variable cutoff, the absence coefficient decreased, becoming insignificant as far as OLS results. Furthermore, the RD models, Stage 2 coefficients were small and not significant, regardless of cohort, bandwidth, and, for most, grade level. One exception was RFEP students were more likely to be suspended on-campus ($0.59\sigma$, $p <.05$) than ELL students in tenth grade. The second exception was RFEP students were more likely to be suspended off-campus ($0.13\sigma$, $p <.05$) than ELL students in eleventh grade. The effect sizes may be small, and they are not consistently significant in all bandwidths, however, it demonstrates that in some circumstances RFEP students were more likely to be suspended due to their language classification status.

As a robustness checks, on- and off-campus suspension outcomes were dichotomized into (1) never suspended, and (0) suspended once or more. These effect sizes were not statistically significant and can be provided upon request.

**Robustness Checks**

Similar effect sizes were obtained across the various bandwidth choices of the regression discontinuity models as well as the regression models; thus the magnitude of these results is not an artifact of the model specification. RD results were provided for students who were about 0.5, 1.0, or 1.5 standard deviations away from the CELDT and CST cutoff scores, but varied based on the Nichols (2007) RD package. However, as robustness checks, several other bandwidths were also selected. I found that the results remained the same, regardless of bandwidth. Further, all RD models were run separately for each cohort, but the results remained the same. For this reason, when possible, I combined the cohorts.

## Discussion

In California, language minorities are assessed to determine whether they need additional language support to become English proficient. Based on these assessments, those designated as ELL students should be receiving curricula that are different from their RFEP peers and specifically tailored to them. However, prior research has demonstrated that RFEP students outperform ELL students in English and math on every measure examined, even when controlling for prior achievement (Edwards et al., 2008; Grimssom, 2004; Gandara & Rumberger, 2009; Mosqueda & Maldonado, 2013; Slama, 2014; Saunders & Marcelletti, 2013). Not only are ELLs English skills incomparable to RFEP skills, their math skills are also lower. The difference between RFEP and ELL students may be explained by the fact that higher achieving students are reclassified while lower achieving students remain ELL. However, the language classification itself may have unintended consequences, where academically inclined ELL students may underperform because they are not given adequate opportunities to learn. Moreover, ELL students may be treated differently than RFEP students, which can lead to a greater number of absences and suspensions (Losen & Martinez, 2013; Uriarte et al., 2009).

This study's OLS estimates also demonstrate that RFEP students have higher CST and CAHSEE scores, and, additionally, that RFEP students are more likely to be placed in more advanced math courses in high school than the ELL students. Moreover, OLS models with interactions demonstrate that high school CST ELA and CAHSEE ELA scores increase when students are placed in one period English course instead of two period English courses in middle school. CAHSEE math scores also increase when students are placed in accelerated math courses instead of basic math courses in middle school. ELL students who are placed in a one period English course and an accelerated math course, in particular, see an increase in their English and math achievement outcomes. Furthermore, OLS models show ELL students are more likely to be absent and have more on-campus suspensions than RFEP students. However, the RD models provide a less biased estimate, showing that in most cases these differences are due to unmeasured factors. In most instances, academic and behavioral differences between ELL and students who RFEP in the seventh or eighth grades are spurious and not due to language classification itself. There were only two instances where language classification mattered, and in both cases RFEP student performance was worse than ELL. In specific RD bandwidths, RFEP students were less likely to pass the CAHSEE ELA portion, and were more likely to be suspended.

This study's results coincide with the Robinson studies, although this study focused on a different school district with different language classification policies. First, Robinson (2011) finds RFEP high school students score lower than ELL high school students on the CST ELA exam, but that there is no difference for elementary or middle school students. The same study shows that language classification does not affect high school course placement in English and math, nor attendance. Second, the Robinson-Cimpian and Thompson (2015) RD model demonstrates that more rigorous classification policies making it more difficult to reclassify increases high school students' CST ELA scores, but that there is no effect on middle school students' CST ELA scores. Robinson's studies show no differences amongst ELL and RFEP students in elementary and middle school. However, RFEP high school students score lower on the CST ELA exam compared with ELL students unless the reclassification policies are more rigorous. Here, I also find language classification by the end of eighth grade did not affect students' CST ELA scores, CAHSEE math scores, high school math course placement, and attendance. However, in a few instances classification by the end of middle school did negatively affect the RFEP students' probability of passing the CAHSEE ELA portion, and whether they got suspended. These negative effects were only statistically significant when using some RD bandwidths.

The present study expands our knowledge regarding the consequences of language classification itself. First, the majority of the literature on language classification conducts OLS models with non-randomized data where researchers cannot make causal inferences. Based on

OLS, researchers conclude that reclassifying is beneficial for ELL students. This study's RD models demonstrate that, for the most part, both ELL and students who RFEP in middle school have comparable achievement and behavioral outcomes. Language classification policies seem to be working appropriately at MUSD. One possible reason for this is students are receiving appropriate English and math curricula and, therefore, there is no difference. Academically inclined students seem to be reclassified, and any differences found in OLS models seem to be based on unmeasured factors that highly correlate with student language classification. Second, this study also expands on Robinson's studies that also use RD models (Robinson-Cimpian & Thompson, 2015; Robinson, 2011). This study also examines students' CST ELA scores, math course placement, and attendance, while also examining student CAHSEE ELA, CAHSEE math, and suspension outcomes. Third, this study focuses on students who specifically reclassified in middle school, and does not aggregate all RFEP students. The RFEP student body composition changes each year because each school year's ELL students have the opportunity to become RFEP. In the end, this study demonstrates that, generally speaking, simply reclassifying a student in middle school will not increase their achievement and behavioral outcomes.

Starting in 2014-2015, central language classification assessments, such as the CELDT and CST, were replaced by the English Language Proficiency Assessments for California (ELPAC) and Smarter Balanced Assessment Consortium (SBAC) (Umansky et al., 2015). Furthermore, the state adopted the English language development (CA ELD) standards that are comparable in rigor and specificity to the California Core State Standards (CCSS), the California Common Core State Standards for Mathematics (CA CCSSM), and the Next Generation Science Standards (CA NGSS) (Laguoff, Spycher, Linquanti, Carroll, & DiRanna, 2015). California now makes recommendations as to what ELL students should be learning in their mathematics and science classrooms. Here, the content should be as rigorous as non-ELL classrooms, and students' proficiencies should be developed in all subject areas. These changes will definitely alter how ELL students are reclassified, as well as the type of content they will learn. District administrators should be encouraged to provide English and math courses that will develop ELL students' English proficiency and provide them access to rigorous content knowledge. When making language classification and course placement policies, administrators need to consider the fact that ELL and RFEP students have different language support needs. Further, they should see that although ELL students may not have reached full English proficiency yet, they should still be given opportunities to become as college ready and academically prepared as their RFEP counterparts. Future research should address the effects of these new policies and compare middle school ELL students' achievement and behavioral outcomes to different groups of students (e.g., RFEP in middle school, RFEP in elementary, English Only).

**Limitations**

In interpreting this study's findings, I note the following empirical limitations. First, the quantitative analyses focuses on the 18% of students who have CELDT and CST scores—the required reclassification assessments. About 8% of sixth grade ELL students did not have the necessary scores to be included in the RD models. A second limitation is the lack of Individualize Education Program (IEP) information for special education students. IEP's vary greatly and the specification for special education students can influence there probability of course placement, and ultimately their achievement and behavioral outcomes. Special education students are normally dropped from analysis (for example Hill, Weston, et al., 2014) but they remained in these analyses due to the high representation of special education students in middle school. Many special education students are long-term English language learners and the intersection must be further research particularly in middle school. Future research should address the intersection of special education and ELL classification especially when several of these students cannot RFEP since many are exempted from taking the required reclassification exams. A third limitation is that I only compare students who reclassified in seventh and eighth grade to current ELL students. I did not have access to elementary ELL students CELDT and CST scores to create the necessary assignment variable. Prior research should examine each grade level but account and specify the timing of reclassification when conducting RD models. A final limitation of this study, or any study that attempts to disentangle the effects of language classification on achievement, is the variability of language classification policies. These results are generalizable to school districts that have similar language classification policies as Manzanita district.

**Conclusion**

As most language classification literature has found, I also find a linear association between language classification and English and math achievement scores, as well as behavioral outcomes. However, in most cases students' middle school language classification did not directly affect their high school achievement or behavior outcomes. This study demonstrates that language classification itself is not contributing to the achievement gap and behavioral differences when comparing ELL with students who reclassified in middle school. In the data used in this study it appears that simply reclassifying students will not improve their educational outcomes.

References

Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in US schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoro, S. (2005). The new demography of America's schools: Immigration and the No Child Left Behind Act. *Urban Institute (NJ1)*.

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming High and Falling Short California's Eighth-Grade Algebra-for-All Effort. *Educational Evaluation and Policy Analysis, 37*(3), 275-295.

Edwards, B., Leichty, J., & Wilson, K. (2008). English Learners in California: What the Numbers Say. *EdSource*.

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources, Unequal outcomes. *Education Policy Analysis Archives, 11*(36), 1-54.

Gandara, P., & Rumberger, R. W. (2009). Immigration, language, and education: How does language policy structure opportunity. *Teachers College Record, 111*(3), 750-782.

Grissom, J. B. (2004). Reclassification of English Learners. *Education Policy Analysis Archives, 12*(36), n36.

Hahnel, C., Wolf, L., Banks, A., & LaFors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. The Education Trust-West.

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.

Hampden-Thompson, G., Mulligan, G., Kinukawa, A., & Halle, T. (2008). Mathematics Achievement of Language-Minority Students During the Elementary Years.

Hill, L. E., Betts, J. R., Chavez, B., Zau, A. C., & Bachofer, K. V. (2014). Pathways to Fluency: Examining the Link between Language Reclassication Policies and Student Success. *Public Policy Institute of California*.

Hill, L. E., Weston, M., & Hayes, J. M. (2014). Reclassification of English Learner Students in California. *Public Policy Institute of California. Retrieved from www. ppic. org/main/publication. asp*

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615-635.

Kohler, A. D., & Lazarín, M. (2007). Hispanic education in the United States. *Statistical Brief, 8*.

Krashen, S., & Brown, C. L. (2005). The ameliorating effects of high socioeconomic status: A secondary analysis. *Bilingual Research Journal, 29*(1), 185-196.

Laguoff, R., Spycher, P., Linquanti, R., Carroll, C., & DiRanna, K. (2015). *Integrating the CA ELD Standards into K-12 Mathematics and Science Teaching and Learning* Retrieved from WestEd: www.cde.ca.gov

Lee, D. S., & Lemieux, T. (2010). *Regression discontinuity designs in economics*. Retrieved from http://www.jstor.org/stable/20778728

Losen, D. J., & Martinez, T. E. (2013). Out of school and off track: The overuse of suspensions in American middle and high schools. *K-12 Racial Disparities in School Discipline*.

Lovaglia, M. J., Lucas, J. W., & Thye, S. R. (1998). Status Processes and Mental Ability Test Scores 1. *American Journal of Sociology*, *104*(1), AJSv104p195-228.

Maxwell, L. (2014). US school enrollment hits majority-minority milestone. *Education Week*.

Mayer, A. (2008). Understanding How US Secondary Schools Sort Students for Instructional Purposes: Are All Students Being Served Equally? *American Secondary Education*, 7-25.

Mosqueda, E., & Maldonado, S. I. (2013). The Effects of English Language Proficiency and Curricular Pathways: Latina/os' Mathematics Achievement in Secondary Schools. *Equity & Excellence in Education*, *46*(2), 202-219.

Nichols, A. (2007). Causal Inference with Observational Data. *Stata Journal*, *7*(4), 507-541.

Olsen, L. (2010). Reparable Harm Fulfilling the Unkept Promise of Educational Opportunity for California's Long Term English Learners. *California Together (Research Report)*.

Papay, J. P., Murnane, R. J., & Willett, J. B. (2011). *How performance information affects human-capital investment decisions: The impact of test-score labels on educational outcomes*. Retrieved from https://www.researchgate.net/profile/Richard_Murnane/publication/228303490_How_Performance_Information_Affects_Human-Capital_Investment_Decisions_The_Impact_of_Test-Score_Labels_on_Educational_Outcomes/links/0046352854781893380000000.pdf

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, *5*(1), 83-104.

Robinson-Cimpian, J. P., & Thompson, K. D. (2015). The Effects of Changing Test-Based Policies for Reclassifying English Learners. *Journal of Policy Analysis and Management*.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292.

Saunders, W. M., & Marcelletti, D. J. (2012). The Gap That Can't Go Away The Catch-22 of Reclassification in Monitoring the Progress of English Learners. *Educational Evaluation and Policy Analysis*, 0162373712461849.

Schneider, B., Swanson, C. B., & Riegle-Crumb, C. (1997). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2*(1), 25-53.

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.

StataCorp. (2011). Stata Statistical Software: Release 12. StataCorp LP: College Station.

Stein, J. C., Hetzel, J., & Beck, R. (2011). Twice Exceptional? The Plight of the Gifted English Learner. *International Journal for Professional Educators*, *78*(2), 36-41.

Umansky, I. M., Reardon, S. F., Hakuta, K., Thompson, K. D., Estrada, P., Hayes, K., . . . Goldenberg, C. (2015). Improving the Opportunities and Outcomes of California's Students Learning English: Findings from School District-University Collaborative Partnerships. Policy Brief 15-1. *Policy Analysis for California Education, PACE*.

Uriarte, M., Lavan, N., Agusti, N., & Karp, F. (2009). English Learners in Boston Public Schools: Enrollment and Educational Outcomes of Native Spanish Speakers. *Gaston Institute Publications*. Paper 113.

Walqui, A., Koelsch, N., Hamburger, L., Gaarder, D., Insaurralde, A., Schmida, M., & Weiss, S. (2010). What are we doing to middle school English Learners. *Findings and* *recommendations for change from a study of California EL programs (Research Report). San Francisco: WestEd*.

Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school students' mathematics achievement. *The Journal of Educational Research, 97*(1), 3-17.

Wong, V. C., Steiner, P. M., & Cook, T. D. (2013). Analyzing regression-discontinuity designs with multiple assignment variables a comparative study of four estimation methods. *Journal of Educational and Behavioral Statistics, 38*(2), 107-141.

Zaragoza-Petty, A. L., & Zarate, M. E. (2014). College access factors of urban latina girls: The role of math ability perceptions *Journal of Urban Learning, Teaching, and Research, 10*, 64-72.

Table 3.1

*Manzanita Full, Selected, and Final Sample (*2010-2011 to 2012-2013)

| | Full Sample | Selected Sample | Final Sample |
|---|---|---|---|
| *District information* | | | |
| Total 8th grade enrollment | 16,144 | 4,231 (26%) | 2,969 (18%) |
| Average 8th grade cohort | 5,381 | 1,410 | 989 |
| Total # of middle schools | 13 | 8 (TS) | 8 (TS) |
| *Student demographics in 8th grade* | | | |
| % Female | 49.2 | 45.8 | 47.3 |
| % Hispanic or Latino | 65.8 | 87.6 | 88.3 |
| % Asian | 12.3 | 9.3 | 8.8 |
| % White | 13.3 | --- | --- |
| % African American | 3.1 | --- | --- |
| % Other race | 5.2 | 2.9 | 2.8 |
| % Born in United States† | 84.0 | 74.1 | 75.3 |
| % Free- and Reduced-Price Lunch | 71.7 | 90.1 | 90.2 |
| % Special Education | 10.1 | 19.7 | 4.6 |
| *Language Classification* | 8th Grade | 8th Grade | 8th Grade |
| % English Language Learners (ELL) | 22 | 72 | 65 |
| % Reclassified Fluent English Speakers (RFEP) | 40 | 28 | 35 |
| % English Only (EO) and Initially Fluent English Speakers (IFEP) | 38 | --- | --- |
| *Course placement* | 8th Grade | 8th Grade | 8th Grade |
| English Courses | | | |
| Honors English | 22.5 | 3.4 | 4.0 |
| Regular English | 50.4 | 42.3 | 47.6 |
| ELM & Reg. English | 1.1 | 4.2 | 4.4 |
| Two English (Remedial/Regular) | 18.9 | 36.9 | 36.3 |
| Two ELD | 2.7 | 7.0 | 5.5 |
| No English Class | 4.1 | 6.0 | 1.9 |
| Math Courses | | | |
| Trigonometry or Higher | .04 | 0 | 0 |
| Geometry | 5.4 | 1.1 | 1.3 |
| Honor Algebra | 16.3 | 5.2 | 5.7 |
| Regular Algebra | 30.8 | 21.6 | 24.9 |
| Basic Algebra | 27.8 | 38.3 | 41.2 |
| Pre-Algebra | 16.1 | 29.6 | 25.0 |
| No Math Course | 3.3 | 3.9 | 1.5 |

*Note*. The full sample represents averages over three school years of complete cohort data for middle school students provided by the school districts. The selected sample represents students who were classified ELL by 6th grade. It excludes White, African American, EO, IFEP, and elementary RFEP, and also excludes non-traditional schools (TS). The final sample are students classified ELL by 6th grade and who had both CELDT and CST scores.

Table 3.2

*High School Achievement and Behavioral Outcomes by Language Classification*

| | | Final Sample (n=2,969) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | RFEP by 8th Grade (n=1,038) | | | | | ELL in 8th Grade (n=1,931) | | | | |
| | Cohort | N | Mean | SD | Min | Max | N | Mean | SD | Min | Max |
| **Achievement Outcomes** | | | | | | | | | | | |
| CST ELA 9th grade | 1-2 | 853 | 360.0 | 36.9 | 262 | 509 | 1458 | 308.9 | 36.4 | 194 | 478 |
| CST ELA 10th grade | 1 | 408 | 352.3 | 39.5 | 218 | 478 | 747 | 304.9 | 36.3 | 208 | 438 |
| CAHSEE ELA 10th grade | 1-2 | 727 | 382.3 | 23.6 | 320 | 450 | 1,158 | 350.8 | 23.3 | 275 | 437 |
| CAHSEE Math 10th grade | 1-2 | 730 | 388.9 | 29.7 | 317 | 450 | 1,157 | 357.8 | 27.5 | 284 | 450 |
| 9th Gr. Math Course | 1-3 | 961 | | | | | 1,738 | | | | |
| Geometry | | | 30% | --- | --- | --- | | 7% | --- | --- | --- |
| Honor Algebra | | | 35% | --- | --- | --- | | 15% | --- | --- | --- |
| Regular Algebra | | | 36% | --- | --- | --- | | 78% | --- | --- | --- |
| Basic Math | | | 0% | --- | --- | --- | | 1% | --- | --- | --- |
| 11th Gr. Math Course | 1 | 327 | | | | | 560 | | | | |
| Trigonometry | | | 20% | --- | --- | --- | | 5% | --- | --- | --- |
| Algebra II | | | 29% | --- | --- | --- | | 12% | --- | --- | --- |
| Geometry | | | 30% | --- | --- | --- | | 44% | --- | --- | --- |
| Algebra I | | | 21% | --- | --- | --- | | 40% | --- | --- | --- |
| **Behavioral Outcomes** | | | | | | | | | | | |
| Absences 9th | 1-3 | 1038 | 5.5 | 8.6 | 0 | 81 | 1931 | 8.1 | 12.8 | 0 | 128 |
| Absences 10th | 1-2 | 803 | 6.2 | 10.1 | 0 | 113 | 1346 | 9.9 | 14.5 | 0 | 166 |
| Absences 11th | 1 | 368 | 5.9 | 9.2 | 0 | 95 | 671 | 11.0 | 17.7 | 0 | 179 |
| On-Campus Suspension 9th | 1-3 | 1038 | .28 | 1.2 | 0 | 17 | 1931 | .52 | 1.7 | 0 | 32 |
| On-Camp. Suspension 10th | 1-2 | 803 | .18 | .83 | 0 | 8 | 1346 | .26 | 1.0 | 0 | 19 |
| On-Camp. Suspension11th | 1 | 368 | .11 | .67 | 0 | 8 | 671 | .17 | .72 | 0 | 9 |
| Off-Camp. Suspensions 9th | 1-3 | 1038 | .14 | 1.1 | 0 | 26 | 1931 | .39 | 1.5 | 0 | 19 |
| Off-Camp.Suspensions10th | 1-2 | 803 | .10 | .67 | 0 | 10 | 1346 | .29 | 1.3 | 0 | 21 |
| Off-Camp. Suspensions11th | 1 | 368 | .09 | .63 | 0 | 7 | 671 | .16 | 0.8 | 0 | 9 |

*Note*. †This is a longitudinal dataset, therefore, some of the high school outcomes are only available for some of the cohorts. Cohort 1 (n=1,118), Cohort 2 (n=1,043), and Cohort 3 (n=808).

Table 3.3

*CST Standardized Scores, Language Classification, and Course Placement (Regressions)*

| | (1)<br>9th Grade CST ELA Scores | (2)<br>9th Gr. CST ELA Scores (Interaction) | (3)<br>10th Grade CST ELA Scores | (4)<br>10th Gr. CST ELA Scores (Interaction) |
|---|---|---|---|---|
| RFEP by 8th Grade | 0.30*** | ---- | 0.16*** | ---- |
| | (0.03) | ---- | (0.03) | ---- |
| CST ELA 8th Gr. Scores | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| One English Course | 0.17*** | ---- | 0.21*** | ---- |
| | (0.03) | ---- | (0.03) | ---- |
| ELL x Two English (reference) | | | | |
| ELL x One English | | 0.17** | | 0.22** |
| | | (0.04) | | (0.06) |
| RFEP x Two English | | 0.32*** | | 0.18 |
| | | (0.02) | | (0.12) |
| RFEP x One English | | 0.47*** | | 0.37*** |
| | | (0.03) | | (0.04) |
| Constant | -4.21*** | -4.22*** | -4.49*** | -4.49*** |
| | (0.15) | (0.15) | (0.24) | (0.23) |
| N | 1891 | 1891 | 936 | 936 |
| R-sqr | 0.62 | 0.62 | 0.58 | 0.62 |
| df | 8 | 8 | 8 | 8 |

*Note*. All models include gender, race, country born, FRL, special education, cohort, and school fixed effects but were not shown in the table to conserve space. The sample size is smaller than the final sample because Models 1 and 2 only include cohort 1 and 2. Models 3 and 4 only include cohort 1. Models 2 and 4 include an interaction between eighth-grade language classification and eighth-grade English course placement.

Table 3.4

*Language Classification and CST Standardized ELA Scores*

| CST ELA 9th Grade (Two Years) N=2,161 | | | | | CST ELA 10th (One Year) N=1,043 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 2.0 | 1756 | .50*** (.03) | **-.05 (.19)** | .30*** (.03) | 1.6 | 792 | .51*** (.05) | **.03 (.19)** | .16*** (.03) |
| 1.0 | 1193 | .42*** (.05) | **-.08 (.33)** | .23*** (.04) | 0.8 | 472 | .46*** (.08) | **-.11 (.32)** | .08*** (.04) |
| 0.5 | 675 | .35*** (.08) | **-.00 (.11)** | .17*** (.04) | 0.4 | 259 | .41** (.13) | **-.62 (.63)** | .08 (.06) |

*Note*. The assignment variable is based of students CELDT and CST scores. The treatment variable is classified (0) ELL or (1) RFEP by eighth grade. The outcomes are ninth and tenth graders standardized CST ELA scores. OLS estimates are also provided where the sample size is restricted comparable to the RD models. The ninth grade models only include cohort one and two. The tenth grade models only include cohort one. The coefficients do not vary by cohort.

Figure 3.1

*RD for CST ELA Score in 9th Grade*

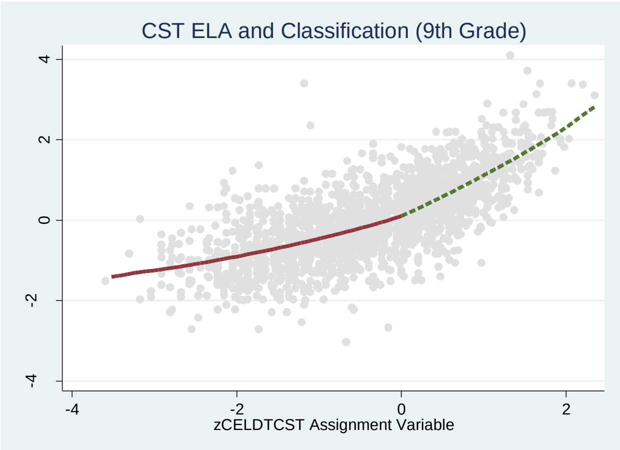



Figure 3.2

*RD for CST ELA Scores in 10th Grade*

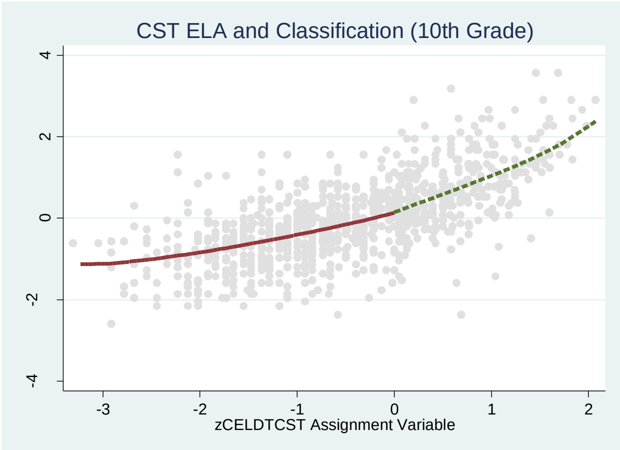


Table 3.5

*CAHSEE Standardized Scores, Language Classification, and Course Placement (Regressions)*

| | (1) CAHSEE ELA | (2) CAHSEE ELA (Interaction) | (3) CAHSEE MATH | (4) CAHSEE MATH (Interaction) |
|---|---|---|---|---|
| RFEP by 8th Grade | **0.25****** | ---- | **0.39****** | ---- |
| | **(0.02)** | ---- | **(0.02)** | ---- |
| CST 8th ± | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| One English Course | 0.17*** | ---- | ---- | ---- |
| | (0.02) | ---- | ---- | ---- |
| Accelerated Math | ---- | ---- | 0.39*** | ---- |
| | ---- | ---- | (0.05) | ---- |
| ELLxTwo English (reference) | | | | |
| ELLxOne English | | 0.26*** | | |
| | | (0.04) | | |
| RFEPxTwo English | | 0.17*** | | |
| | | (0.02) | | |
| RFEPxOne English | | 0.42*** | | |
| | | (0.02) | | |
| ELLxBasic (Reference) | | | | |
| ELLxAccelerated | | | | 0.45*** |
| | | | | (0.06) |
| RFEPxBasic | | | | 0.43*** |
| | | | | (0.02) |
| RFEPxAccelerated | | | | 0.76*** |
| | | | | (0.07) |
| Constant | -3.31*** | -3.31*** | -2.09*** | -2.12*** |
| | (0.13) | (0.16) | (0.21) | (0.21) |
| N | 1849 | 1849 | 1864 | 1864 |
| R-sqr | 0.573 | 0.573 | 0.565 | 0.565 |
| df | 8 | 8 | 8 | 8 |

*Note*. The sample size is smaller than the final sample because the models only include cohort 1 and 2. All models include gender, race, country born, FRL, special education, cohort, and school fixed effects but were not shown in the table to conserve space. ± Models 1 and 2 include students' eighth grade CST ELA scores. Models 3 and 4 include students' eighth grade CST math scores. Models 2 and 4 include an interaction between eighth-grade language classification and eighth-grade English and math course placement.

Table 3.6

*Language Classification and CAHSEE Standardized Scores*

| CAHSEE ELA (Two Years) N=2,133 | | | | | CAHSEE Math (Two Years) N=2,147 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 1.8 | 1647 | .49*** (.03) | **-.08 (.12)** | .25*** (.02) | 2.3 | 1820 | .51*** (.03) | **.05 (.12)** | .36*** (.02) |
| 0.9 | 1065 | .41*** (.05) | **-.22 (.20)** | .18*** (.01) | 1.1 | 1237 | .42** (.05) | **.01 (.20)** | .26*** (.03) |
| 0.4 | 659 | .35*** (.08) | **-.71 (.42)** | .13** (.04) | 0.5 | 750 | .36*** (.07) | **-.07 (.36)** | .17*** (.02) |

*Note*. The assignment variable is based of students CELDT and CST scores. The treatment variable is classified (0) ELL or (1) RFEP by eighth grade. The outcomes are tenth graders standardized CAHSEE ELA and math scores. OLS estimates are also provided where the sample size is restricted comparable to the RD models. All the models only include cohort one and two. The coefficients do not vary by cohort.

Figure 3.3

RD for *CAHSEE ELA Scores*

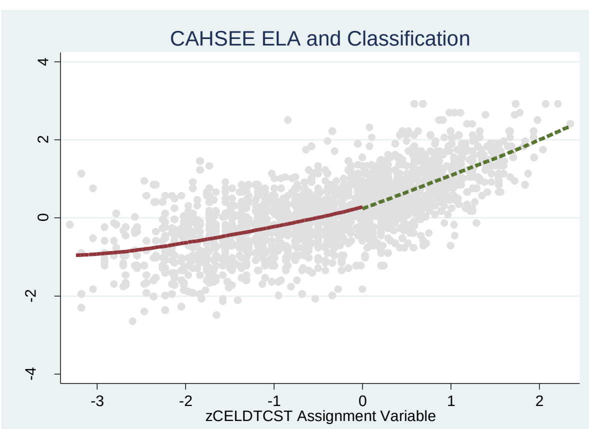



Figure 3.4

RD for *CAHSEE Math Scores*

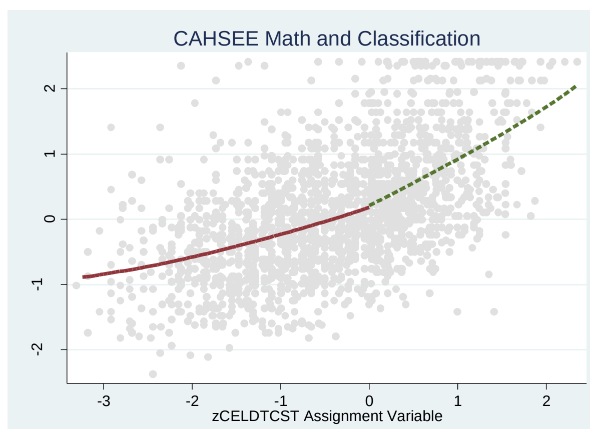


Table 3.7

*Odds Ratios in Placing in Regular Algebra in Ninth Grade (Two MLOGIT)*

| | $9^{th}$ Grade Math Course | | |
|---|---|---|---|
| Regular Algebra (Ref.) | Geometry | Honor Algebra | Basic Math |
| Model 1 | | | |
| RFEP by 8th | 3.95(1.21)*** | 3.22(0.41)*** | 0.24(0.07)*** |
| Accelerated $8^{th}$ Grade Math | 174.67(125.98)*** | 13.86(5.64)*** | 1.71(0.95) |
| Model 2 | | | |
| Interaction | | | |
| ELL x Basic (Reference) | | | |
| RFEP x Accelerated | 6.66(.90)*** | 3.90(.50)*** | 0.07(.92) |
| RFEP x Basic | 1.34(.61)* | 0.99(.11)*** | -2.04(.21)*** |
| ELL x Accelerated | 5.06(.92)*** | 2.42(.32)*** | -0.34(.64) |
| *N* | | | 2690 |
| df_m | | | 6.00 |
| pr2 | | | 0.41 |

*Note*. Both mlogit models include gender, race, country born, FRL, special education, cohort and school fixed effects but were not shown in the table to conserve space.

Table 3.8

*Odds Ratios in Placing in Geometry in Eleventh Grade (Two MLOGIT)*

| | $11^{th}$ Grade Math Course | | |
|---|---|---|---|
| Geometry (Reference) | Trig. or Higher | Algebra II | Algebra I |
| Model 1 | | | |
| RFEP by $8^{th}$ Grade | 2.59(0.73)*** | 1.91(0.32)*** | 0.91(0.21) |
| Accelerated $8^{th}$ Grade Math | 50205608.86*** (31187436.04) | 4.84(1.36)*** | 0.74(0.13) |
| Model 2 | | | |
| Interaction | | | |
| ELL x Basic (Reference) | | | |
| RFEP x Accelerated | 18.49(.68)*** | 2.29(.20)*** | -0.27(.19) |
| RFEP x Basic | -0.01(.23) | 0.78(.32)** | -0.31(.29) |
| ELL x Accelerated | 17.46(.67)*** | 1.62(.24)*** | -0.50(.09)*** |
| *N* | | | 886 |
| df_m | | | 5.00 |
| pr2 | | | 0.16 |

*Note*. Both mlogit models include gender, race, country born, FRL, special education, cohort and school fixed effects but were not shown in the table to conserve space. Models 1 and 2 ($11^{th}$ Grade math courses) had smaller sample size because the models only include cohort 1.

Table 3.9

*Language Classification and High School Math Course Placement*

| Model 1<br>9th Grade Math Course (Three Years)<br>N=2,922 | | | | | Model 2<br>11th Grade Math Course (One Year)<br>N=1,111 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 2.9 | 2898 | .55*** (.02) | **-.12 (.09)** | .14*** (.02) | 2.6 | 1084 | .54*** (.05) | **.00 (.06)** | .12** (.03) |
| 1.4 | 2262 | .49*** (.03) | **-.07 (.08)** | .12*** (.02) | 1.3 | 798 | .46*** (.06) | **-.05 (.17)** | .10* (.03) |
| 0.7 | 1375 | .43*** (.05) | **-.05 (.20)** | .08* (.02) | 0.6 | 431 | .39*** (.09) | **-.13 (.28)** | .05 (.05) |

*Note*. The assignment variable is based on students CELDT and CST scores. The treatment variable is classified ELL (0) or RFEP (1) by the end of middle school. Model 1 outcome is students' math course placement in 9th grade. The outcome is 1=honors algebra or higher, 0=regular algebra or lower. About 66% of ninth graders are placed in regular algebra or lower and 34% are placed in honors algebra or higher. Model 2 outcome is students' math course placement in 11th grade. The outcome is 1=Algebra II or higher, 0=Geometry or lower. This model only includes cohort one.

Figure 3.5

*RD for 9th Grade Math Courses*

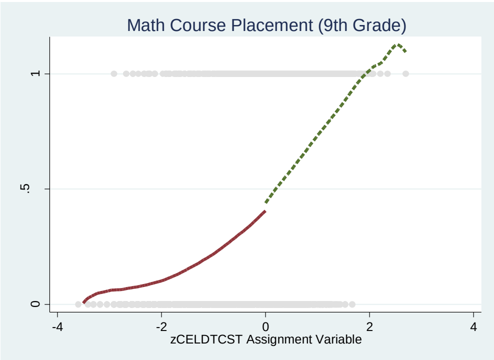



Figure 3.6

*RD for 11th Grade Math Courses*

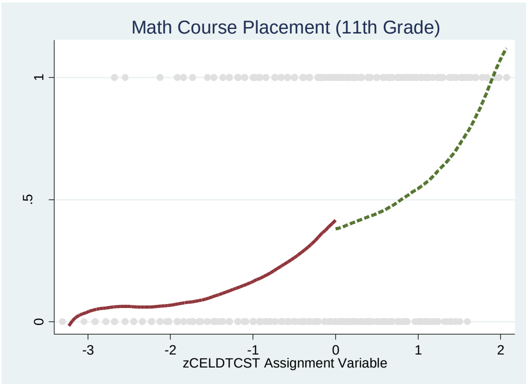


Table 3.10

*Language Classification and Absences in High School*

| 9th Grade Absences (Three Years) N=2,969 | | | | 10th Grade Absences (Two Years) N=2,149 | | | | 11th Grade Absences (One Year) N=1,039 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 2.2 | .51*** (.02) | **-.04 (.12)** | -.16** (.04) | 2.5 | .51*** (.03) | **-.03 (.14)** | -.21*** (.02) | 1.7 | .49*** (.05) | **-.36 (.32)** | -.22*** (.04) |
| 1.1 | .44*** (.04) | **-.13 (.19)** | -.12 (.06) | 1.2 | .42*** (.04) | **-.07 (.23)** | -.12* (.05) | 0.8 | .41*** (.08) | **-.82 (.63)** | -.11* (.04) |
| 0.5 | .39*** (.05) | **-.26 (.32)** | -.07 (.06) | 0.6 | .36*** (.06) | **-.12 (.40)** | -.05 (.09) | 0.4 | .32*** (.12) | **-1.1 (1.3)** | -.10 (.06) |

*Note*. The assignment variable is based on students' CELDT and CST scores. The treatment variable is classified ELL (0) or RFEP (1) by the end of middle school. The outcomes are the number of absences (0-179) in a given school year and the coefficient is standardized.

Table 3.11

*Language Classification and High School On-Campus Suspensions*

| 9th Gr. On-Campus Suspensions (Three Years) N=2,969 | | | | 10th Gr. On-Campus Suspensions (Two Years) N=2,149 | | | | 11th Gr. On-Campus Suspensions (One Year) N=1,039 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 1.6 | .48*** (.03) | **-.15 (.13)** | -.10* (.04) | 1.3 | .43*** (.04) | **.17 (.16)** | -.03 (.04) | 1.5 | .48*** (.05) | **.11 (.27)** | .02 (.07) |
| 0.8 | .42*** (.04) | **.00 (.20)** | -.11 (.05) | 0.6 | .37*** (.06) | **.44 (.23)** | -.01 (.07) | 0.7 | .40*** (.08) | **.45 (.39)** | .03 (.10) |
| 0.4 | .36*** (.07) | **.22 (.26)** | -.08 (.04) | 0.3 | .32*** (.09) | **.59* (.31)** | .05 (.05) | 0.3 | .30** (.14) | **.34 (.59)** | .00 (.18) |

*Note*. The assignment variable is based on students' CELDT and CST scores. The treatment variable is classified ELL (0) or RFEP (1) by the end of middle school. The outcomes are the number of on-campus suspensions (0-32) in a given school year and the coefficient is standardized.

Table 3.12

*Language Classification and High School Off-Campus Suspensions*

| 9th Grade Off-Campus Suspensions (Three Years) N=2,969 | | | | 10th Off-Campus Suspensions (Two Years) N=2,149 | | | | 11th Off-Campus Suspensions (One Year) N=1,039 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 4.0 | .54*** (.02) | **.02 (.11)** | -.12 (.06) | 2.3 | .50*** (.03) | **.22 (.13)** | -.12** (04) | 2.4 | .53*** (.04) | **.13* (.20)** | -.00 (.05) |
| 2.0 | .50*** (.03) | **-.01 (.13)** | -.11 (.06) | 1.1 | .41*** (.04) | **.30 (.23)** | -.05 (.06) | 1.2 | .45*** (.06) | **.45 (.28)** | -.04 (.08) |
| 1.0 | .44*** (.04) | **-.02 (.19)** | -.05 (.05) | 0.5 | .35*** (.07) | **.27 (.40)** | -.02 (.04) | 0.6 | .38*** (.10) | **.99* (.46)** | -.02 (.04) |

*Note*. The assignment variable is based on students CELDT and CST scores. The treatment variable is classified ELL (0) or RFEP (1) by the end of middle school. The outcomes are the number of off-campus suspensions (0-26) in a given school year and the coefficient is standardized.

## Appendix

Appendix Table 3.1A

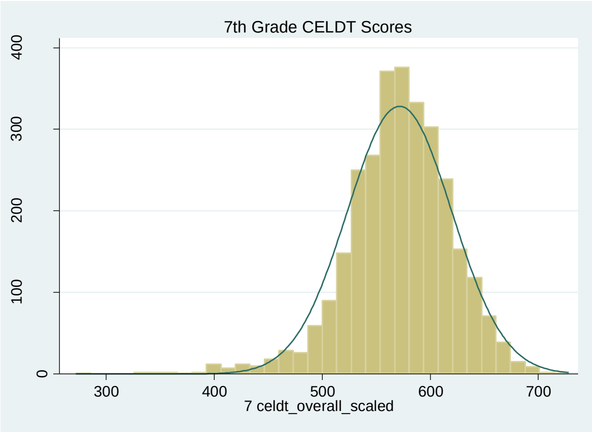
*Special Education and Students without Prior Scores are more likely to be Missing Reclassification Criteria in Manzanita District (MLOGIT/ODD RATIOS)*

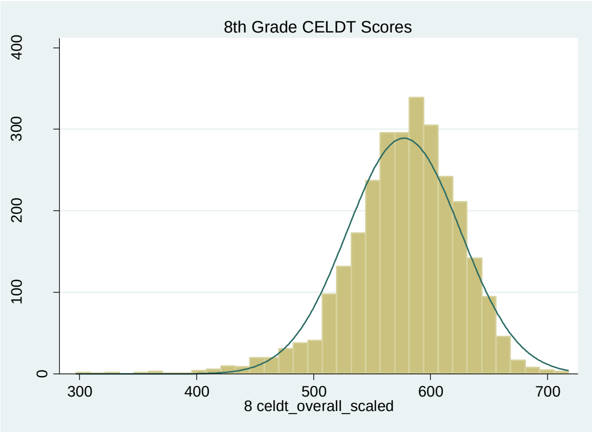
| | (1) 7th Graders (CELDT/CST) | (2) 8th Graders (CELDT/CST/ GPA) | (3) 8th Graders (CELDT/CST) | (1) 7th Graders (CELDT/CST) | (2) 8th Graders (CELDT/CST/ GPA) | (3) 8th Graders (CELDT/CST) |
|---|---|---|---|---|---|---|
| **All Criteria Available** | | | | | | |
| | **One Criterion is Missing** | | | **All Criteria are Missing** | | |
| Female | 1.23*** | 0.88 | 0.86 | 1.29* | 0.92 | 1.45 |
| | (0.07) | (0.09) | (0.09) | (0.14) | (0.55) | (0.54) |
| Hispanic (Ref.) | | | | | | |
| Asian | 0.83 | 1.22 | 1.38 | 1.01 | 0.32 | 0.36 |
| | (0.16) | (0.35) | (0.40) | (0.26) | (0.26) | (0.32) |
| Other | 0.84 | 0.52 | 0.57 | 1.39 | 0.91 | 0.78 |
| | (0.16) | (0.22) | (0.24) | (0.54) | (0.59) | (0.56) |
| Born in the U.S. | 0.71*** | 1.05 | 0.99 | 1.01 | 0.78 | 0.95 |
| | (0.07) | (0.16) | (0.18) | (0.16) | (0.51) | (0.27) |
| FRL | 0.71 | 0.84 | 0.93 | 0.57 | 0.27*** | 0.29 |
| | (0.16) | (0.29) | (0.29) | (0.17) | (0.09) | (0.20) |
| Special Education | 46.41*** | 18.14*** | 19.04*** | 17.38*** | 4.25** | 45.75*** |
| | (13.87) | (3.12) | (3.23) | (6.22) | (1.97) | (11.59) |
| Prior CELDT unavailable† | 3.58*** | 5.97*** | 5.87*** | 21.77*** | 26.99*** | 40.70*** |
| | (0.45) | (1.48) | (1.63) | (12.33) | (10.65) | (14.44) |
| Prior CST unavailable† | 15.70*** | 4.08*** | 4.61*** | 118.95*** | 0.19*** | 1.16 |
| | (2.62) | (0.47) | (0.50) | (52.52) | (0.09) | (0.47) |
| | | | *N* | 4231 | 3430 | 3430 |
| | | | df_m | 6.00 | 6.00 | 6.00 |
| | | | pr2 | .47 | .41 | .41 |

*Note.* All the models include students' cohort and school fixed effects. Model 1 outcome includes three categories for 7th graders: 1) all criteria available (reference group) means they have CELDT/CST scores, 2) student has CELDT or CST scores, or 3) the student does not have any scores. Model 2 outcome includes three categories for 8th graders: 1) all criteria available (reference group) means they have CELDT, CST and GPA versus 2) they are missing either CELDT, CST, or GPA 3) they are missing both scores and GPA. Model 3 outcome only includes the CELDT and CST for 8th graders. † For 7th graders prior scores are 6th grade CELDT and 5th grade CST, and for 8th graders prior scores are 7th grade CELDT and 6th grade CST. Special education students and students who are missing prior years scores (referred unidentifiable students) are more likely to be missing one or all reclassification. Furthermore, in some instances students born in the United States and those that qualify for free and reduce lunch (FRL) are less likely to be missing one or all reclassification criteria.

Appendix Figure 3.1A

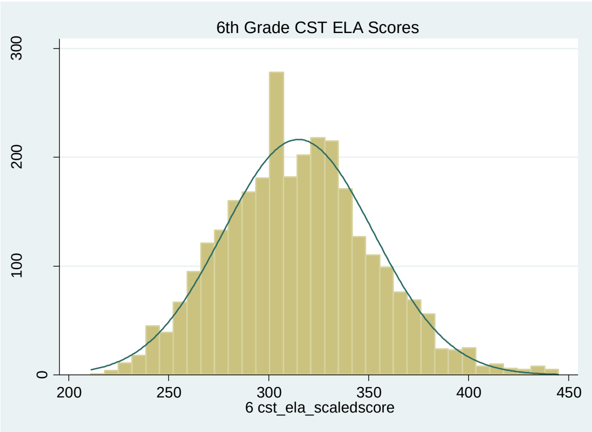
*CELDT Scores (Passing 556 +)*

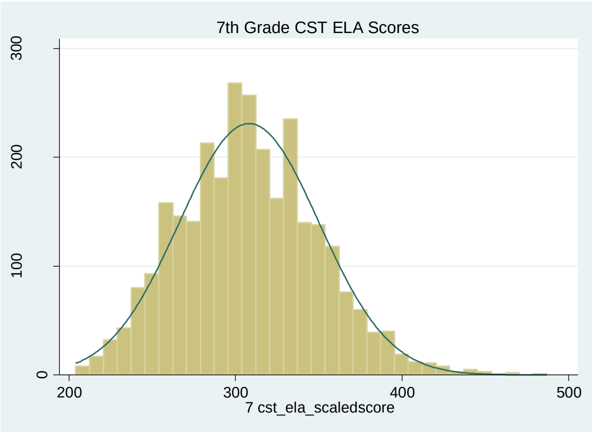



Appendix Figure 3.2A

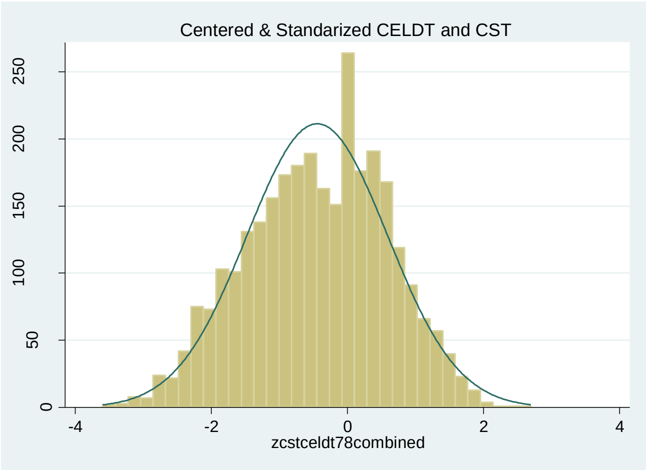
*CST ELA Scores (Passing 325+)*




Appendix Figure 3.3A

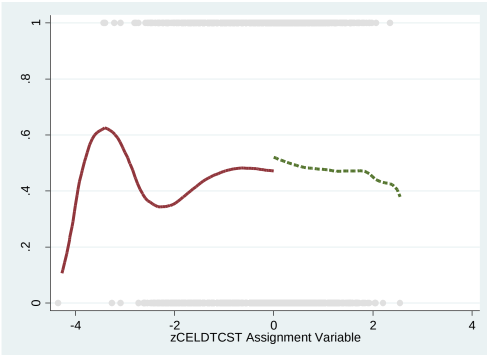
*CELDT and CST ELA (Centered at 556 and 325 and Standardized)*



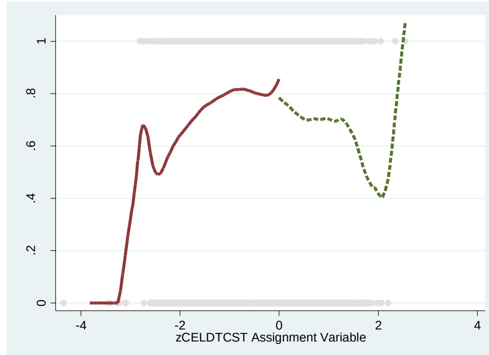
Appendix Figure 3.4A

*Students Demographics and Reclassification Assignment Variable*

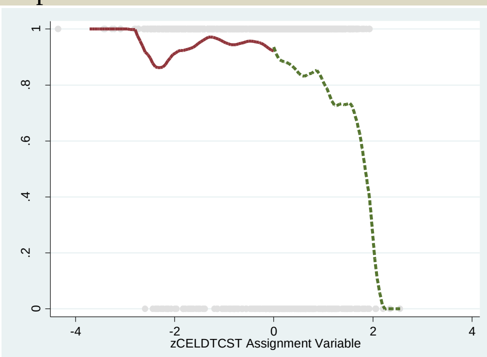
Female



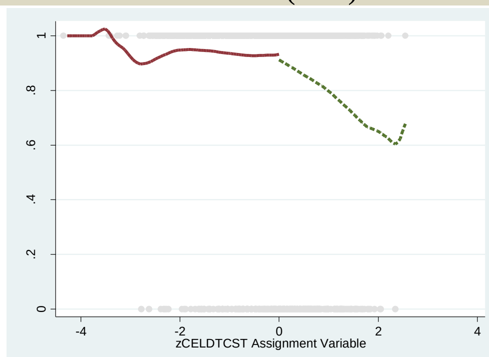

Born in the United States



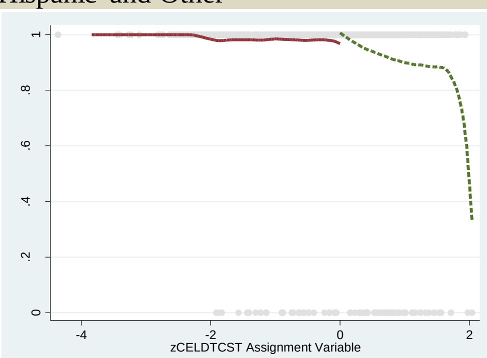

Hispanic and Asian American



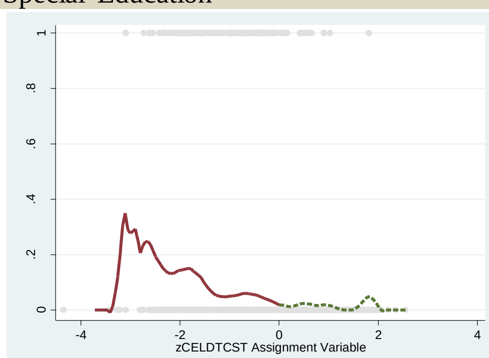

Free or Reduced Lunch (FRL)



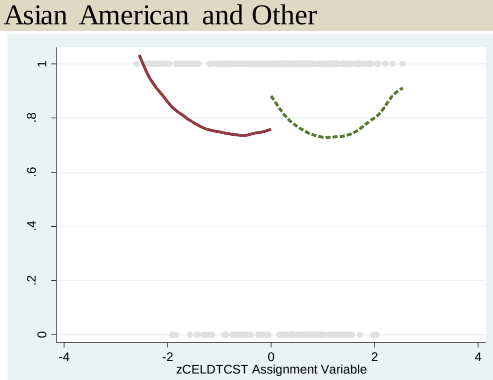

Hispanic and Other




Special Education




Asian American and Other



Appendix Table 3.2A

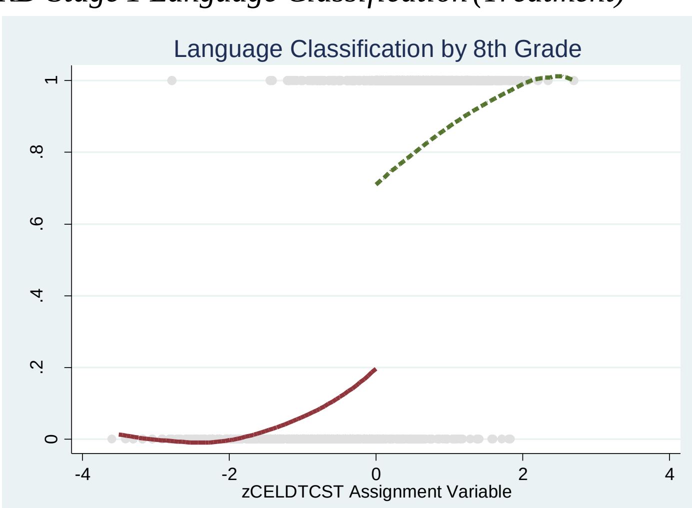
*Eighth Grade Math and English Course Placement by Language Classification*

| | 7th Gr. RFEP N=764 | 7th Gr. ELL N=2,205 | Chi-Square Test |
|---|---|---|---|
| **English Courses** | % | % | $X^2(5, N=2{,}269)=754.6$, $p < .001$) |
| Honors English | 13.4 | 0.7 | |
| Regular English | 77.3 | 37.3 | |
| ELM & Reg. English | 0.0 | 5.9 | |
| Two English (Remedial/Regular) | 7.5 | 46.3 | |
| Two ELD | 0.0 | 7.4 | |
| No English Class | 1.5 | 2.1 | |
| **Math Courses** | % | % | $X^2(5, N=2{,}269)=549.7$, $p < .001$) |
| Geometry | 4.1 | 0.3 | |
| Honor Algebra | 15.8 | 2.3 | |
| Regular Algebra | 43.9 | 18.3 | |
| Basic Algebra | 23.1 | 47.4 | |
| Pre-Algebra | 11.2 | 29.8 | |
| No Math Course | 1.5 | 1.5 | |

*Note*. These are students who remained ELL versus students who reclassified RFEP in seventh grade. ELL students were overrepresented in "Pre-Algebra" and "Basic Algebra", and RFEP students were overrepresented in "Regular Algebra." About 60% of ELL students took either two English mainstream courses or two ELD courses. The rest of ELL students were enrolled in one English course. In comparison 91% of RFEP students were enrolled in one English course.

Appendix Figure 3.5A

*RD Stage 1 Language Classification (Treatment)*



Appendix Table 3.3A

*Language Classification and Passing the CAHSEE*

| CAHSEE ELA (Two Years) N=2,133 | | | | | CAHSEE Math (Two Years) N=2,147 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS | Bandwidths | N | Stage 1 Coefficient (SE) | Stage 2 Coefficient (SE) | OLS |
| 1.3 | 1647 | .45*** (.09) | **-.09 (.12)** | .04 (.02) | 2.1 | 1820 | .51*** (.03) | **-.02 (.06)** | .16*** (.01) |
| 0.6 | 1065 | .39*** (.06) | **-.30* (.14)** | .00 (.03) | 1.0 | 1237 | .42** (.05) | **-.08 (.11)** | .10*** (.02) |
| 0.3 | 659 | .35*** (.10) | **-.66* (.31)** | -.03 (.04) | 0.5 | 750 | .35*** (.07) | **-.12 (.18)** | .05*** (.01) |

*Note*. The assignment variable is based of students CELDT and CST scores. The treatment variable is classified (0) ELL or (1) RFEP by eighth grade. The outcomes are whether tenth graders passed (1) or failed (0) the CAHSEE ELA and math exam. OLS estimates are also provided where the sample size is restricted comparable to the RD models. All the models only include cohort one and two. The coefficients do not vary by cohort.

# CONCLUSION

The United States is a nation that protects the educational rights of all children regardless of their backgrounds, and this includes language classification. In fact, the federal No Child Left Behind (NCLB) act made ELL students a priority, holding states, districts, and schools accountable for ELLs' opportunities to learn, access to educational resources, and reaching high levels of achievement (Hess & Petrilli, 2006). However, the NCLB assumption was "that by the time English Learners get to secondary school, they would already have developed the skills to participate on an equal footing with English proficient students – and that English Learners who are enrolled in middle and high school are more newly arrived immigrants" (Olsen, 2010, p. 6). The work of this dissertation shows that 26% of the Manzanita sixth graders and 38% of the Granada sixth graders were classified as ELL. Of those, respectively, only 28% and 26% were reclassified by the eighth grade. Furthermore, 73% of these language minorities were born in the United States; essentially, most of them were so-called "Long-Term English Language Learners" (LTELLs) and not foreign-born students. This three-study dissertation focuses on middle school ELL students who are less frequently studied than ELL elementary or ELL high school students. In short, my work shows that a district's language classification policy can prevent students from reclassifying; however, remaining classified ELL is not as disadvantageous to students' opportunities to learn as most prior researchers have concluded.

Study one illustrated how two districts with different language classification policies decreased middle school student chances of reclassification. Because language classification categories are socially constructed, I examined the categories themselves before examining their effects. I asked whether ELL students are struggling to become English proficient, or if the classification policies themselves are unnecessarily preventing them from reclassifying. The first study of this work illustrated both districts prevented students from reclassifying regardless of their two different approaches—MUSD administrators stated they favored teacher/parent recommendations, and GUSD administrators declared they privileged test scores. Both districts, however, set higher CST cutoffs than the state mandated, and Granada also required students to pass a district essay exam, thus essentially making it more difficult for ELL students to be considered English proficient. During the interviews, both sets of administrators expressed concern that ELL students may be reclassifying too early and losing essential language support too soon. They also discussed that they felt keeping students classified ELL protected them from academic failure. Ultimately, in both districts the assessments were found to be gatekeepers, where students had to pass these hurdles even before being considered for reclassification.

In both districts, certain students who passed most, or even all, of the requirements were sometimes not reclassified. My work points out that, regardless of the language classification policies in both districts, teachers made the final decisions, and demographic differences continued to persist. In MUSD, male, Hispanic, and low-income students were less likely to reclassify, even after having passed the CELDT and CST. In GUSD, low-income students were less likely to be reclassified when considering students who passed the CELDT, CST, and essay. Furthermore, in both districts exceptions were made for some students who only met the state's lower reclassification requirements. These exceptions were most likely for the district's female, Asian American, and high-income students. GUSD administrators may assume that formulaic placement would lead to fair language classification. Yet, they did not consider that teachers can make exceptions for those that meet the state's minimum requirements. And they also did not consider that certain students were less likely to take the essay test, or, alternatively, to take and fail this exam.

District administrators may be motivated to establish language classification policies to equitably reclassify students, but their classification process is a loosely coupled process. The way administrators establish language classification policies, and how those policies are implemented in actuality are disjointed. In both districts, regardless of district administrator guidelines, English teachers reached conclusions based on their own experiences instead of using assessments scores that represent English proficiency. Thus, we find that educators' interactions and expectations of certain students may influence final reclassification decisions. Teachers may view male, Hispanic, and low-income students negatively, and incapable of succeeding in mainstream classrooms. Prior research demonstrates that the ELL classification can have negative implications for student opportunity to learn (e.g., Edwards, Leichty, & Wilson, 2008; Gandara & Rumberger, 2009; Grissom, 2004; Mosqueda & Maldonado, 2013), therefore, in the second and third study, I assess the effects of reclassifying in middle school. In the next two studies, I only focused on MUSD students because GUSD data did not meet all the RD assumptions.[30] RD models show language classification can affect English and math course placement in some circumstances, but ultimately reclassifying did not increase student educational outcomes.

In study two, I expected ELL students to receive different English courses than RFEP students because the purpose of the language classification is to provide each group suitable and tailored courses to become English proficient. However, a district administrator explained that language classification did not determine eighth-grade English course placement because it was based on prior achievement scores and teacher recommendations. In theory, Manzanita's ELL

[30] I did not use GUSD data in this investigation because it did not meet one of the RD assumptions: RD requires that the assignment variables include continuous variables, and GUSD's essay exam was scored on a scale of 1-4. Furthermore, a frontier RD could not be carried out because the essay exam was more difficult to pass than either the CELDT or CST; I could not conduct an RD with only those students who passed the essay exam.

and RFEP students had an equal opportunity to be placed in one mainstream English course (e.g., regular or Honors) versus two period remedial English courses (e.g., Intensive Literacy, ELD, ELM) that prepared students for mainstream English courses. Typical OLS methods demonstrated that RFEP students were more likely than ELL students to be enrolled in a mainstream English course for all three cohorts. However, RD estimates demonstrated that language classification only determined English course placement in the third cohort. Here, MUSD began to offer a two-period ELM course intended specifically for LTELL. In the first two cohorts of the study, LTELLs were placed in Intensive Literacy courses intended for students with low English skills regardless of language classification.

As for math courses, based on interviews and state policies, I did not expect language classification to determine eighth-grade math course placement. However, prior research demonstrates that ELL students are placed in lower math tracks (Callahan, 2005; Callahan, Wilkinson, & Muller, 2010; Mosqueda, 2010). The OLS models demonstrate that Manzanita's RFEP students were more likely than ELL students to be enrolled in an accelerated math course (e.g., algebra, honors algebra) than a basic math course (e.g., basic algebra and pre-algebra) in all three cohorts. However, again for the first two cohorts, ELL and RFEP students were equally likely to be placed into accelerated math courses, as derived from the RD estimates. Until the third year, academically comparable ELL and RFEP students were put into different math courses based on their language classification categories. In 2012-2013, Manzanita moved away from the algebra-for-all concept. The district explicitly started distinguishing between algebra and pre-algebra, where ELL students were more likely to be placed in the latter.

Clearly, all students should be enrolled and have access to rigorous content courses, along with the proper language assistance to help them address their particular needs. The first two years MUSD educators made English and math course placement decisions based on a student's prior performance, and not the language classification itself. However, they did not offer designed courses for LTELL, and most MUSD students were placed in some type of algebra course. In the last school year, MUSD educators stop placing ELL students in two period Intensive Literacy courses that were not tailored to their particular linguistic needs and started to place them in tailored, two period ELM courses. However, that same year administrators also began placing ELL students in pre-algebra. I found that when ELL students were grouped and placed into appropriate English courses, this, unfortunately, led to them being grouped into lower-level math courses. MUSD educators may be concerned that eighth grade ELLs do not have the proficiency to succeed in algebra, however the ELL category was meant only to be an indicator of the types of required resources necessary to ensure English proficiency. No evidence that pre-algebra courses provide more language support than algebra courses currently exists. Middle school is the first point at which students are sorted into distinct educational tracks, and previous research demonstrates lower tracks may have a lasting influence on high school achievement scores (e.g., Gandara & Rumberger, 2009; Mosqueda & Maldonado, 2013). In the third study, I assessed whether reclassifying in middle school affects student achievement and behavioral outcome, mainly for the first MUSD cohorts.

Study three demonstrates, based on OLS models, that Manzanita's ELL students had lower CST ELA, CAHSEE ELA, CAHSEE math, high school math course placement, attendance, and greater behavioral problems than students who reclassified in seventh or eighth grade. These results coincide with other researchers' work that provided descriptive results based on different measures (Edwards et al., 2008; Gandara & Rumberger, 2009; Grissom, 2004; Losen & Martinez, 2013; Mosqueda & Maldonado, 2013; Saunders & Marcelletti, 2012; Slama, 2014; Uriarte, Lavan, Agusti, & Karp, 2009). However, RD models show, in most cases, middle school ELL and RFEP students did not have different academic and behavioral outcomes. Assigned English and math courses and resources appear to adequately support language minorities. There were two exceptions. RFEP students were less likely than ELL students to pass the CAHSEE ELA portion, and were more likely to be suspended. However, these negative results were only statistically significant based on some bandwidths. MUSD middle school language minorities may be receiving appropriate English and math curricula and resources, and the ELL classification may not be as detrimental as prior researchers have concluded.

In sum, this dissertation demonstrates a district's language classification policies can prevent students from reclassifying, and male, Hispanic, and low-income students are less likely to be reclassified. At first, this may be concerning, but my work also demonstrates that the ELL classification does not necessarily affect student opportunity to learn. For two cohorts, MUSD equitably placed students in appropriate English and math courses, despite their language classification. Only in cohort three were ELL students more likely to be placed in a two period ELM course and pre-algebra, despite their academic abilities. Furthermore, the final study demonstrated language classification did not affect high school academic and behavioral outcomes. The null effects coincide with the only two other studies that use RD models to estimate the effects of language classification in two Southern California school districts (Robinson-Cimpian & Thompson, 2015; Robinson, 2011). Robinson's work shows no differences between language minorities in elementary and middle school, but he found high school RFEP students scored lower than ELL students on the CST ELA exam. He explains lower English scores maybe caused by RFEP removal from English language development courses and the move back into mainstream English courses with no support. Robinson (2011) also shows language classification did not affect high school course placement, or attendance. He also explains null effects does not imply “best” practices because “a null effect cannot tell us whether a third unobserved setting is preferred to the current nonreclassified and reclassified setting” (Robinson, 2011, p. 283). My results agree with Robinson’s work, and although I focused on MUSD (that set higher reclassification requirements than his districts), and I specifically focused on students who reclassified in seventh and eighth grade. MUSD provided another example where ELL and RFEP students may be receiving the appropriate courses and resources; or, minimally, the ELL classification is not harming student achievement and behavioral outcome relative to their peers who reclassified in middle school.

**Generalizability**

Many factors must be taken into consideration when generalizing these results. The findings in the present work are generalizable to school districts with similar classification policies and student body compositions as MUSD and GUSD. Most California districts set higher reclassification requirements than the state mandates, and they give preference to different requirements (e.g., teacher recommendation, course grades, and attendance), thus each of these factors need to be taken into consideration in any future work on the subject (Hill, Weston, & Hayes, 2014). These results may be generalizable to districts outside of California, but educators and researchers must take into consideration that other states set their own language classification policies. Also, these results are based on the CELDT and CST–exams designed especially for California students. Moreover, districts with fewer minority and low-income students tend to have higher reclassification rates (Hanhnel, Wolf, Banks, & Lafors, 2014), and so these results may not be generalizable to districts with fewer minority and low-income students. MUSD and GUSD had a high percentage of minorities, low-income, and ELL students compared with the average California school district. The variability of language classification policies and implementation of those policies, in addition to the variability of student body population, must be considered when generalizing these results.

Furthermore, the effects of language classification may be different in districts with different English and math courses, and with different course placement policies. For example, Granada offered students a wider variety of math courses than Manzanita, and they did not have an algebra-for-all policy. Granada English and math course placement policy was also more formulaic than Manzanita's policies, which set exact cutoff requirements on the CST, the district's benchmarks, and the essay exam. To make generalizations, one should take into consideration that districts may offer language minority students different courses and resources.

The RD analyses tell us little about the extent to which various district language classification policies moderate the effects of educational outcomes in districts that do not follow similar policies. It is only possible to estimate regression discontinuity analyses in districts that have implemented a formula-based placement system. Furthermore, the effects of language classification are generalizable to students who have the required CELDT and CST scores, and those near the cutoff threshold of 556 and 325, respectively. The RD models do not include students without the required assessments, and these analyses only estimate the effects of language classification near the classification threshold, providing limited evidence for those students who scored either very high or very low on the CELDT and CST exams.

There are many aspects to consider when generalizing these results. However, it is important to reiterate that these results coincide with Robinson's (2011) and Robinson-Cimpian and Thompson's (2015) work focusing on different school districts, and used somewhat different

RD methods than the ones used in this dissertation. In general, the ELL classification itself may not explain ELL students' underperformance.

**Policy Implications**

Based on these findings, I recommend our nation's school districts implement a tightly coupled process where administrators and teachers work together to establish and implement the language classification policies. Language classification polices are variable by state and district, making it imperative for administrators to evaluate how their classification policies are implemented, modifying the process if their objectives are not met. Further, as we have seen, district-mandated reclassification criteria can be put into place without any empirical data, supporting the supposition that their assessments accurately measure English proficiency. Complicated and changing reclassification criteria can make the pathway out of ELL classification extremely difficult, placing an undue burden on many students, particularly given the biased methods of entry into the system. On the other hand, undue burden may be caused for students who reclassify too early. In a few instances RFEP students' performance was worse than ELL students. Therefore, district evaluation of the language classification policy should also include the effects on achievement and behavioral outcome. As MUSD results show, language classification and course placement policies may be appropriately working, but these findings were only evident using rigorous statistical methods.

Now more than ever school districts nationwide need to evaluate their language classification and course placement policies for language minorities. Recently, Congress passed the Every Student Succeeds Act (ESSA), requiring all schools to demonstrate that they are improving the English language proficiency of their ELL students. ESSA is intended to strengthen the accountability provisions, and to increase funding targeted at ELL students. With this act, the government will provide more money per ELL pupil, and districts will now have the freedom to allocate those resources as they choose, thus creating a greater financial incentive for school districts to keep students classified ELL. More locally, California has also implemented the Common Core State Standards (CCSS) and the English language development standards, both of which are aimed at improving academic rigor in all subject areas, and increasing the English language requirements for both ELL and non-ELL students. Furthermore, central language classification assessments, such as the CELDT and CST, are now being replaced by the English Language Proficiency Assessments and Smarter Balanced Assessment Consortium (Umansky & Reardon, 2014). These changes can create additional barriers for ELL students to become RFEP, because the standards and assessments are more language intensive, and there is financial gain in keeping students classified ELL; or these changes can lead to more students being reclassified who would continue to benefit from the additional language support.

**Limitations and Future Research**

A number of future research directions are suggested by the work presented here. In future work, I would interview middle school English teachers, middle school ELLs, and the students' parents. The literature on language classification focuses on school district differences, but, in actuality, district administrators' descriptions do not entirely explain how their policies are being implemented. It is important to understand English teachers' roles, and their perspectives on language classification policy. English teachers are asked to review the district's language classification recommendation for each individual student, but, in both districts, teachers prevented students who passed the district's requirement from reclassifying, yet also made exceptions for certain students who only met the state's minimum requirements. It is also important to interview middle school language minorities and their parents to gauge their understanding of language classification policies. The principal research objective here would be to ascertain whether students and parents are aware that the CELDT, CST, and, in Granada, the essay exam, impact ability to reclassify.

Most literature on language reclassification, course placement, and achievement has selection bias limitations. More research needs to be conducted that compares the achievement and behavioral outcomes for ELL and RFEP students. District language classification policies and course placement policies differ by district and within districts, and, therefore, effects may be unique. In fact, in MUSD courses and course placement policies changed for the third cohort. Unfortunately, most of the third study's analyses did not include MUSD's third cohort because they had not reached high school grade levels. For example, I could not estimate the classification effects on the CAHSEE ELA and math scores for the third cohort because they had not yet reached tenth grade. More investigations should compare ELL and RFEP student outcomes using rigorous statistical methods, such as RD, as utilized here. At the same time, researchers must identify when students reclassify, and not aggregate all RFEP students. Here, reclassification in MUSD middle schools may not make any difference because neither ELL nor students who reclassified in middle school are necessarily receiving adequate opportunities to learn rigorous content when compared with EO, IFEP, and students who reclassified in elementary school. ELL achievement and behavioral outcomes should be compared to multiple sets of students under different circumstances.

Additionally, future work needs to examine the intersection between special education and ELL students, particularly in middle school. I found that a number of special education students were classified ELL, but had not taken the required assessments, thus preventing them from reclassifying. Students without the required assessments were also not included in the RD models. Special education students have Individualized Education Programs that determine their individual, required resources and course placement. It was beyond the scope of this dissertation to investigate the types of courses offered to ELL special education students, but this data also should be examined for a complete understanding of middle school ELL students.

As mentioned, many language classification policies are in flux. The CST and soon the CELDT will be replaced. In future work, it would be important to examine the new developing language classification policies to determine if gender, racial, and SES differences continue to exist. Furthermore, the CCSS now requires math teachers to develop all students' English skills in their math courses while simultaneously providing rigorous math content. District administrators may decide to place ELL middle school students with non-ELL students in the same math classrooms now that English development has become a priority for all students. Alternatively, district administrators may determine that ELL students should be put in different math courses in order for them to receive more individualized language development. Thus, administrators may be encouraged to integrate, or further segregate, ELL students in both English and math classes. In turn, these new course placement policies could affect language minorities' achievement and behavioral outcomes, providing different results than those found in the present work.

**Conclusion**

The nation's accountability system includes ELL students, but this inclusion may not be beneficial if their academic needs are not clearly defined or fully addressed. As I have demonstrated here, current language classification policies can prevent male, Hispanic, and low-income students from reclassifying. However, the significance of student language classification itself may not be as important as prior researchers have concluded. Only for one cohort did language classification determine student English and math course placement, when ELLs were placed in remedial English and basic math courses. However, for the other two cohorts, language classification did not affect course placement. For these two cohorts, language classification did not directly affect high school achievement and behavioral outcomes. This provides evidence that ELL classification does not necessarily harm a student's educational outcomes.

References


Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in US schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Edwards, B., Leichty, J., & Wilson, K. (2008). English Learners in California: What the Numbers Say. *EdSource*.

Gandara, P., & Rumberger, R. W. (2009). Immigration, language, and education: How does language policy structure opportunity. *Teachers College Record, 111*(3), 750-782.

Grissom, J. B. (2004). Reclassification of English Learners. *Education Policy Analysis Archives, 12*(36), n36.

Hanhnel, C., Wolf, L., Banks, A., & Lafors, J. (2014). *The language of reform: English learners in California's shifting education landscape*. Retrieved from The Education Trust-West:

Hess, F. M., & Petrilli, M. J. (2006). *No child left behind primer*: Peter Lang.

Hill, L. E., Weston, M., & Hayes, J. M. (2014). Reclassification of English Learner Students in California. *Public Policy Institute of California. Retrieved from www. ppic. org/main/publication. asp*

Losen, D. J., & Martinez, T. E. (2013). Out of school and off track: The overuse of suspensions in American middle and high schools. *K-12 Racial Disparities in School Discipline*.

Mosqueda, E. (2010). Compounding inequalities: English proficiency and tracking and their relation to mathematics performance among Latina/o secondary school youth. *Journal of Urban Mathematics Education, 3*(1), 57-81.

Mosqueda, E., & Maldonado, S. I. (2013). The Effects of English Language Proficiency and Curricular Pathways: Latina/os' Mathematics Achievement in Secondary Schools. *Equity & Excellence in Education, 46*(2), 202-219.

Olsen, L. (2010). Reparable Harm Fulfilling the Unkept Promise of Educational Opportunity for California's Long Term English Learners. *California Together (Research Report)*.

Robinson-Cimpian, J. P., & Thompson, K. D. (2015). The Effects of Changing Test-Based Policies for Reclassifying English Learners. *Journal of Policy Analysis and Management*.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292.

Saunders, W. M., & Marcelletti, D. J. (2012). The Gap That Can't Go Away The Catch-22 of Reclassification in Monitoring the Progress of English Learners. *Educational Evaluation and Policy Analysis*, 0162373712461849.

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.

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.

Uriarte, M., Lavan, N., Agusti, N., & Karp, F. (2009). English Learners in Boston Public Schools: Enrollment and Educational Outcomes of Native Spanish Speakers. *Gaston Institute Publications*. Paper 113.