American Educational Research Journal
Month XXXX, Vol. XX, No. X, pp. 1–34
DOI: 10.3102/0002831214545110

Reclassification Patterns Among Latino English Learner Students in Bilingual, Dual Immersion, and English Immersion Classrooms

Ilana M. Umansky
University of Oregon
Sean F. Reardon
Stanford University

Schools are under increasing pressure to reclassify their English learner (EL) students to "fluent English proficient" status as quickly as possible. This article examines timing to reclassification among Latino ELs in four distinct linguistic instructional environments: English immersion, transitional bilingual, maintenance bilingual, and dual immersion. Using hazard analysis and 12 years of data from a large school district, the study investigates whether reclassification timing, patterns, or barriers differ by linguistic program. We find that Latino EL students enrolled in two-language programs are reclassified at a slower pace in elementary school but have higher overall reclassification, English proficiency, and academic threshold passage by the end of high school. We discuss the implications of these findings for accountability policies and educational opportunities in EL programs.

Keywords: reclassification, English learners, event history analysis, bilingual education, Latino students

---

Ilana M. Umansky is an incoming assistant professor in the College of Education, Department of Educational Methodology, Policy and Leadership at the University of Oregon, 5267 University of Oregon, Eugene, OR 97403; e-mail: *ilanau@uoregon.edu*. Her research uses quantitative methods and sociological theory to examine the educational opportunities and outcomes of immigrant students and students classified as English learners.

Sean F. Reardon is the endowed Professor of Poverty and Inequality in Education and professor (by courtesy) of sociology at Stanford University. His research focuses on the causes, patterns, trends, and consequences of social and educational inequality; the effects of educational policy on educational and social inequality; and applied statistical methods for educational research.

For many students learning English, one of the most important milestones in their educational lives is when they are reclassified from the status of an "English learner" to that of "fluent English proficient." This change in official status marks a student as proficient in English and able to enter the educational mainstream. Although merely an administrative checking of a box, reclassification carries with it profound implications for students. English learner status is designed to support students learning English with specially prepared teachers, content instruction taught with modifications to increase English learner accessibility, English language development classes, and regular monitoring and English language proficiency assessments. Once reclassified, students lose access to these specialized services but gain access to mainstream classes including the full breadth of courses, teachers, and peers.

Many accountability mechanisms, educational services, and policies are organized to promote and incentivize rapid and universal English learner reclassification. Federal Title III reporting requires that schools and districts report the proportion of their English learners (ELs) who attain district benchmarks for English proficiency each year (Ramsey & O'Day, 2010). Some state laws also promote rapid reclassification, as exemplified by a 2006 Arizona law that requires that English learners receive four hours of daily English language development in order to promote reclassification after one year (Gándara & Orfield, 2010).

A growing body of literature suggests that the road to reclassification is a long one, taking many students 4 to 10 years or more (Conger, 2009; Grissom, 2004; Hakuta, Butler, & Witt, 2000; Parrish et al., 2006; Salazar, 2007; Thompson, 2012). These studies also show that reclassification rates are slower among Spanish-speaking students, poor students, and students with lower English proficiency levels (August & Shanahan, 2006; Conger, 2009; Grissom, 2004; Thompson, 2012). Aside from this, we know relatively little about what factors are related to faster or slower reclassification and higher or lower reclassification. In particular, we know almost nothing about how reclassification rates relate to instructional practices and other factors that are under the control of schools and districts. This is important given that schools have the potential to alter reclassification patterns through programmatic changes yet next to no influence over the poverty, ethnicity, or other background factors of incoming students. This article addresses that gap in the literature by examining how reclassification patterns differ between different instructional models—specifically different models of bilingual versus monolingual instruction.

Prior research suggests that students in bilingual classrooms do slightly to moderately better than equivalent students in monolingual English classrooms on English literacy outcomes (August & Shanahan, 2006; Cheung & Slavin, 2012). Several studies suggest that this bilingual advantage, however, may manifest in the medium term; in the short term, English immersion

students may outperform students in bilingual classrooms on English literacy outcomes (Genesee, 2006).

While linguistic instructional program is a factor that is under the direct control of school districts, it is not immediately clear what kind of linguistic environment should result in higher reclassification rates. English immersion classrooms may have higher reclassification rates due to their concentrated focus on English acquisition. Two-language environments (including both bilingual and dual immersion programs)[1] may result in higher reclassification rates due to students' more facile access to content that is provided in their home language or because these students are able to acquire English more effectively due to linguistic transfer from the home language (Cummins, 1991).

We examine Latino English learners in one large urban school district. We focus on Latino English learners because they constitute 90% of English learners nationally and are a diverse and growing segment of public schools (Hugo Lopez, Gonzalez-Barrera, & Cuddington, 2013; Ramsey & O'Day, 2010). Furthermore, Latino ELs face large and enduring achievement and attainment gaps as compared with other students (Gándara & Contreras, 2009; Kao & Thompson, 2003; D. Lee, Perreira, & Harris, 2006; Reardon & Galindo, 2009).

Using discrete-time survival analysis and following nine cohorts of students for up to 12 years (2000–2012) we ask how reclassification patterns differ in English immersion, transitional bilingual, maintenance bilingual, and dual immersion instructional settings. We also examine barriers to reclassification, examining how students in the four programs progress toward linguistic and academic requirements for reclassification. Implications of the findings for policy and practice are discussed.

Prior Literature

Timing and Prevalence of Reclassification and Attainment of English Proficiency

Reclassification marks a student's official change in status from an English learner to a fluent English proficient student. In order to be reclassified, a student must clear one or more established thresholds. One of the complications in understanding reclassification timing and prevalence is that districts and states set their own reclassification criteria. In all states, districts use an English proficiency assessment as a criterion for reclassification, but many states and districts use additional criteria as well, including standardized academic test scores in math or English language arts, grades, and/or teacher and parent approval (Abedi, 2008; Linquanti, 2001; Ragan & Lesaux, 2006). In addition, states use different English proficiency and academic assessments and districts within a state may set different thresholds on those assessments.

Despite these complications, there is a small but growing literature on the timing and prevalence of reclassification among English learners. Due to the paucity of available longitudinal data, many of the studies of reclassification use cross-sectional data. When longitudinal, most studies rely on relatively few years of data, limiting their ability to discern longer term reclassification patterns, especially cumulative reclassification rates. Limitations aside, these studies have established general notions of the rates and timing of reclassification and English proficiency. Studies of timing to reclassification typically find that roughly half of ELs who enter school in kindergarten are reclassified by the end of elementary school, with another 25% to 30% reclassified by the end of middle school (Grissom, 2004; Salazar, 2007; Thompson, 2012; Warren, 2004).

By federal law, all districts must assess English learner proficiency in four domains of the English language: reading, writing, speaking, and listening. While there are no agreed upon definitions of English proficiency or how to measure these domains (Abedi, 2004; Linquanti, 2001), research suggests that oral (speaking and listening) English proficiency typically develops more quickly than English literacy (reading and writing) or academic English (the English skills necessary to be successful in an all-English academic environment) (Conger, 2009; Hakuta et al., 2000; Thompson, 2012). Estimates vary, but oral proficiency may develop after 2 to 5 years while academic proficiency may take 4 to 10 years or longer (Collier & Thomas, 1989; Hakuta et al., 2000).

Student and school background characteristics predict some of the variation in students' timing to reclassification (Conger, 2009; Grissom, 2004; Hakuta et al., 2000; Jepsen & De Alth, 2005). A considerable amount of variation remains unexplained, however, and is probably due to unobservable or unobserved unique individual, school, or district characteristics (August & Shanahan, 2006; Bialystok, 2001).

Latino EL students typically are reclassified at a slower pace than students of other ethnic and linguistic backgrounds (Conger, 2009; Grissom, 2004; Jepsen & De Alth, 2005; Warren, 2004). For example, Grissom (2004) finds that the proportion of Spanish-speaking students reclassified by fourth grade in California is roughly two-thirds that of non–Spanish-speaking students. Much of this disadvantage may be due to greater poverty and social disadvantage faced by Latino immigrant families (Grissom, 2004).

Language of Instruction and Timing and Prevalence of Reclassification

U.S. public schools typically offer one or more of four main instructional models for English learners: English immersion, transitional bilingual instruction (typically K–2 or K–3), maintenance bilingual instruction (typically K–5 or longer), and dual immersion. In English immersion classrooms EL students are instructed solely in English with the goal of promoting English

proficiency and providing academic content that is accessible to ELs. Transitional and maintenance bilingual classrooms are designed specifically for English learners and provide instruction in English and in a target home language. Transitional bilingual programs focus on using the home language to support English acquisition and access to curricular content while maintenance bilingual programs prioritize full bilingualism in English and the home language. Dual immersion classrooms, like bilingual classrooms, offer instruction in English and a target language but student composition includes both English learners and English-only speakers (EOs) with the goal that both groups develop proficiency in both languages. Within each of these four broad models there is extensive heterogeneity including program goals, class composition, instructional techniques, program structure, fidelity to program design, and program quality (August & Shanahan, 2006).

Many studies and meta-analyses, albeit of varying quality, have investigated the relative effects of two-language versus English immersion programs. Taken together, these studies find that there is a significant small to moderate benefit of two-language instruction for English learners on English literacy outcomes (August & Shanahan, 2006; Cheung & Slavin, 2012; Genesee, 2006; Greene, 1997; Slavin, Madden, Calderón, Chamberlain, & Hennessy, 2011; Thomas & Collier, 2002; Willig, 1985).

Several studies examine the question of whether this effect varies based on length of exposure to two-language instruction (Gersten & Woodward, 1995; Saldate, Mishra, & Medina, 1985; Slavin et al., 2011). Typically, these studies suggest that the two-language advantage emerges only after several years of two-language instruction (Genesee, 2006). This finding makes intuitive sense given that many bilingual programs are conducted almost exclusively in the target language in the early years, with the intent of providing a foundation in home language literacy before teaching English literacy.

The dominant theories explaining the benefit of instruction in primary/home language literacy on English literacy outcomes are transfer theory and underlying proficiency theory. Closely related, these two theories suggest that when an individual acquires literacy skills in one language he or she applies those skills to a second language. Furthermore, students learn new literacy concepts more easily and successfully in their primary language than in a new language, and armed with the knowledge in their primary language (or in an underlying linguistic knowledge that transcends any one particular language), they readily transfer what they learn into their new language, English (Cummins, 1991; Goldenberg & Coleman, 2010).

Two studies have compared reclassification between English immersion students and bilingual students. Both find that students in bilingual classrooms are less likely to be reclassified by the end of seventh grade than are students in English immersion programs (Conger, 2010; Thompson, 2012). Thompson (2012), however, finds that this difference disappears by the end of eighth grade, at which time slightly (and insignificantly) more

students from bilingual classrooms have been reclassified than have students from English immersion classrooms. Because neither study had access to students' reclassification patterns beyond eighth grade, it is not clear whether the higher rates of reclassification in English immersion programs persist or are reversed as students progress through high school.

In addition, both studies compare English immersion to bilingual instruction without differentiating between two-language programs. The studies are conducted in two districts, both of which offer transitional bilingual as the predominant two-language program. The results of their studies may reflect patterns unique to transitional bilingual instruction. Finally, differential student sorting into program makes it difficult to disentangle the effect of bilingual instruction on reclassification. Conger (2010) investigates this in her study and concludes that her findings are unlikely to capture the unbiased impact of bilingual instruction on reclassification.

In summary, there is evidence that reclassification may be higher initially in English immersion, compared with two-language programs. Nonetheless, several important questions remain. First, it is unclear whether this pattern changes in middle or high school. Second, it is unclear how different types of two-language programs affect reclassification timing. Finally, the existing studies may not adequately control for differences in the characteristics of students who enroll in different EL instructional programs.

Barriers to Reclassification

The most common requirements for reclassification, as detailed previously, are meeting thresholds for English proficiency and academic achievement (Wolf et al., 2008). The English proficiency requirement for a status of "fluent English proficient" is the most basic and universal criterion for reclassification and is typically assessed by measuring students' English reading, writing, speaking, and listening skills.

The academic requirement is included to ensure that students are academically prepared to succeed in mainstream classes without English language support (Linquanti, 2001). Policymakers and scholars have debated the appropriateness of an academic criterion for reclassification since it introduces a requirement for entry that native English speakers are not held to and because academic performance may, in part, be the result of educational opportunities or lack thereof experienced as an EL (Linquanti, 2001; Linquanti & Cook, 2013).

Recent research suggests that students at the elementary level are most frequently held back from reclassification by English proficiency, rather than academic criteria, while the reverse is true at the secondary level (Robinson, 2011; Thompson, 2012). Among the English proficiency domains, students are most commonly held back by the English reading subtest, though it is worth noting that this last finding may be an artifact of the California

English proficiency test, given that all the studies examining this issue are from California (Hill, Betts, Chavez, Zau, & Volz Bachofer, 2014; Parrish et al., 2006; Robinson, 2011).

Conceptual Framework

The review of the literature suggests several competing hypotheses regarding how reclassification patterns may differ in different linguistic instructional settings. A first hypothesis is that reclassification rates will be higher in English immersion programs because of this program's focused attention on English acquisition. Furthermore, students are exposed to academic instruction in English, which in theory maps closely onto academic assessments since those assessments are also given in English. Finally, if native English speakers comprise part of the composition of English immersion classrooms—as is the case in the district examined here—EL students in English immersion should have ample access to English-speaking peers and the opportunities this affords for meaningful, contextualized interactions in English.

A competing hypothesis is that instructional features of two-language instruction may lead to more successful reclassification patterns. As reviewed previously, experimental and quasi-experimental research suggests that there is a benefit of two-language instruction on English literacy outcomes. Applying this to reclassification we might expect that students in two-language classrooms will have higher reclassification rates compared to students in monolingual instructional environments due to a transfer of literacy skills from Spanish to English. Given existing research that shows that the two-language English literacy advantage may only emerge after several years of two-language instruction, it may be that reclassification rates are initially lower in two-language classrooms but higher by the time students reach middle or high school.

Two other mechanisms may lead to higher reclassification rates in two-language versus monolingual English environments. First, students in two-language classrooms may have more access to academic content because much of that content is offered in a language that they know. This may lead to better academic outcomes and higher pass rates of the academic criterion for reclassification. Second, two-language classrooms may be more likely to create a social and academic environment that is beneficial for ELs. For instance, they may be more likely than monolingual English classrooms to value cultural and linguistic diversity, and teachers and peers may have higher and more favorable expectations of ELs in a two-language environment (Gándara & Orfield, 2010; Harklau, 1994).

There may also be important differences in reclassification patterns between different types of two-language instruction. Maintenance bilingual and dual immersion programs are typically longer term programs than

transitional bilingual. Reclassification rates may be higher in the longer term programs due to greater home language literacy and therefore greater transfer into English. On the other hand there may be little incentive to reclassify students in the longer term programs, which would result in higher reclassification in the transitional bilingual program. Finally, the inclusion of native English-speaking students in the dual immersion program might speed reclassification for ELs due to increased access to contextualized English, or it may slow reclassification if English speaker inclusion shifts the social dynamic toward greater discrimination against English learners.

School District Background Information

The school district examined in this study is a large, diverse, urban school district in California with an enrollment of over 50,000 students. In the 2011–2012 school year 50% of the student population was EL or reclassified as fluent English proficient (RFEP). Unlike many districts whose EL population is predominantly from one language background, this district has a very diverse EL population. Latinos make up a quarter of the student population and just under half of the EL population.

This school district offers a unique opportunity to examine reclassification outcomes by EL instructional program because unlike many districts, it has had in place robust versions of the four main instructional programs for Latino ELs for at least the past 15 years. These four programs include (1) a traditional English immersion program, (2) a Spanish transitional bilingual program, (3) a Spanish maintenance bilingual program, and (4) a Spanish dual immersion program.

Parents of students who speak a language other than English at home can list their preferred programs on the district's enrollment form. The district then uses an algorithm to assign students to programs based on these preferences. In oversubscribed programs, the district relies in part on a randomized lottery system. Program enrollment can be thought of as partially random and partially based on family preferences, neighborhood of residence, and other factors.

Due to a unique set of historical, legal, and social factors, this school district has a strong commitment to providing equitable and high-quality services and opportunities for ELs in all instructional programs. It has actively sought to provide multiple linguistic instructional programs and to give parents choice regarding program enrollment for their children. Over time, the district has increasingly attended to fidelity of program design. The district's EL services are influenced as well by active parent and community organizations, which advocate for strong and effective services for ELs and opportunities for ELs to receive culturally inclusive instruction in their home languages. Nonetheless, ELs, and Latino ELs in particular, perform far below their English-speaking peers, a pattern seen across the United States.

While the district is relatively unique in having four linguistic instructional programs in place, the program models are quite typical of English immersion, transitional bilingual, maintenance bilingual, and dual immersion programs across the country. The English immersion program in this district is designed to teach EL students English as well as provide academic instruction that is accessible to English learners. English immersion classrooms are typically general education classrooms in which ELs share their classroom with mainstream, non-EL students, and instruction is solely in English. Teachers in English immersion need to hold special certification to work with ELs and are expected to use teaching strategies that increase content accessibility to non–English proficient students. English immersion is the largest and most widely available EL program being offered at all school sites.

The goal of the Spanish transitional bilingual program is to develop English proficiency with the use of primary language support and full access to core content through primary language instruction. Only students with Spanish as their home language can enroll in the Spanish transitional bilingual program. Instruction is predominantly given in Spanish in kindergarten (80%–90%) with increasing proportions of English in each subsequent grade. The program ends after the third grade at which point students transfer to an all-English environment.

The objective of the Spanish maintenance bilingual program is for students to develop fluency in both Spanish and English. All students in the maintenance program are Spanish-speaking language minority students, and instruction is 80% to 90% Spanish in kindergarten, transitioning to 50% to 65% English by the fifth grade. Unlike the transitional bilingual program, the maintenance bilingual program continues throughout elementary school. Some middle and high schools have bilingual programs in which students can continue to take core content area classes in Spanish.

The Spanish dual immersion program is designed to serve both EL students and English-speaking students together in the same classroom. The goal of the program is for both groups of students to become bilingual in Spanish and English. As in the maintenance program, elementary school instruction shifts from predominantly in Spanish in kindergarten to half Spanish and half English by the end of elementary. Also similar to maintenance bilingual, some dual immersion students continue with Spanish core content classes in middle and high school.

Teachers in all three of the two-language programs are required to hold certification to teach in two languages, and classrooms use textbooks and other educational materials in Spanish for the subject areas that are taught in Spanish. The district determines which content area classes are taught in which language in every grade.

We did not observe classrooms in this study to assess the fidelity with which the different instructional programs were implemented. However, we did conduct 22 in-depth interviews with teachers and administrators

regarding the district's four linguistic instructional programs. Some of the EL teachers and district administrators suggested that the programs have not always been implemented with perfect fidelity, both because of practical constraints and resource constraints. Some of these constraints may have led to weaknesses in ELD provision, access to learning materials, and availability of effective teachers. We have no means of assessing the accuracy of these descriptions of classroom implementation. The constraints that some teachers and administrators described to us are, unfortunately, typical of large urban school districts. In this study, we cannot determine how they affected EL students or whether they are responsible for any of the differences we observed in the outcomes of students in different EL pathways.

The district is representative of other school districts across the country in its reclassification criteria (Wolf et al., 2008). Beginning in the second grade, students are assessed annually for reclassification. In order to be reclassified, a student must have: (a) an overall California English Language Development Test (CELDT) score of 4 (early advanced) or 5 (advanced) and no subscore (reading, writing, speaking, or listening) below 3 (intermediate) and (b) a California Standards Test in English language arts (CST-ELA) score of 325 or higher (mid-basic).[2] The district has two additional criteria for reclassification, teacher approval and in middle and high school a grade point average (GPA) threshold, but we do not examine those criteria in the article. We do not have data to adequately analyze either GPA or teacher approval, although our interview data suggest that only rarely does either criterion prevent students from being reclassified.

The CELDT is taken annually in the fall, beginning in kindergarten, and the CST-ELA is taken annually in the spring, beginning in second grade. Hence, students are eligible for reclassification twice yearly, first, when CELDT scores come in and again when CST scores come in. The first point at which a student is eligible for reclassification is in the fall of third grade, when the spring of second-grade CST results become available. Students need to pass all criteria in adjacent sittings in order to be reclassified. For instance, if a student passes the CELDT in fall of fifth grade he or she must pass the CST in either spring of fourth grade or spring of fifth grade in order to be eligible for reclassification.

Both the CELDT and the CST change across grades. There is a separate CST test in every grade, meaning the academic proficiency criteria for reclassification increases in each grade. There are separate versions of the CELDT for Grades K–2, 3–5, 6–8, and 9–12.

Data and Methods

Data

For this study we use school district administrative data for a 12-year period spanning from fall 2000 to spring 2012. We include in the sample

nine cohorts of Latino EL students who entered the district in kindergarten between fall 2000 and spring 2009 and who were enrolled in one of the district's four EL programs. We include only kindergarten entrants because these students are much more likely to get a robust dosage of their language program. We do not include incoming kindergartners entering the district after the 2008–2009 cohort because these students do not reach third grade, the first point at which they can be reclassified, by spring 2012 when our data end. In total, 5,423 students meet these criteria. The first cohort, the 2000–2001 cohort, completed 11th grade at the end of our data while the last cohort, the 2008–2009 cohort, completed the 3rd grade.

Table 1 shows descriptive statistics on the analytic sample for the full group and by initial EL program. There are significant differences in the background characteristics of students in the different programs. In particular, students who enroll in the bilingual programs are more likely to be first-generation students who speak Spanish at home, have lower levels of English proficiency, and come from economically disadvantaged families (as measured by eligibility for free or reduced price lunch). Many of these factors are linked to slower reclassification in prior literature, suggesting that there are important selection issues that we will need to take into account in our analysis.

Table 2 shows the proportion of students entering each of the linguistic instructional programs by academic entry year. The English immersion program has the highest initial entry among Latino EL kindergarten enrollees with 38% of the sample. Over time enrollment in dual immersion and English immersion has grown, enrollment in transitional bilingual has remained relatively constant, and enrollment in maintenance bilingual has declined sharply.

The outcome variables in the analysis are, in the first order, reclassification, and secondarily, each criterion for reclassification including CELDT overall and component scores and CST-ELA score. Using these, we can analyze how instructional program predicts both reclassification patterns as well as criteria passage patterns.

Methods

We model the association between instructional programs and the timing of reclassification (and the timing of meeting the various individual criteria for reclassification) using discrete-time event history analysis (Reardon, Brennan, & Buka, 2002; Singer & Willett, 2003). We denote the hazard of reclassification for student i in semester t (the probability that student i is reclassified in semester t, conditional on his or her not having yet been reclassified) as h_{it}. The discrete-time hazard model expresses the logit of h_{it} as a function of semester (indicated by the dummy variables S_{it}^1 to S_{it}^K) and additional variables, including program type (denoted by a vector $\mathbf{P}_i$

Table 1
Descriptive Statistics, Analytic Sample

	Total Sample	Initial Instructional Program				*F* Test Comparing Means
		Dual Immersion	English Immersion	Transitional Bilingual	Maintenance Bilingual	
Number of students	5,423	927	2,061	1,106	1,329	
Student-level variables						
Initial California English Language Development Test (CELDT) proficiency level (1-5)	2.10	2.04	2.26	1.90	1.90	***
Speaks Spanish at home (%)	94	93	89	97	99	***
Female (%)	49	49	46	51	51	*
Born in USA (%)	81	83	84	77	77	***
Born in Mexico (%)	8	6	7	11	11	***
Born in El Salvador (%)	2	2	1	4	2	***
Born in Guatemala (%)	1	1	1	2	1	**
Born in Nicaragua (%)	0	0	0	0	0	
No birth country given (%)	6	6	5	5	8	**
Other birth country (%)	1	1	1	1	1	
Gifted-classified student (%)	13	13	12	12	16	**
Special education-classified student (%)	22	21	23	20	23	
Number of years in district	6.9	6.5	6.6	7.0	7.7	***
Entry age	5.6	5.6	5.6	5.5	5.5	*
Highest grade reached	6.0	5.6	5.8	6.0	6.8	***
Entry year	2004	2005	2004	2004	2003	***
Eligible for free/reduced lunch (%)	97	95	96	99	99	***
School-level variables (for students' first elementary school)						
School size	374	375	363	369	394	***
Latino (%)	46	57	36	43	54	***
Chinese (%)	16	6	19	21	15	***
African American (%)	12	9	15	9	13	***
English learner (%)	48	44	43	52	54	***
Special education (%)	11	11	11	9	12	***
Free/reduced lunch (%)	49	51	51	54	41	***

~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table 2
Proportion of Kindergarten Students Entering Each Instructional Program, by Academic Year

	Initial Instructional Program			
Entry Year	Dual Immersion	English Immersion	Transitional Bilingual	Maintenance Bilingual
2000–2001	7.0	33.2	20.5	39.3
2001–2002	10.0	27.4	19.5	43.1
2002–2003	14.9	35.7	18.4	31.0
2003–2004	15.6	40.9	22.9	20.6
2004–2005	18.5	44.4	18.8	18.3
2005–2006	21.3	33.5	23.8	21.5
2006–2007	20.2	39.2	20.9	19.7
2007–2008	21.3	41.9	19.8	17.0
2008–2009	21.9	42.8	19.3	16.1
Total	17.1	38.0	20.4	24.5

of dummy variables) and student and school covariates measured in the year a student entered kindergarten (the vectors $\mathbf{X}_i$ and $\mathbf{W}_{is}$, respectively):

$$\text{logit}(h_{it}) = \sum_{k=1}^{K} \alpha_k S_{it}^k + \mathbf{P}_i\mathbf{B}_P + \mathbf{X}_i\mathbf{B}_X + \mathbf{W}_{is}\mathbf{B}_W + (\mathbf{P}_i \cdot \ T_{it})\mathbf{\Gamma} + \Delta_y + \Lambda_s \qquad (1)$$

where T_{it} is a continuous time measure (measuring the number of semesters elapsed from fall of third grade), Δ_y is a set of kindergarten entry year fixed effects, and Λ_s is a set of school fixed effects. In this model, the α_k's indicate the baseline hazard rate—the log odds of reclassification in semester k for a student with values of 0 on the other variables in the model. The coefficients of interest are the vectors $\mathbf{B}_P$ and $\mathbf{\Gamma}$, which describe the differences in hazard rates between the programs, as a function of time. We include terms interacting EL program type with time $(\mathbf{P}_i \cdot T_{it})$ in order to test the hypothesis that hazard of reclassification grows or narrows differently across time in each program.

A concern with this method of modeling EL program is that it characterizes students by the EL program they enroll in upon entry into the school district rather than using all the information we have about what program each student is in each year. In supplementary analyses (available on request), we found that the overwhelming majority of students remain in their initial EL program until that program ends by design or they are reclassified.

One of the main issues we want to account for is selection into instructional program. We do this by including a range of student and school background characteristics in the model, as well as cohort and (in our preferred model) school fixed effects. The school fixed effects allow us to examine differences in reclassification patterns within rather than between schools. This allows us to control for one of the main selection issues: selection into different schools based on neighborhood of residence, parent information, and so on. However, the results of the model with school fixed effects only apply to schools where there is more than one linguistic instructional program. The cohort fixed effects control for changes in reclassification patterns over time.

The additional control variables attempt to account for remaining selection issues between programs within schools. Student background controls include various measures of student English proficiency upon entry into kindergarten, student free/reduced price lunch eligibility, home language, country of origin, and other factors. School-level controls include the demographic composition of the school and the proportion of the school that is EL. In addition, we conduct a sensitivity analysis in which we control for parents' listed choices for school and linguistic instructional program.

Although we attempt to control for as much variation in student selection into pathways as we can, our results should not be interpreted as robust causal estimates. Instead, they should be interpreted as a first attempt at establishing different patterns in reclassification by language of instruction in the school district examined. Future research using other quasi-experimental or experimental designs is necessary, as is research looking at other geographic areas and student populations.

Results

Reclassification

As other studies have shown, it takes students many years to be reclassified as a mainstream fluent English student. Figure 1A describes the unconditional probability that a not yet reclassified Latino EL student will be reclassified in each grade level; Figure 1B shows the cumulative proportion of Latino EL students reclassified by grade level. These are based on unconditional descriptive models that include no covariates; they reflect the average observed patterns in the district among the cohorts we include in the models.

Figure 1A illustrates that Latino ELs are most likely to be reclassified at or toward the end of each schooling cycle: 5th grade, 8th grade, and 11th grade. The 5th grade peak is the highest. The 11th grade peak should be interpreted with caution given that only one cohort of students in our sample is present in 11th grade. Cumulatively, 38% of Latino EL kindergarten entrants are reclassified by the end of elementary school, 6 years after

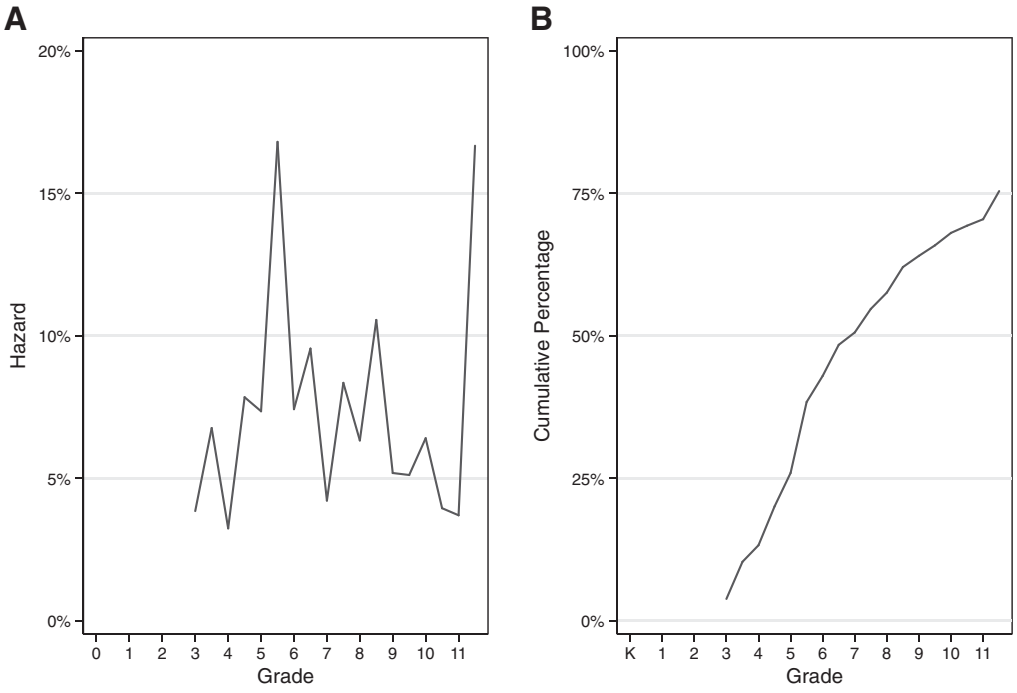

Figure 1. **Unconditional likelihood of reclassification (A) and cumulative reclassification (B), by grade.**

entering the school district. A common definition of a long-term English learner—a variant of which is used as the official definition in the state of California—is a student who is classified as an EL for over six years (Lara, 2011). Over 60% of Latino English learners in this school district become long-term English learners.

Once in middle school, reclassification slows but does not stop. In this district, 62% of Latino ELs have been reclassified by the end of middle school (8th grade), and 75% of Latino ELs have been reclassified by the end of 11th grade.

We estimate that the median time to reclassification for Latino kindergarten entrant EL students is 8 years (meaning that 50% are reclassified by seventh grade). This is somewhat longer than Warren's (2004) estimate of 6.4 years, and Thompson's (2012) estimate of between 5 and 6 years, both of which are based on samples comprised of predominantly, but not exclusively, Spanish-speaking students.

Two other points are worth noting regarding these baseline analyses. First, one in four Latino ELs is not reclassified by the end of 11th grade. While our longitudinal data do not allow us to examine the 12th grade, these patterns suggest that close to 25% of Latino ELs who remain in school through the 12th grade are never reclassified. This is in addition to students who drop out prior to reclassification. Second, our analysis shows very

Table 3
Coefficient Estimates, Reclassification Discrete Time Hazard Models

	Model 1	Model 2	Model 3	Model 4
Dual immersion	–0.53**	–0.50**	–0.58***	–0.46***
	(0.17)	(0.17)	(0.17)	(0.14)
Transitional bilingual	–0.31~	–0.10	–0.16	–0.19~
	(0.17)	(0.17)	(0.16)	(0.12)
Maintenance bilingual	–0.64***	–0.32*	–0.37**	–0.30**
	(0.14)	(0.15)	(0.13)	(0.11)
Dual Immersion × Time	0.08***	0.08***	0.08***	0.08***
	(0.02)	(0.02)	(0.02)	(0.02)
Transitional Bilingual × Time	0.05~	0.04	0.04	0.03
	(0.03)	(0.03)	(0.03)	(0.02)
Maintenance Bilingual × Time	0.08***	0.07***	0.07***	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)
Student and cohort controls		X	X	X
Initial school controls			X	X
Initial school fixed effects				X
Log likelihood	–7,839	–7,625	–7,593	–7,497
N (students)	5,423	5,423	5,423	5,423
N (student-semester observations)	32,103	32,103	32,103	32,103

~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

uneven reclassification rates across grade. The reclassification rate in 5th grade is double that of the 3rd or 4th grades, for example.

Table 3 presents the estimates from the hazard models that analyze reclassification rates by linguistic instructional program. In each model, the reference category is English immersion. The point estimates for the program dummy variables indicate the difference in the estimated log odds of reclassification in a given program, compared to English immersion, in the first semester of third grade (the first time a student can be eligible for reclassification). The interaction terms of each program with the continuous variable time can be interpreted as the average change in the likelihood of reclassification each semester in a given program compared to English immersion. Throughout the table we see a pattern in which two-language programs have lower log odds of reclassification compared to English immersion initially (negative point estimates on the dummy variables) but higher log odds of reclassification over time (positive point estimates on the interaction terms). Model 4 is our preferred model; it includes cohort and school fixed effects and controls for student and school characteristics. Although the sample in that model is restricted to schools with more than one linguistic program, regression results are very consistent with the other three models.

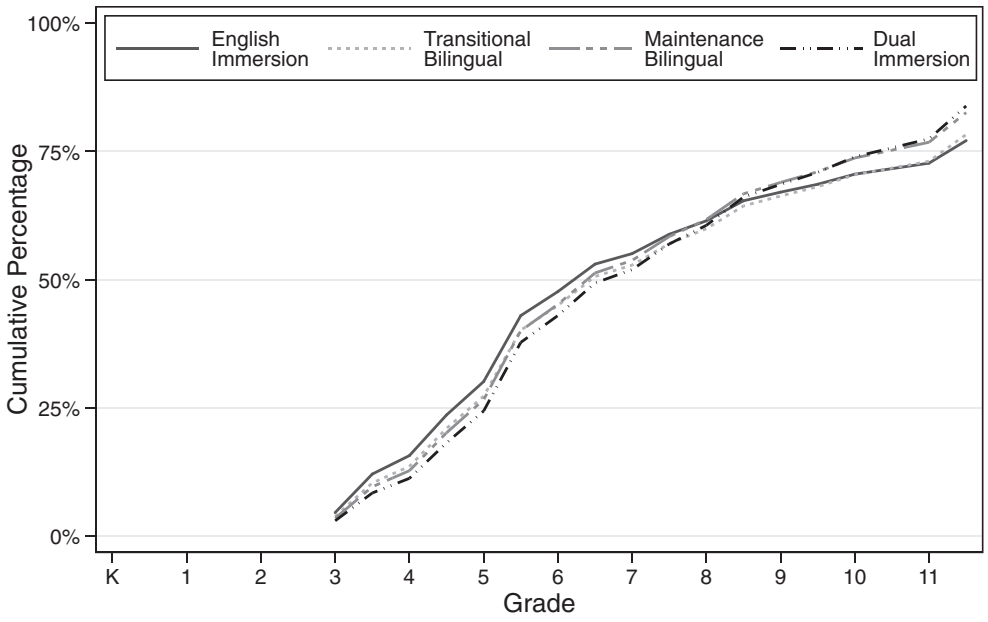

Figure 2. **Adjusted cumulative percentage of students reclassified, by grade and initial linguistic instructional program.**

Because the point estimates from a logit hazard model are log odds they do not lend themselves to easy interpretation without visual accompaniment. Figure 2 shows Model 4 results, reporting the estimated cumulative proportion of reclassified students by grade and initial instructional program. More students are reclassified in the early years in English immersion, but as students reach middle school, reclassification in the two-language programs catches up with and (in maintenance bilingual and dual immersion) surpasses that of English immersion. The differences are relatively modest in magnitude; cumulative reclassification rates are about 7 percentage points higher in maintenance bilingual and dual immersion than in English immersion in the 11th grade.

As noted in the previous data section, English immersion students have characteristics that are linked to higher reclassification rates in our analysis and prior literature. Because of this we would expect the direction of any bias in our results to favor students in the English immersion program. This suggests that the higher cumulative reclassification that we observe in the medium to long term in two-language programs may, if anything, underestimate the differences between the programs.

Barriers to Reclassification

As described earlier, there are six main criteria for reclassification. Five of the criteria come from the CELDT test: the four subtests for reading, writing, speaking, and listening and the overall CELDT score. The sixth criterion is

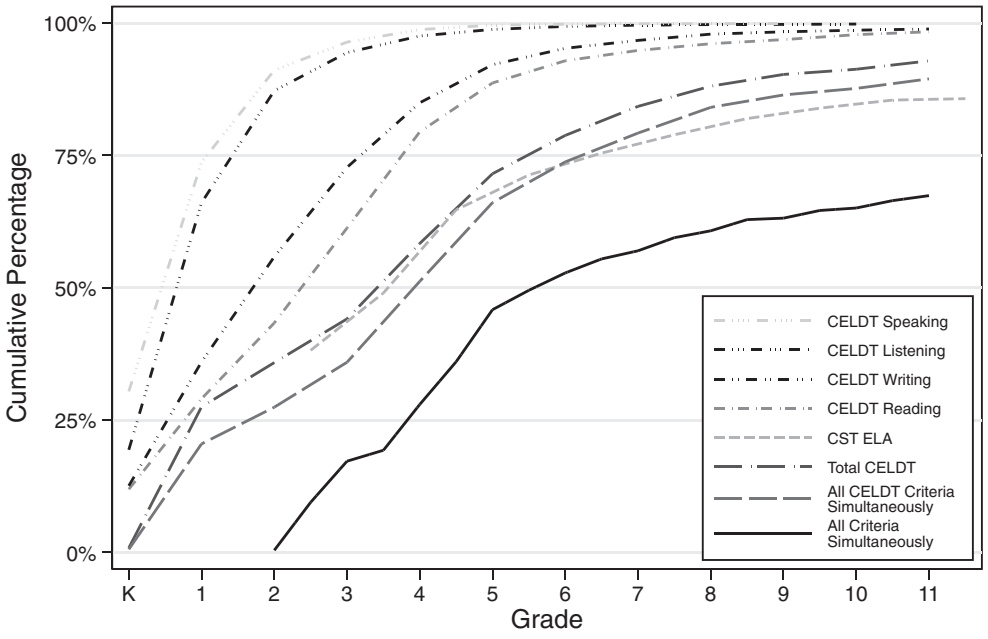

Figure 3. **Unconditional cumulative percentage of students meeting each reclassification criterion.**

the academic measure of English language arts ability (the CST-ELA test). Figure 3 plots the cumulative proportion of students meeting each of the criteria by grade level, based on a baseline hazard model with no controls; that is, it simply reflects the estimated proportions of students who have met each criterion at least once by the end of each semester. Some lines represent more than one criterion, for example, the "All CELDT Criteria Simultaneously" line. The "All Criteria Simultaneously" line in the figure represents the proportion of students who have met all six main reclassification criteria in adjacent sittings by a given semester and are therefore eligible for reclassification.

Prior research suggests that listening and speaking skills advance more quickly than reading and writing skills (Collier & Thomas, 1989; Hakuta et al., 2000). Our results support these findings. We find that roughly 90% of Latino EL students reach proficiency in English speaking and listening by the end of second grade, while it takes until the fifth or sixth grade for the same to be true of English reading and writing.

Recall that the reclassification criterion for each subscale on the CELDT is to reach Level 3 (intermediate) while the overall CELDT score criterion is to reach at least Level 4 (early advanced). This explains why the "Total CELDT" passage line in the graph is much lower than the individual subtests. Fewer than 70% of Latino ELs have reached all CELDT criteria simultaneously

by the end of fifth grade, at which point the pace at which new students reach the threshold slows.

Comparing the CELDT English proficiency criteria with the CST-ELA academic criterion for reclassification, we find that at every grade level through the fifth grade English proficiency is a larger barrier for students than is the academic English language arts academic criterion. Beginning in the sixth grade, the academic criterion becomes a larger barrier to reclassification than English proficiency, although by a relatively small margin. This pattern is also found in other studies (Parrish et al., 2002; Robinson, 2011; Thompson, 2012).

The largest hurdle is passing all criteria in adjacent sittings. A student may pass the reading test in one year but not the writing test or may pass all the English proficiency criteria but not pass the academic criterion. Furthermore, if a student passes all criteria but is not reclassified due to administrative reasons (e.g., not getting all the required signatures on a reclassification form), that student must pass all the criteria again in the next sittings in order to be eligible for reclassification again. Figure 3 shows that just under 50% of Latino ELs meet all criteria in adjacent sittings at least once by the end of fifth grade and that after fifth grade, the likelihood of new students meeting this benchmark declines.

Before moving on to look at barriers to reclassification by linguistic program, we first compare reclassification eligibility to the likelihood of being reclassified. In this analysis, we consider a student to be eligible for reclassification the semester after he or she has cleared all reclassification hurdles in adjacent sittings. This represents the first semester in which we would expect to observe that student reclassified in our data. Figure 4A displays the likelihood of being reclassified and the likelihood of becoming eligible for reclassification in each semester. Figure 4B displays these data cumulatively. One feature stands out. In the early grades more students are eligible for reclassification than are reclassified. But this trend reverses over time. In middle and high school students are more likely to be reclassified than to be reclassification eligible. So while we find that 38% of students are reclassified by the end of 5th grade, nearly half have been eligible. In 11th grade, by contrast, fewer than 70% of students have been reclassification eligible, but three quarters of students have been reclassified.

As previously described, we determined that reclassification patterns are meaningfully different between English immersion and two-language instructional programs. Analyzing patterns in how students in each program progress toward each criterion for reclassification can shed light on those differences in reclassification outcomes.

Table 4 reports results from our preferred model using reclassification criteria as outcome variables. The first column reports results for the five combined English proficiency criteria, and Figure 5 uses these results to plot the cumulative proportion of students we estimate reach the five English proficiency reclassification criteria by grade and instructional program.

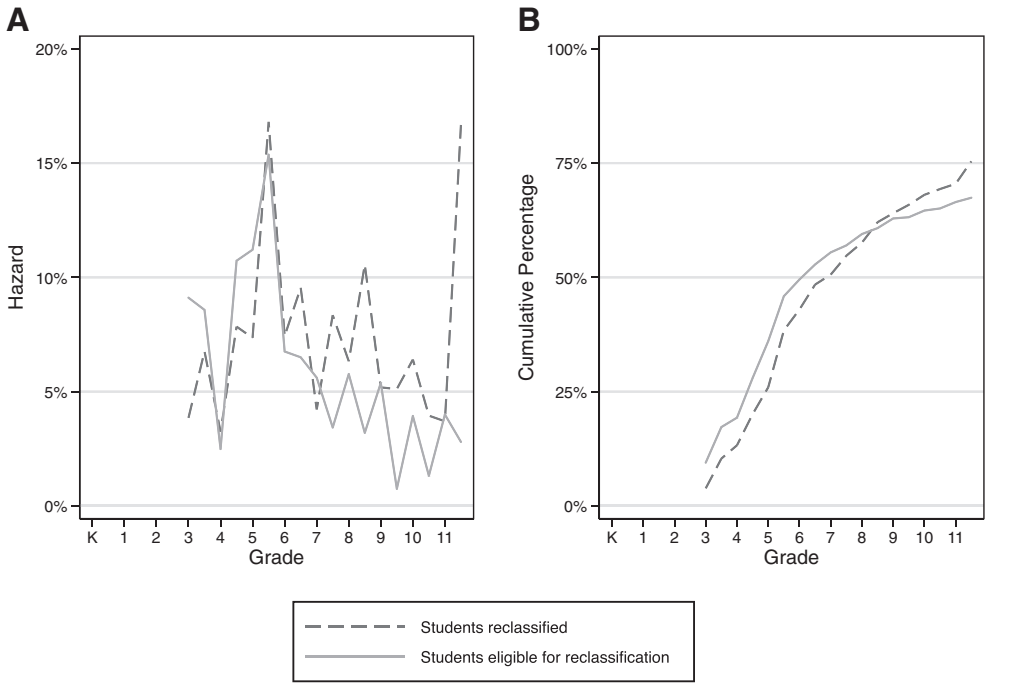

Figure 4. **Comparing reclassification and reclassification eligibility likelihood (A) and cumulative percentage (B), by semester.**

In elementary school, larger proportions of English immersion students reach English proficiency compared to two-language program students. The English immersion advantage is large and relatively persistent across elementary grades. But, as with reclassification rates, that early advantage disappears over time. Both bilingual programs surpass English immersion in cumulative proportion of English proficient students toward the end of middle school. Dual immersion students catch up to English immersion students in high school. The same overall pattern is true when separately examining the speaking, listening, reading, and writing components of the CELDT (results available on request).

The second column of Table 4 and Figure 6 present the same analysis for the academic English language arts (CST-ELA) criterion for reclassification. The findings here are generally the same, albeit with a more modest English immersion advantage in elementary school and with a larger two-language advantage in middle and high school (with the exception of the transitional bilingual program). The crossover also happens earlier for this academic outcome, with dual immersion students outperforming English immersion students by the fifth grade and maintenance bilingual students outperforming English immersion students in the sixth grade. Dual

Table 4
Coefficient Estimates, Reclassification Criteria Passage Discrete Time Hazard Models

	CELDT Criteria[a]	CST Criterion[b]	All Criteria[b]
Dual immersion	–0.70***	–0.46***	–0.87***
	(0.12)	(0.11)	(0.15)
Transitional bilingual	–0.79***	–0.14	–0.45***
	(0.12)	(0.09)	(0.12)
Maintenance bilingual	–1.07***	–0.43***	–1.08***
	(0.11)	(0.08)	(0.13)
Dual Immersion × Time	0.07***	0.18***	0.17***
	(0.02)	(0.02)	(0.03)
Transitional Bilingual × Time	0.09***	0.04~	0.08***
	(0.01)	(0.02)	(0.02)
Maintenance Bilingual × Time	0.12***	0.13***	0.17***
	(0.01)	(0.02)	(0.02)
Student and cohort controls	X	X	X
Initial school controls	X	X	X
Initial school fixed effects	X	X	X
Log likelihood	–8,376	–7,113	–7,044
N (students)	7,870	6,366	6,325
N (student-period observations)	25,544	13,536	35,317

Note. CELDT = California English Language Development Test; CST = California Standards Test.
[a]Point estimates for the program dummy variables for the CELDT analysis indicate the difference in log odds of reclassification for each program, compared to English immersion, in the fall of kindergarten, the first time students take the CELDT.
[b]Point estimates for the program dummy variables for the CST and 'All Criteria' analyses indicate the difference in log odds of reclassification for each program, compared to English immersion, in the spring of 2nd grade, the first time students take the CST and can become eligible for reclassification.
~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

immersion students, in particular, show a clear and large advantage over English immersion students throughout middle and high school.

Figure 7 and the last column of Table 4 show the cumulative proportion of students meeting all reclassification criteria by grade and instructional program. This represents full eligibility for reclassification. Here, again, we see the consistent pattern of gaps between programs reversing in direction over time. What accounts for the gap reversal is that English immersion students reach a virtual plateau as they enter middle school, while students in the two-language programs—especially the dual immersion and maintenance programs—continue progressing once they enter middle school. The dual immersion program, in particular, has thirteen percentage points more

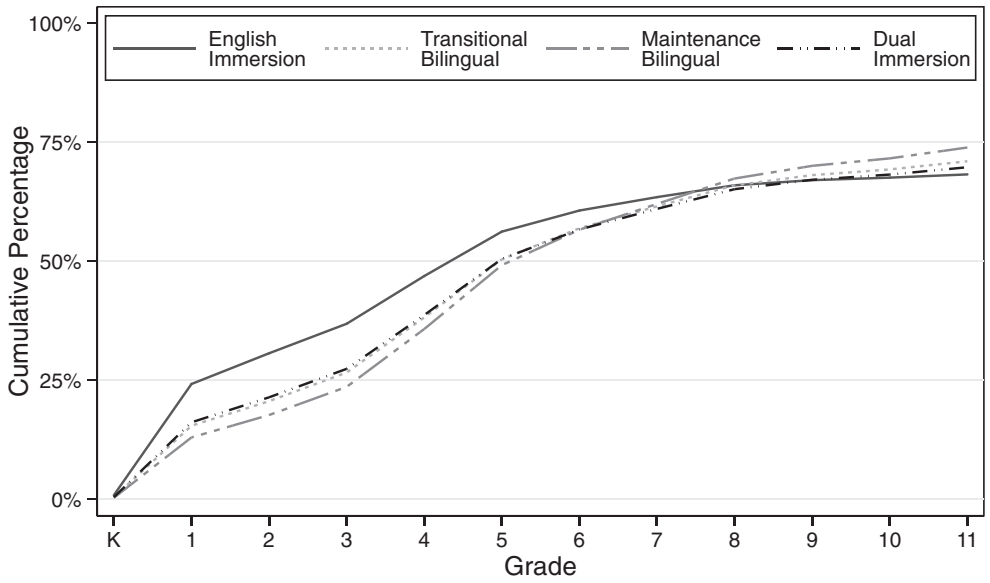

Figure 5. **Adjusted cumulative percentage of students meeting all California English Language Development Test (CELDT) subtest criteria simultaneously, by grade and initial linguistic instructional program.**

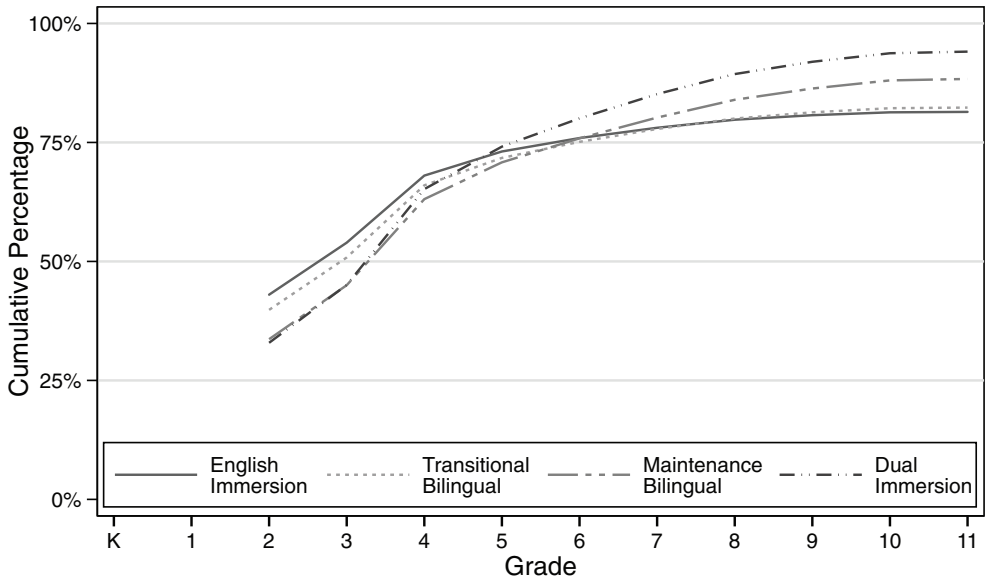

Figure 6. **Adjusted cumulative percentage of students meeting English language arts California Standards Test (CST) criterion, by grade and initial linguistic instructional program.**

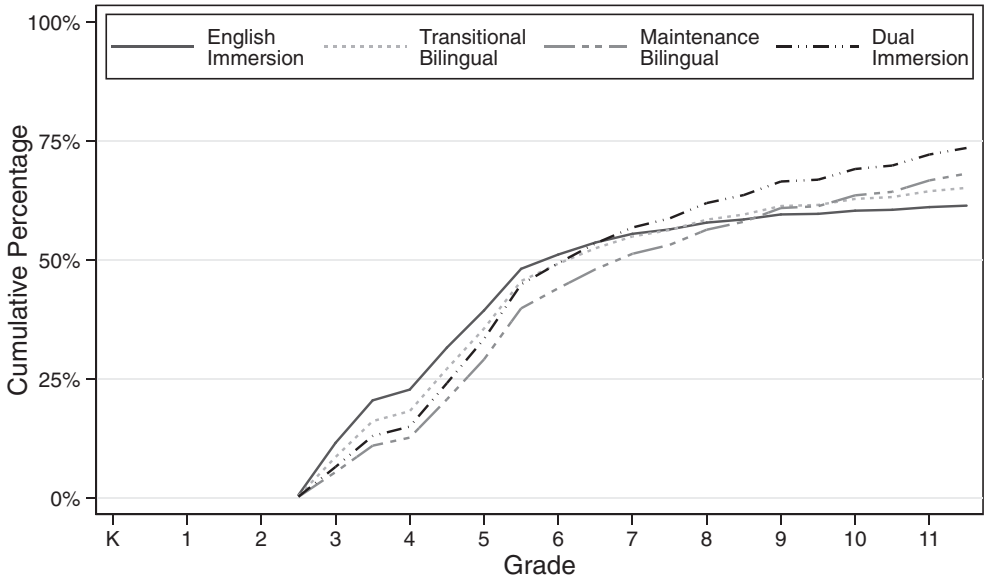

Figure 7. **Adjusted cumulative percentage of students meeting all reclassification criteria simultaneously, by grade and initial linguistic instructional program.**

students reaching reclassification eligibility by the end of high school, compared to English immersion students.

Sensitivity Analysis

As discussed earlier, one concern in a study like this is differential selection into the various instructional programs. In the main analysis we attempt to limit or remove selection bias through student and school controls and school and cohort fixed effects. As a check on our main results we conduct a sensitivity analysis using data on parents' choices for school and program for their children (recall that in this district parents can rank order their school and program choices).

In this model we include fixed effects for parents' first choice for school and linguistic instructional program. In doing so, we compare outcomes among students whose parents selected the same school and program but who, due to the student assignment and the lottery system, were assigned to different programs. We do not include these fixed effects in our main analysis because we have parental choice data for only five of our nine cohorts of ELs.

The results (available on request) support the findings in our main analysis: English immersion students show an early advantage in all outcomes while students in two-language programs catch up and in some instances surpass their peers in all English environments in later grades. The results differ from results in our main analysis in two notable ways. First, results

controlling for parental preferences tend to show a two-language advantage earlier than in our main analysis. For example, controlling for parental preference, Latino EL students in transitional bilingual reach their English immersion peers' cumulative reclassification rates in 4th grade, as compared to around 10th grade in the main analysis.

Second, while the results of our main analysis show that long-term reclassification rates are highest in the dual immersion and maintenance bilingual programs, the estimates from the models including parental choices indicate that reclassification rates are highest among students enrolled in the transitional bilingual program. Students in the maintenance bilingual program have the highest cumulative rates of CST and CELDT criteria passage. These differences across analyses in relative program outcomes are not large, but they do suggest that our primary results do not fully control for differences among students who enroll in the four programs. Nonetheless, all of the models indicate that EL students in the two-language programs generally catch up, or surpass, their peers in the English immersion program by middle school.

Discussion

The conclusions of this article build on those of prior work on reclassification (Conger, 2009; Hakuta et al., 2000; Parrish et al., 2006; Thompson, 2012; Warren, 2004), language of instruction (August & Shanahan, 2006; Genesee, 2006; Goldenberg & Coleman, 2010), and the educational experiences of Latino students (Gándara & Contreras, 2009; C. Lee, 2006). Our results reflect Latino ELs in one large urban district and we caution that results may differ for other populations and in other districts or states.

Timing and Rates of Reclassification

Our study confirms prior research findings that it takes most EL students many years to become proficient in English and to be reclassified into mainstream status in school. Specifically, we find that in this district it takes eight years in school for 50% of Latino ELs to be reclassified into mainstream status. A full 60% of Latino ELs become long-term ELs and roughly a quarter are never reclassified.

Expectations that students can be appropriately reclassified after a year or two of EL services, such as Arizona's model of one year of intensive ELD, appear to be extremely unrealistic (Gándara & Orfield, 2010). Even in an all-English environment, in a program designed to promote rapid acquisition of English, we find that only about 40% of students are reclassified by the end of fifth grade and 45% have not been reclassified as they enter high school.

For students in all instructional programs, English or two-language, reclassification slows in middle school. This is likely for a number of reasons. Chief among them are that students with the highest levels of English

proficiency and academic skills are most likely to be reclassified in the early grades, leaving a population of EL students with lower levels of English proficiency and academic skills over time (Hopkins, Thompson, Linquanti, Hakuta, & August, 2013; Saunders & Marcelletti, 2012). Another reason that reclassification rates may slow beginning in middle school is increased EL tracking practices. Recent studies suggest that EL students are disproportionately placed into low-level academic classes once they reach middle and high school (Callahan, 2005; Callahan, Wilkinson, & Muller, 2008, 2010; Kanno & Kangas, 2014). If students are not exposed to academic rigor they may fail to meet the academic reclassification criterion. Indeed, we see that the academic criterion becomes the largest barrier to reclassification once students enter middle school. Likewise, if ELs are isolated in classes with few native English speakers and low-level academic English they may fail to meet English proficiency reclassification criteria.

Our findings also suggest that reclassification tends to happen in predictable moments—namely, the end of each schooling cycle (5th, 8th, and 11th grades) and that reclassification patterns often do not parallel reclassification eligibility patterns. Higher reclassification in 5th grade corresponds to higher reclassification eligibility in that grade, but the same is not true in 8th or 11th grade. Because we do not see parallel "peaks" in the likelihood of students becoming reclassification eligible and the likelihood of being reclassified, it suggests that factors other than reclassification criteria may come into play in reclassification decisions. For example, the reclassification peak in 8th grade may reflect a push on the part of teachers or administrators to reclassify students prior to entering high school. Another factor may be program design. In the dual immersion and maintenance bilingual programs teachers may have little incentive to reclassify EL students prior to the 5th grade given that students remain in the program through 5th grade regardless of their reclassification status.

Reclassification and Linguistic Instructional Program

Many factors impacting timing to reclassification and English proficiency are not under the control of schools and teachers. Other factors, however, are. One of those factors is the type of linguistic instructional environment EL students are exposed to. We find that reclassification patterns differ meaningfully between English immersion and two-language programs with the pattern that English immersion students have more favorable outcomes in elementary grades while students in two-language programs catch up and surpass their English immersion peers in middle school. This pattern holds not only when looking at reclassification as an outcome, but also when examining academic ELA achievement and English reading, writing, speaking, and listening proficiency. Indeed, the apparent benefits of the two-language programs on academic ELA achievement and English proficiency are larger than on reclassification.

These findings of a medium- to long-term academic and linguistic advantage for students in two-language programs potentially contribute to a large body of research on the broad benefits of bilingualism, ranging from neurological to economic advantages (Bialystok, 2011; Callahan & Gándara, 2014; Craik, Bialystok, & Freedman, 2010; de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012; Kovács & Mehler, 2009). Our study was unable to measure bilingualism due to a lack of data on student Spanish ability. For this reason we do not know if our results are due to bilingualism per se or are due to another factor or factors including increased accessibility to core content, more welcoming and engaged learning environments, or some other benefit of the two-language programs in this district.

Our results are consistent with prior reclassification literature that finds an English immersion advantage in the early grades (Conger, 2010; Thompson, 2012). This study adds to existing knowledge by showing that the English immersion advantage for this population disappears and reverses by the time students are in high school. Moreover, the results from our sensitivity analysis using parental preference data suggest that the two-language advantage may appear even earlier—in late elementary or middle school. We also show that important differences exist between different types of two-language instruction in this population, differences that can be hidden when grouping all two-language programs together and when looking only at reclassification as an outcome rather than also examining progress toward reclassification requirements.

Importantly, our analysis suggests that at least in this district, Latino EL students sort into instructional programs in ways that bias simple estimates of the programs' differential effects. Descriptive statistics on students in the four programs show that students in the two-language programs have characteristics that are associated with lower reclassification and related outcomes. While we are able to control for some of these differences, the results from our sensitivity analysis using parental choice data suggest that meaningful differences in student sorting remain in our initial models.

Across both the main analysis and the sensitivity analysis, however, we consistently find that students in the maintenance bilingual program are the most likely to reach English proficiency thresholds. These students are also among the most likely to be reclassified, to reach the academic ELA threshold, and to become eligible for reclassification.

Our findings support theory and research on second language acquisition and bilingual instruction. Transfer theory and underlying proficiency theory both suggest that acquiring a solid foundation in one's native language supports one's ability to acquire proficiency in a second language (Cummins, 1991; Goldenberg & Coleman, 2010). Studies have found a transfer effect of home language to English in areas including phonological awareness (López & Greenfield, 2004), vocabulary (Ordóñez, Carlo, Snow,

& McLaughlin, 2002), and reading (Paez & Rinaldi, 2006); research also suggests that linguistic transfer may be particularly high between Spanish and English due to their similar alphabet systems (Bialystok, Luk, & Kwan, 2005).

These theories may explain why we see sustained linguistic and academic growth among students in two-language programs and superior outcomes in middle and high school compared to English immersion students. They may also explain the inferior outcomes in two-language programs in elementary grades. These programs focus on home-language instruction in the early grades, perhaps delaying students' English acquisition in those years.

Furthermore, the fact that we observe less of an academic disadvantage than a linguistic disadvantage among two-language students in elementary may suggest that students in two-language programs benefit academically from content instruction in the home language. Language arts content, as well as other academic content, is likely to be more accessible to students when provided in their native language.

Unlike two-language program students, Latino EL students in English immersion in this district reach a virtual plateau when they enter middle school. Two hypotheses may explain this phenomenon. First, English immersion students may face difficulties in acquiring full English proficiency and advancing academically without a strong base in their home language. A second, and possibly complementary, hypothesis is that English immersion ELs in middle and high school face more isolation, stigma, and detrimental tracking than ELs in two-language programs. This may particularly apply to Latino ELs in English immersion settings, as research has found that Latino immigrants often face more acute stigma and hostile societal reception than some other linguistic and ethnic immigrant groups (Portes & Rumbaut, 2006).

Barriers to Reclassification

Across linguistic programs, we find that the largest barrier to reclassification in elementary school is English proficiency, as measured by the CELDT. By middle school, however, the academic (CST) requirement for reclassification becomes the biggest hurdle. This pattern has been found elsewhere (Abedi, 2008; Robinson, 2011). Comparatively low CELDT threshold passage in elementary school may reflect the difficult nature of acquiring a second language and the amount of time it takes to master that language. On the other hand, it may also reflect potential problems with the CELDT such as overly high reclassification cut-points or a lack of correspondence between what the test measures and what it should measure in order to predict student readiness to exit EL status (Department of Education, 2011; García Bedolla & Rodriguez, 2011).

There are two likely explanations for why passage of the CST criterion appears to get more difficult with grade level. First, students take a different

CST test in every grade, and the tests become increasingly advanced with each test. Second, the CST may become comparatively more difficult over time because EL students fall behind academically as they progress through school either due to extensive time devoted to language instruction (Valdés, 1998; Valenzuela, 1999) or a lack of access to rigorous content instruction (Callahan, 2005; Dabach, 2009; Kanno & Kangas, 2014).

A related finding is that there is a sizable discrepancy between the cumulative proportion of Latino EL students who are eligible for reclassification and those that are actually reclassified. Comparing these two figures over time we find that more students are eligible than are reclassified in the elementary years and more students are reclassified than are eligible in the high school years. This pattern has been noted before and deserves future research (Abedi, 2008). In our sample, 50% of Latino EL students have reached the CST and CELDT reclassification criteria in adjacent sittings by the end of fifth grade; only 38% of Latino EL students have been reclassified by then, however. Eligible students may not be reclassified because (a) teachers decide to hold them back because they don't think they are ready for general education classes, (b) there is a monetary incentive to not reclassify students since EL students receive additional funds, (c) there is administrative slippage in which the appropriate steps are not completed for student reclassification, and/or (d) there is little motivation to reclassify students in elementary school because only rarely does it result in changes in classroom or instructional placement.

This finding has potentially serious policy implications given prior research suggesting that carrying the EL classification, particularly among students with relatively high English language and academic skills, can be detrimental to students' access to courses and academic outcomes (Callahan et al., 2008; Callahan, Wilkinson, Muller, & Frisco, 2009; Umansky, 2014).

Conclusion

While timing to reclassification is certainly a reflection of the quality of an EL student's educational experience, we find that it also likely stems from the type of instructional program a student is exposed to. Two-language programs, particularly those that focus on home language acquisition in the early grades, may result in longer durations of EL status prior to reclassification. But this study further shows that remaining an EL longer is not associated with inferior outcomes in the long term. We find that EL students in two-language programs—transitional bilingual, maintenance bilingual, and dual immersion—have a higher long-term likelihood of becoming proficient in English, meeting an academic ELA threshold, and being reclassified. In fact, reclassification outcomes between programs understates the long-term two-language program advantage. EL students in two-language programs have a larger advantage both in terms of academic outcomes and

reclassification eligibility than is evident when looking only at the proportion of students reclassified in each program.

This finding has significant implications. First, it underscores the importance of examining the efficacy of two-language instruction using longitudinal methods. Many important educational outcomes—both linguistic and academic—may show very different patterns when looking at the short versus the long term. Current federal accountability systems including Annual Measureable Achievement Objectives (AMAOs) under Title III regulations may bias toward more favorable results in English immersion over two-language instructional programs despite evidence that two-language classrooms produce more beneficial results for more students in the longer term.

Furthermore, rather than one set of goals for all instructional models, it may make sense to have differentiated benchmarks based on language of instruction and theories of development. For example, goals for English immersion programs might be stable annual growth in English proficiency and academic outcomes across grades while goals for bilingual and dual immersion programs would have slower English proficiency growth levels in early grades and more rapid acceleration in later grades. Measures of home language proficiency and academic outcome measures in the home language would also be ideal, especially in the early grades and depending on the unique theories behind different two-language instructional models.

A second important implication of this study relates to English learners' academic opportunity to learn. Rather than prioritizing reclassification so that students learning English can enter mainstream classes, EL status should be designed such that it does not inhibit full access to rigorous content and interaction with English-speaking peers. In practice, this would mean ensuring that ELD classes do not prevent enrollment in other classes, that teachers are properly prepared to teach ELs within mainstream classes, that English language instruction is embedded in content area classes, that less concern is placed on ELs solely based on how long they have been classified as such, and that ELs can and do enroll in any and all classes (Bunch, 2013; Bunch, Kibler, & Pimentel, 2012; Lara, 2011; Moschkovich, 2012; Olsen, 2010).

This study has significant limitations that we hope will be addressed in future research. First, the study presents something of a "black box" view of instructional programs. We do not know the specific pedagogical details or mechanisms in the two-language instructional models that result in higher eventual reclassification rates. Bilingual and two-language instruction can vary widely in implementation (August & Shanahan, 2006), and future research should disentangle what characteristics of two-language programs are responsible for the patterns we observed. Second, although we controlled for important student and school covariates and included cohort and school fixed effects in order to tackle selection issues into different instructional programs, this study does not provide as strong a causal warrant as would a randomized experiment or rigorous quasi-experimental design. Third, questions

remain regarding barriers to reclassification. In particular, future research should examine the role of teacher approval as a criterion for reclassification and the relationship between reclassification eligibility and reclassification. Finally, this study does not examine how reclassification impacts students.

In research and practice there is an implicit assumption that the more quickly students are reclassified, the better the academic and linguistic outcome. Faster reclassification, according to this underlying belief, implies more effective instruction and better-served English learners. This study shows that the speed with which students are reclassified is not necessarily a good indicator of how well students progress linguistically or academically. The median amount of time a student in dual immersion takes to be reclassified, for example, is a full semester more than that of an English immersion student. But by the time a dual immersion student reaches the ninth grade, he or she is over 10 percentage points more likely to meet grade-level academic requirements than is the student in English immersion. Moreover, his or her English proficiency level is higher as well.

Rather than focusing on rapid reclassification, the findings of this study point policymakers and practitioners in an alternate direction: ensuring high-quality instruction and full, rigorous access to the curriculum regardless of language status. If exiting EL status is a de facto requirement for quality instruction and access to content, then EL students will continue to struggle in school with large achievement gaps between themselves and their non-EL counterparts (Callahan, 2005; Fry, 2007; Gándara & Contreras, 2009; Kanno & Kangas, 2014; Reardon & Galindo, 2009; Valdés, 1998). If, instead, EL students are ensured quality instruction and full access to content, longer periods spent in the EL classification could actually result in higher linguistic and academic outcomes by the end of high school.

Notes

The authors are grateful and indebted to the following individuals for their thoughtful feedback and reflections: Kenji Hakuta, Claude Goldenberg, Martin Carnoy, Rachel Valentino, Karen Thompson, Robert Linquanti, and Patrícia Gándara. In addition, we extend deep thanks to several individuals at the school district examined in this article. We will refrain for mentioning them by name in order to protect district anonymity, but these individuals provided invaluable insights into data features, program characteristics, and interpretation of findings. Any remaining errors are solely those of the authors. This study was funded in part through two grants from the Institute of Education Sciences, award numbers R305B090016 and R305A110670. The views expressed in this article are those of the authors and do not necessarily represent the opinions and positions of the Institute of Education Sciences or the U.S. Department of Education.

[1]In this article we use the term *two-language* to refer to bilingual programs that teach English learners (ELs) in both their home language and in English and dual immersion programs that teach both native English speakers and English learners in the same classrooms using instruction in English and the EL student's home language.

[2]The California English Language Development Test (CELDT) measures California's English language development standards, which are organized around the concepts of (a) interacting in meaningful ways, (b) learning about how English works, and (c) using

foundational literacy skills. The California Standards Test in English language arts (CST-ELA), by contrast, is a test of California's English language arts standards, including aspects such as literary response and analysis, reading comprehension, writing strategies, and genres.

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, *33*(1), 4.

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

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge, UK: Cambridge University Press.

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*(4), 229–235.

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, *9*(1), 43–61.

Bunch, G. C. (2013). Pedagogical language knowledge preparing mainstream teachers for English learners in the new standards era. *Review of Research in Education*, *37*(1), 298–341.

Bunch, G. C., Kibler, A., & Pimentel, S. (2012). Realizing opportunities for English learners in the Common Core English language arts and disciplinary literacy standards. *Commissioned Papers on Language and Literacy Issues in the Common Core State Standards and Next Generation Science Standards*, *94*, 1.

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

Callahan, R., & Gándara, P. (Eds.). (2014). *Bilingual advantage: Language, literacy, and the labor market*. Clevedon, UK: Multilingual Matters.

Callahan, R., Wilkinson, L., & Muller, C. (2008). School context and the effect of ESL placement on Mexican-origin adolescents' achievement. *Social Science Quarterly*, *89*(1), 177–198.

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.

Callahan, R., Wilkinson, L., Muller, C., & Frisco, M. (2009). ESL placement and schools: Effects on immigrant achievement. *Educational Policy*, *23*(2), 355–384.

Cheung, A., & Slavin, R. E. (2012). Effective reading programs for Spanish-dominant English language learners (ELLs) in the elementary grades: A synthesis of research. *Review of Educational Research*, *82*(4), 351–395.

Collier, V. P., & Thomas, W. P. (1989). How quickly can immigrants become proficient in school English? *Journal of Educational Issues of Language Minority Students*, *5*(1), 26–38.

Conger, D. (2009). Testing, time limits, and English learners: Does age of school entry affect how quickly students can learn English? *Social Science Research*, *38*(2), 383–396.

Conger, D. (2010). Does bilingual education interfere with English-language acquisition? *Social Science Quarterly*, *91*(4), 1103–1122.

Craik, F. I., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: Bilingualism as a form of cognitive reserve. *Neurology*, *75*(19), 1726–1729.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge, UK: Cambridge University Press.

Dabach, D. (2009). *Teachers as a context of reception for immigrant youth: Adaptations in "sheltered" and "mainstream" classrooms*. Unpublished manuscript.

De Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism enriches the poor enhanced cognitive control in low-income minority children. *Psychological Science*, *23*(11), 1364–1371.

Department of Education. (2011). *A comparison study of kindergarten and grade 1 English-fluent students and English learners on the 2010–11 edition of the CELDT*. Sacramento, CA: State of California.

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

Gándara, P., & Contreras, F. (2009). *The Latino education crisis: The consequences of failed social policies*. Cambridge, MA: Harvard University Press.

Gándara, P., & Orfield, G. (2010). *A return to the Mexican room: The segregation of Arizona's English learners*. Los Angeles, CA: Civil Rights Project/Proyecto Derechos Civiles. University of California, Los Angeles.

García Bedolla, L., & Rodriguez, R. (2011). *Classifying California's English learners: Is the CELDT too blunt an instrument?* Berkeley, CA: Center for Latino Policy Research.

Genesee, F. (2006). *Educating English language learners: A synthesis of research evidence*. Cambridge, UK: Cambridge University Press.

Gersten, R., & Woodward, J. (1995). A longitudinal study of transitional and immersion bilingual education programs in one district. *The Elementary School Journal*, *95*(3), 223–239.

Goldenberg, C., & Coleman, R. (2010). *Promoting academic achievement among English learners: A guide to the research*. Thousand Oaks, CA: Sage.

Greene, J. P. (1997). A meta-analysis of the Rossell and Baker review of bilingual education research. *Bilingual Research Journal*, *21*, 103–122.

Grissom, J. B. (2004). Reclassification of English learners. *Education Policy Analysis Archives*, *12*(36), 1–36.

Hakuta, K., Butler, Y., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (No. 2000-1). Berkeley, CA: University of California Linguistic Minority Research Institute.

Harklau, L. (1994). "Jumping tracks": How language-minority students negotiate evaluations of ability. *Anthropology & Education Quarterly*, *25*(3), 347–363.

Hill, L., Betts, J., Chavez, B., Zau, A., & Volz Bachofer, K. (2014). *Pathways to fluency: Examining the link between language reclassification policies and student success*. San Francisco, CA: Public Policy Institute of California.

Hopkins, M., Thompson, K., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance a key issue in ESEA reauthorization. *Educational Researcher*, *42*(2), 101–108.

Hugo Lopez, M., Gonzalez-Barrera, A., & Cuddington, D. (2013). *Diverse origins: The nation's 14 largest Hispanic-origin groups*. Washington, DC: Pew Hispanic Center.

Jepsen, C., & De Alth, S. (2005). *English learners in California schools*. San Francisco, CA: Public Policy Institute of California.

Kanno, Y., & Kangas, S. (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*, *XX*, XXX–XXX.

Kao, G., & Thompson, J. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*, 417–443.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560.

Lara, A. Long-term English learners, Pub. L. No. AB-2193 (2011).

Lee, C. (2006). *Denver public schools: Resegregation, Latino style*. Cambridge, MA: Harvard University, Civil Rights Project.

Lee, D., Perreira, K. M., & Harris, K. M. (2006). Making it in America: High school completion by immigrant and native youth. *Demography*, *43*(3), 511–536.

Linquanti, R. (2001). *The redesignation dilemma: Challenges and choices in fostering meaningful accountability for English learners*. Berkeley, CA: The University of California Linguistic Minority Research Institute.

Linquanti, R., & Cook, G. (2013). *Toward a "common definition of English learner": A brief defining policy and technical issues and opportunities for state assessment consortia*. Washington, DC: Council of Chief State School Officers.

López, L. M., & Greenfield, D. B. (2004). The cross-language transfer of phonological skills of Hispanic head start children. *Bilingual Research Journal*, *28*(1), 1–18.

Moschkovich, J. (2012). Mathematics, the Common Core, and language: Recommendations for mathematics instruction for ELs aligned with the Common Core. *Commissioned Papers on Language and Literacy Issues in the Common Core State Standards and Next Generation Science Standards*, *94*, 17.

Olsen, L. (2010). *Reparable harm: Fullfilling the unkept promise of educational opportunity for California's long term English learners*. Long Beach, CA: Californians Together.

Ordóñez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology*, *94*(4), 719–728.

Paez, M., & Rinaldi, C. (2006). Predicting English word reading skills for Spanish-speaking students in first grade. *Topics in Language Disorders*, *26*(4), 338–350.

Parrish, T., Linquanti, R., Merickel, A., Quick, H., Laird, J., & Esra, P. (2002). *Effects of the implementation of Proposition 227 on the education of English learners, K–12, year 2 report*. Oakland, CA: AIR & WestEd.

Parrish, T., Merickel, A., Perez, M., Linquanti, R., Socias, M., Spain, A., & . . . Delancey, D. (2006). *Effects of the implementation of Proposition 227 on the education of English learners, K-12: Findings from a five-year evaluation. Final report for AB 56 and AB 1116*. Washington, DC: American Institutes for Research and WestEd.

Portes, A., & Rumbaut, R. G. (2006). *Immigrant America: A portrait*. Berkeley, CA: University of California Press.

Ragan, A., & Lesaux, N. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives*, *14*, 20.

Ramsey, A., & O'Day, J. (2010). *Title III policy: State of the states*. Washington, DC: U.S. Department of Education.

Reardon, S., Brennan, R., & Buka, S. (2002). Estimating multi-level discrete-time hazard models using cross-sectional data: Neighborhood effects on the onset of adolescent cigarette use. *Multivariate Behavioral Research*, *37*(3), 297–330.

Reardon, S., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, *46*(3), 853–891.

Robinson, J. (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.

Salazar, J. J. (2007). *Master plan evaluation report for English learners—2005/2006*. Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Assessment Branch.

Saldate, M., Mishra, S. P., & Medina, M. (1985). Bilingual instruction and academic achievement: A longitudinal study. *Journal of Instructional Psychology*, *12*, 24–30.

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*, *35*(2), 139–156.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Slavin, R., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Educational Evaluation and Policy Analysis*, *33*(1), 47–58.

Thomas, W., & Collier, V. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz, CA: Center for Research on Education, Diversity and Excellence, University of California-Santa Cruz.

Thompson, K. (2012). *Are we there yet?: Exploring English learners' journey to reclassification and beyond*. Unpublished manuscript, Stanford University.

Umansky, I. (2014). *Peeling back the label: Does English learner classification help or hurt students learning English?* (Stanford University Center for Education Policy Analysis Working Paper). Stanford, CA: Stanford University.

Valdés, G. (1998). The world outside and inside schools: Language and immigrant children. *Educational Researcher*, *27*(6), 4–18.

Valenzuela, A. (1999). *Subtractive schooling: US-Mexican youth and the politics of caring. SUNY Series, The Social Context of Education*. Albany, NY: State University of New York Press.

Warren, P. (2004). *A look at the progress of English learner students*. Sacramento, CA: Legislative Analyst's Office.

Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, *55*(3), 269–317.

Wolf, M. K., Kao, J. C., Griffin, N., Herman, J. L., Bachman, P., Chang, S. M., & Farnsworth, T. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses* (No. 732). Los Angeles, CA: CRESST.

Manuscript received November 30, 2013
Final revision received June 5, 2014
Accepted July 3, 2014