Journal of Educational Psychology
2012, Vol. 104, No. 2, 265–285

0022-0663/11/$12.00 DOI: 10.1037/a0025861

A Longitudinal Analysis of Academic English Proficiency Outcomes for Adolescent English Language Learners in the United States

Rachel B. Slama
Harvard University

The recent availability of nationally mandated academic English proficiency data on all English language learners (ELLs) from landmark United States federal education legislation now makes it possible to track these learners' academic progress longitudinally. Using 5 waves of 9th through 12th grade academic English proficiency data from 2004–2008 for a statewide cohort of 9th grade ELLs (n = 3,702), I employed growth modeling to fit a multilevel model for change in academic English proficiency (Singer & Willett, 2003). I found that the average ELL in my sample started high school performing at an early intermediate level of academic English proficiency and was not projected to reach the score indicating proficiency until the end of 11th grade. Further, U.S.-born ELLs began high school with significantly higher levels of academic English proficiency than their foreign-born ELL peers, but foreign-born ELLs caught up by the end of high school. However, 60% of high school ELLs were born in the U.S.—implying that large numbers of these students had spent 9 or more years in U.S. schools without developing sufficient academic language needed to perform mainstream academic work in English. The findings emphasize the need for academic language interventions for adolescent ELLs. This study has implications for countries struggling to promote the language development and academic achievement of large numbers of language minority learners.

Keywords: academic language development, English language learner, generational status, adolescent, longitudinal study

Large numbers of the United States' growing population of English language learners (ELLs)—students who speak a non-English language at home and who have not acquired sufficient academic English to perform ordinary classroom work in English—are U.S.-born children of immigrant parents. U.S.-born children of immigrants represent 6% of public-school students nationwide and are the most rapidly growing segment of the U.S. school-age population (Capps et al., 2005). Furthermore, the population of U.S.-born children of immigrants is expanding the fastest at the high school level (Ruiz-de-Velasco, Fix, & Chu Clewell, 2000), where ELLs are at the greatest risk for academic failure (Consentino de Cohen, Deterding, & Chu Clewell, 2005; Short & Fitzsimmons, 2007). Large percentages of U.S.-born children of immigrants—an estimated 57% of ELLs nationwide—are still designated ELL, requiring specialized language learning services (Batalova, Fix, & Murray, 2007; Glick & Hohmann-Marriott, 2007; Goldenberg, 2008). Remaining in language learning programs throughout a student's entire school trajectory signals a failure to develop sufficient academic language—and specifically, academic English in the U.S. setting—to be successful in mainstream classrooms. Meeting these goals is a cornerstone of school success that goes far beyond every day conversational language (Scarcella, 2003). To be clear, these trends are most alarming for those ELLs who have spent enough time in U.S. schools to have developed grade-level academic English proficiency but have not met this goal. However, empirical studies and policies linked to ELLs' educational outcomes, broadly speaking, and to literacy and academic language, more specifically, do not account for students' generational status (Goldenberg, Rueda, & August, 2006). Generational status refers to whether the student and their parents were born in the United States or abroad and, specifically, whether these students were U.S.-born to at least one immigrant parent (second generation), U.S.-born to second-generation parents (third generation), or foreign-born (first generation; as commonly defined in the immigration literature; see Callahan, Wilkinson, Muller, & Frisco, 2009, for a review). Generational status is an important source of variation in ELLs' language development because it shapes these students' linguistic and school experiences (Goldenberg, Rueda, & August, 2006), yet seldom has its role in academic language been studied empirically. Thus, the present study examines the role of generational status in ELLs' academic language development.

In the present study, I examine two competing hypotheses concerning the impact of generational status on immigrant stu-

This article was published Online First December 5, 2011.

Support for this analysis was funded by the Harvard Graduate School of Education Dean's Summer Fellowship and the Gordon M. Ambach Fellowship. I am grateful to John Willett and Pia Caronongan for their contributions to the study design and analysis, Catherine Snow and Nonie Lesaux for their expertise on English language learners and language development, and Carrie Conaway for her help with data access and her in-depth knowledge of Massachusetts state policy. I thank students in Catherine Snow and Nonie Lesaux's lab groups for their comments on drafts of this article and Nealia Khan for her assistance in constructing the longitudinal dataset.

Correspondence concerning this article should be addressed to Rachel B. Slama, Graduate School of Education, Harvard University, Larsen Hall 407, 14 Appian Way, Cambridge, MA 02128. E-mail: Rachel_Slama@mail.harvard.edu

dents' academic outcomes. On one hand, and consistent with the immigrant optimism hypothesis (Kao & Tienda, 1995), researchers have argued that immigrant students' outcomes have the potential to improve with each successive generation as students learn English and maintain high aspirations for educational achievement and upward mobility (Portes & Rumbaut, 2001). On the other hand, immigrant student performance may follow a downward trajectory across generations as immigrant students begin to behave more like low-achieving, native-born peers (Suárez-Orozco & Suárez-Orozco, 2001) and start to exhibit a higher prevalence of academic, social, and health-related risk behaviors (National Research Council [NRC], 1998). Furthermore, immigrants whose families have been in the United States for many generations are more likely to have encountered discrimination and thus become disillusioned about the prospect of upward mobility (for reviews, see Goldenberg, Rueda, & August, 2006; Kao & Tienda, 1995; Ogbu, 1987; Suárez-Orozco & Suárez-Orozco, 2001). While these scholars cite broader academic, health, and social outcomes for immigrant youth, very few empirical studies have compared literacy or academic language proficiency outcomes among successive generations of immigrant students (Goldenberg et al., 2006). The one study which did include literacy outcomes (Buriel & Cardoza, 1988) examined cross-sectional data to determine the effect of generational status on vocabulary and reading achievement for a sample of Mexican American students only. No studies to date have examined academic language development within a large, representative sample of second language learners.

Extending this previous work, I use student-level demographic data and English proficiency test score data on both U.S. and foreign-born students from a state in the Northeastern United States for the 2004 cohort of ninth grade ELLs to conduct the first longitudinal analysis of adolescent ELLs' growth in academic English proficiency. With the exception of a few high quality longitudinal studies of second language learners' educational outcomes (for a review, see Kieffer, 2008), the majority of studies of immigrant students' educational outcomes frequently used cross-sectional data (e.g., Buriel & Cardozo, 1988; Collier & Thomas, 1989; Fernández-Kelly & Schauffler, 1994) or analyzed longitudinal data sets but examined achievement outcomes at only one point in time, controlling for baseline test scores (e.g., Glick & White, 2003; Kao & Tienda, 1995). In addition, there is a lack of empirical work focusing on immigrant students' academic language outcomes (Scarcella, 2003). By contrast, I include five waves of data on the same set of individuals over time and model their individual growth trajectories (Singer & Willett, 2003). The present longitudinal analysis is a contribution to the field because there is an overall dearth of longitudinal work on adolescent ELLs' academic language development (Short & Fitzsimmons, 2007)—a subgroup of ELLs who are at increased risk for educational failure. Investigating growth in academic language proficiency over time is especially important for this population of learners, given that language development is, by nature, dynamic.

Academic English is commonly defined as the English typically used in academic settings such as in school classrooms in order to acquire knowledge (e.g., Snow & Uccelli, 2009). It implies use of decontextualized discourse (Wong-Fillmore & Snow, 2000), as opposed to more automatically acquired, context-embedded conversational language (Scarcella, 2003; Snow & Uccelli, 2009; see Cummins, 1979, 1981a, for early theories on academic English acquisition). Developing academic language required for school success is particularly challenging for ELLs; it is estimated to take 3–5 years for second language learners to become proficient in conversational English but at least 4–7 years for students to develop academic proficiency (e.g., Hakuta, Goto Butler, & Witt, 2000) and 6–8 years for ELLs who immigrated between ages 12–15 years (Collier, 1987). Thus, even after many years in the United States, both U.S.-born and foreign-born high school ELLs may still struggle to develop academic English proficiency, and many remain designated as ELL throughout middle and high school.

Adolescent ELLs who have not developed adequate academic language skills to be successful in school are at elevated risk of dropping out of school before graduation, compared with their nonimmigrant peers (Suárez-Orozco & Suárez-Orozco, 2001; White & Kaufman, 1997). These students leave school ill-prepared to compete in the new global economy (Suárez-Orozco & Suárez-Orozco, 2001). ELLs who have failed to acquire sufficient academic English may be unable to pass high school exit exams—increasingly required to receive a secondary school diploma in the United States (Center on Education Policy, 2008)—and may be further encouraged to drop out of school along with other high-risk peers (Jones, Jones, & Hargrove, 2003; Thomas, 2005). Furthermore, adolescent ELLs' increased risk for drop out has large implications for these students' life trajectories because failure to obtain a high school diploma leads to a host of negative life outcomes, such as lower paying jobs and a decreased probability of attending college (Belfield & Levin, 2007; Rumberger & Lamb, 2003), which may limit these learners to a life of poverty (National Center for Education Statistics [NCES], 2004).

Theoretical Framework: Generational Status as an Importance Source of Variation in Second Language Learners' Development of Academic Language

The theoretical framework of the present study draws on research on language development and academic achievement outcomes for monolingual and ELL students. Understanding the developmental trajectory of academic language in the population of second language learners—with special attention to the existing heterogeneity within this group—is particularly salient to these learners' academic success and subsequent life outcomes.

Academic Language and Students' Life Trajectories

Academic English—and its entailed mastery of reading, writing, listening, and speaking skills (Scarcella, 2003)—is required for school success, yet this presents a challenge to native speakers of English and to second language learners (Snow & Uccelli, 2009; Wong-Fillmore & Snow, 2000). Well-developed proficiency in these four areas—oral and listening communication skills and reading and writing skills—are required for students to communicate effectively in everyday and academic situations. Therefore, in the present study, I examine proficiency in academic English, treating this construct as a composite measure of reading, writing, speaking, and listening skills.

In the big picture, promoting the development of academic English for all learners and for ELLs in particular is a cause for concern because mastering academic English underpins students'

economic and social opportunities (Scarcella, 2003). Monitoring the English development of ELLs is especially critical in the current global economy which demands students' mastery of complex reading and vocabulary skills (Snow & Kim, 2007)—skills that have traditionally proved challenging for ELLs. Indeed, well-developed academic language skills are paramount to ELL and monolingual students' success because they are required in order to pass high stakes standardized mathematics, English, and science assessments, receive at least a high school diploma, and have access to a decent quality of life as productive members of society (Oreopoulos, 2007; Scarcella, 2003). Unfortunately, status as second language learners puts ELLs at increased risk of lagging behind non-ELL peers academically (Kieffer, Lesaux, & Snow, 2008), dropping out of high school, failing to obtain postsecondary degrees, and subsequently limiting them to low-wage jobs and a life of poverty (NCES, 2004). Despite the challenge that ELLs confront in developing academic English, very little is known about its normative development over time among second language learners (Snow & Uccelli, 2009).

Further, student's academic performance—which is closely linked to ELLs' academic English language development—is an important predictor of high school completion for all students. Poor academic performance and grade retention are associated with high school drop out among immigrants of diverse ethnic backgrounds (Fernandez, Paulsen, & Hirano-Nakanishi, 1989; Velez, 1989). Thus, the inability to pass academic language-dependent tests (see Wong-Fillmore & Snow, 2000, for a review) puts these students at risk of failing to obtain a high school diploma and excludes them from the economic and social benefits that additional educational attainment confers (e.g., NCES, 2004). Further, among ELLs who do obtain a high school diploma and enter postsecondary institutions, many have inadequate academic language skills to be successful in college courses, despite having completed their entire elementary and secondary education in U.S. schools (Scarcella, 2003).

Generational Status Shapes Students' L1 and L2 Linguistic and School Experiences

Generational status is one of many group-level sociocultural influences that are an important source of variation in students' first (L1) and second (L2) language experiences and, in turn, their school experiences and academic achievement. Empirical research on the influence of generational status on students' academic language development and achievement is limited (Goldenberg, Rueda, & August, 2006), yet there are several hypothesized mechanisms that may explain why generational status is an important predictor of L2 academic language development.

First, ELLs' generational status may influence the pace at which these students acquire academic English in that it is largely related to students' ability to speak the home language and their oral language proficiently in English upon school entry (Driscoll, 1999; NRC, 1998; Portes & Rumbaut, 1990; Portes & Rumbaut, 2001). With respect to home language (L1) proficiency, it is well documented that home language literacy development supports second language literacy (Cahill, 1987; Dolson, 1985; Hancock, 2002; for relevant theory, see Cummins, 1986; Cummins, 2001). Further, there is evidence of language shifts toward English with each successive generation, until English "effectively became the mother tongue for subsequent generations" (Portes & Schauffler, 1996, p. 11). In other words, at school entry, we would expect greater initial proficiency in English with each subsequent generation of immigrant students.

Second, generational status may play an important role in academic achievement through its influence on motivation and educational expectations. There is evidence that more recent immigrants arrive in the United States with high aspirations for academic achievement (e.g., Driscoll, 1999) that often fade with time in the United States as they become disillusioned by limited opportunities and societal discrimination (e.g., Ogbu & Matute-Bianchi, 1986; Rumbaut, 1995; Suárez-Orozco & Suárez-Orozco, 1996). Subsequently, some scholars hypothesize that time spent in the United States—which, by definition, is linked to generational status—is an important predictor of reading scores (Ima & Rumbaut, 1989).

However, there are many complex factors, such as age at arrival in the host country, which moderate the impact of generational status on students' academic language and school outcomes. White and Kaufman (1997) concluded that more recently arrived adolescent immigrant students were at greater risk of dropping out of high school than their immigrant peers who had spent more time in the United States. Consistent with this finding, two additional studies (Glick & White, 2003; Short & Fitzsimmons, 2007) found that among foreign-born students, those who entered U.S. schools as young children had more promising academic outcomes than their immigrant peers who arrived in the United States as adolescents. Thus, in addition to generational status, age at arrival in the host country plays a role in predicting future academic success. These studies illustrate the complex nature of the relationship between generational status and academic outcomes for high school immigrant students and the importance of hypothesis-generating studies with large, representative samples.

The Present Study

In the present study, I link student-level records of performance on a standardized, unidimensional measure of academic language—representing a composite of students' reading, writing, speaking, and listening proficiencies—over time for a cohort of high school ELLs. This measure is well-suited for examining a precise measure of change in academic language development over time because it is vertically equatable—a psychometric property required for longitudinal analysis (Singer & Willett, 2003). Specifically, I examine heterogeneity in academic language development over students' high school trajectories on the basis of ELLs' generational status. I operationalize generational status by whether a student was U.S.-born or foreign-born. Examining variation in academic language development on the basis of generational status will generate new hypotheses for the field because of the variation implied in time spent in U.S. schools, exposure to academic English, and overall time spent in an English-speaking country. Based on available data on parent country of birth, disaggregation of findings was limited to two sociodemographic groups: first generation students and second and higher generation. I highlight the importance of collecting more fine-grained data on generational status, students' access to formal schooling in the United States and abroad, degree of domestic and international

mobility, age at arrival, home language literacy experiences and parent educational attainment for future analyses.

Research Questions

In the present study, I examined longitudinal trajectories of academic English proficiency for the 2004 ninth grade ELL cohort (n = 3,702) from 2004–2008 in one northeast U.S. state and investigated the impact of these students' generational status on these trajectories. My two specific research questions are as follows:

> *Research Question 1:* What is the shape of the growth trajectories in academic English proficiency for English language learners (ELLs) between ninth and 12th grade?
>
> *Research Question 2:* Are the trajectories in academic English proficiency steeper for U.S.-born children of immigrants—those who have plausibly spent more time in U.S. classrooms—or for first-generation immigrants who have had more time to develop literacy in their first language?

Method

Dataset

To address these questions, I constructed a 4-year longitudinal dataset comprising restricted, student-level academic English proficiency test-score data and demographic information on the 2004 ninth grade Massachusetts ELL cohort. Massachusetts, a state located in the northeastern United States, can serve as a case example from which to examine within-group diversity in ELLs' academic English proficiency outcomes because it is a state with significant linguistic diversity, a relatively homogeneous, English-only instructional environment (as opposed to use of students' native language as a medium of instruction), and a state system that maintains student-level performance and demographic data.

To create this dataset, I combined two unique sources of data. First, I obtained 2004–2008 ninth through 12th grade academic English proficiency test-score data on the *Massachusetts English Proficiency Assessment* (MEPA)—the annual, federally mandated proficiency test designed to measure ELLs' progress in reading, writing, speaking, and listening skills (García, McKoon, & August, 2006). These proficiency tests were mandated with the goal of promoting ELLs' English language acquisition and academic achievement (U.S. Department of Education, 2009) and, thus, are well-suited as a measure of ELLs' academic English proficiency development. In Table A1, in Appendix A, I present the Massachusetts Department of Elementary and Secondary Education-determined scaled scores as they correspond to four academic English performance levels. I report select characteristics in the four domains of academic English proficiency that a given ELL scoring at that level should demonstrate. The performance levels—beginning, early intermediate, intermediate, and transitioning—describe qualitatively how a given scaled score maps onto individual ELLs' academic language proficiency in the classroom on reading, writing, speaking, and listening tasks.

Second, I included corresponding student demographic data from the Massachusetts' *Student Information Management System* (SIMS) for students who took the MEPA in the fall of 2004. Fall 2004 marked the first administration of the MEPA. Thus, 2004 was the first year when it was possible to analyze individual ELLs' English proficiency outcomes disaggregated by student demographic characteristics.

Analytic Sample

I included all 3,702 ELLs who took the MEPA upon entry into ninth grade in Massachusetts in the fall of 2004 (n observations = 18,510; n schools = 193). I incorporated data on a student's generational status as reported at baseline and students' attendance in a language learning program as reported at each wave. In Table B1, in Appendix B, I present descriptive statistics on the demographic characteristics, program status, and language background of my sample. In Table B1, I show that my sample is predominantly low income (78%; n = 2,873), Spanish speaking (58%; n = 2,144), receiving language learning services in English (76%; n = 2,821), and U.S.-born (59%; n = 2,186). The aggregate features of my sample are consistent with the demographic profile of ELLs nationwide—immigrant students are three times more likely to come from low socioeconomic backgrounds as their peers from nonimmigrant backgrounds and more than 75% come from Latin America or Asia (Ruiz-de-Velasco, Fix, & Chu Clewell, 2000). In Table C1, in Appendix C, I show that the majority of ELLs received language learning services in English, consistent with Massachusetts' English-only instructional policy. Further, I show that large numbers of ELLs remained in an all-English language learning program, sheltered English-immersion, for their entire high school careers (n = 602).

Measures

I organized my longitudinal data in a person–period format (Singer & Willett, 2003). In the dataset, therefore, each student contributes one row for every occasion of measurement. The person–period format permits values to be recorded for both time-varying (e.g., ELL status) and time-invariant variables (e.g., generational status).

Academic English language proficiency. My time-varying outcome of interest, MEPA, is a continuous measure of academic English proficiency. I used MEPA scaled scores that represent a composite reading, writing, speaking, and listening score, ranging from 300–400. MEPA scaled-scores met the criteria for vertical equatability required for multilevel modeling because scores were calibrated to a common year using item response theory (IRT), and all students in Grades 9 through 12 took a common English proficiency assessment (Massachusetts Department of Elementary and Secondary Education [MDESE], 2005). Thus, MEPA scores within the ninth through 12th grade reading span can be used to model individual growth in academic English proficiency over time (Singer & Willett, 2003).

All ELLs took the MEPA in both the fall and the spring of the first year of administration (2004) and every spring thereafter from 2004–2008, such that each ELL in my sample could contribute a maximum of five unequally spaced waves of assessment data to the dataset. Academic English proficiency is linked strongly to academic performance for immigrant students (Glick & Hohmann-

Marriott, 2007). Thus growth in academic English proficiency is a useful proxy for gauging immigrant students' academic success.

Time. I recorded the continuous time-varying variable Time as the number of years that had passed since entry into ninth grade. Year has values of 0.5, 1, 2, 3, and 4, corresponding to Fall 2004, Spring 2005, Spring 2006, Spring 2007, and Spring 2008 MEPA assessments, respectively. Using Year as my principal measure of time was more appropriate than using a student's grade-level because Year captures students' total number of years spent in U.S. high schools, regardless of grade repetition.

Generational status. My main question predictor for Research Question 2, U.S.-born was a dichotomous variable that indicated whether the student was U.S.-born or foreign-born, based on the country of origin variable in the SIMS data set (1 = born in the United States; 0 = otherwise). Nearly 60% of ninth grade ELL students in my sample were U.S.-born (n = 2,186; see Table B1 in the Appendix). U.S.-born is the best available proxy for generational status in the present dataset; however, a noteworthy limitation to the dataset is that it does not include accurate information on the length of time that the student has spent in U.S. schools. In other words, I cannot distinguish between a foreign-born ELL who entered U.S. schools in kindergarten and a U.S.-born ELL who may have left the country during their early elementary grades and later returned as an adolescent.

In Table 1, I display the sample mean MEPA scores for the cohort of ninth grade ELLs in each wave by program, income, and generational status. In the last two rows of the table, I display mean academic English proficiency scores, as measured by the MEPA for U.S.-born ELLs and their foreign-born ELL peers. Comparing the mean scores for these two groups of students in the fall of ninth grade (Fall 2004), U.S.-born students, on average, scored 349.41 on the MEPA, compared with their foreign-born peers who scored 332.81—a difference of 16.6 points or, approximately, half of a standard deviation. However, by the end of 12th grade (Spring 2008), among those students who remained designated as ELL, U.S.- and foreign-born ELLs performed at indistinguishable levels of academic proficiency. According to Table A1 in the Appendix, U.S.-born ELLs remained at an intermediate level of English proficiency throughout their high school careers, while foreign-born ELLs progressed from an early intermediate level to an intermediate level. This suggests that among students in the study sample that were designated ELL at the end of high school, neither foreign-born nor U.S.-born ELLs had developed the academic language skills required to move to the transitioning level and to ultimately be moved out of specialized language learning programs.

In Table 2, I present summary statistics on the 2004 ninth grade ELL cohort's participation in one of the four language learning

Table 1
Sample Mean Composite Scores, Standard Deviations, and Number of Test Takers on the Massachusetts English Proficiency Assessment (MEPA) for the Cohort of Ninth Grade ELL Students, by Program and Generational Status

Variable	Fall 2004	Spring 2005	Spring 2006	Spring 2007	Spring 2008
Mean MEPA score (all ELLs)	342.08	349.08	358.97	363.50	366.77
SD	32.02	30.49	29.04	26.37	24.92
N	2,848	2,808	1,842	1,109	672
Program status					
No program	354.05	358.19	378.02	378.73	377.47
SD	28.29	27.30	22.01	21.0	20.63
N	266	338	95	89	93
Sheltered English immersion	341.82	349.04	358.43	362.12	364.88
SD	32.19	30.59	28.79	26.59	25.27
N	2,218	2,212	1,581	830	500
Bilingual	323.09	332.37	347.09	358.42	357.06
SD	25.45	27.59	23.09	25.49	31.21
N	230	214	123	72	18
Opt out	355.17	361.79	370.93	364.88	368.79
SD	29.62	27.77	27.62	25.39	21.60
N	134	43	43	117	61
Income status					
Low income	341.25	347.25	357.83	362.75	366.23
SD	31.36	30.03	28.97	26.37	25.21
N	2,240	2,232	1,480	896	559
Not low-income	345.13	353.53	363.65	366.66	369.44
SD	34.17	31.84	28.89	26.18	23.39
N	608	576	362	213	113
Generational status					
U.S.-born	349.41	353.40	361.18	365.07	367.31
SD	30.56	29.19	28.38	26.58	24.18
N	1,590	1,549	949	528	313
Foreign-born	332.81	343.76	356.63	362.98	366.30
SD	31.41	31.22	23.56	26.19	25.58
N	1,258	1,259	893	581	359

Note. N students = 3,702; N schools = 193.I treat program status as time varying, while generational and income status (U.S.-born) are time-invariant. ELL = English language learner.

Table 2
Participation in Language Learning Programs, by Year and Generational Status, for the 2004 Ninth Grade Massachusetts English Language Learner Cohort

	U.S.-born ELLs				Foreign-born ELLs			
Time period	No program	SEI	Bilingual	Opt out	No program	SEI	Bilingual	Opt out
Fall 2004	16.70% ($N = 365$)	71.87% ($N = 1,571$)	4.57% ($N = 100$)	6.86% ($N = 150$)	3.96% ($N = 60$)	82.45% ($N = 1,250$)	10.42% ($N = 158$)	3.17% ($N = 48$)
Spring 2005	22.50% ($N = 491$)	71.13% ($N = 1,552$)	3.99% ($N = 87$)	2.38% ($N = 52$)	8.97% ($N = 136$)	79.55% ($N = 1,206$)	9.96% ($N = 151$)	1.52% ($N = 23$)
Spring 2006	28.86% ($N = 538$)	66.04% ($N = 1,231$)	2.36% ($N = 44$)	2.74% ($N = 51$)	18.93% ($N = 250$)	72.29% ($N = 955$)	7.65% ($N = 101$)	1.14% ($N = 15$)
Spring 2007	44.87% ($N = 691$)	44.16% ($N = 680$)	1.75% ($N = 27$)	9.22% ($N = 142$)	36.18% ($N = 407$)	52.98% ($N = 596$)	6.13% ($N = 69$)	4.71% ($N = 53$)
Spring 2008	60.05% ($N = 708$)	31.89% ($N = 376$)	—	7.29% ($N = 86$)	56.15% ($N = 507$)	38.87% ($N = 351$)	1.55% ($N = 14$)	3.43% ($N = 31$)

Note. N students = 3,702; N schools = 193. ELL = English language learner; SEI = sheltered English immersion. Dash indicates groups that are too small to report. Consistent with Massachusetts Department of Elementary and Secondary Education policy, I do not report data for groups smaller than 10 students.

program options to emphasize (a) that the ELL population's participation in language learning programs changes over time as the students develop academic English skills and move into mainstream classrooms and (b) that consistent with the data presented in Tables A1 and B1, by 2004, Massachusetts was a largely English-only instructional context for the majority of ELLs. In the second row of Table 2, I display the four program options: sheltered English immersion (SEI), transitional bilingual program, and no language learning services for families who either opted for their child not to participate in a program (Opt out), or for students who were placed in mainstream classes (No program). I organized this table by MEPA assessment wave for the entire ELL sample (columns 2–5), U.S.-born ELLs (columns 6–9), and foreign-born ELLs (columns 10–13). In the respective No program columns (2, 6, 10), as we would expect, over time, increasing numbers of ELLs—both U.S.- and foreign-born—are exiting language learning programs into mainstream classes. This pattern underscores the fact that the ELL designation, unlike any other federal accountability category (e.g., race/ethnicity, special education, low income), is temporary by design (Kieffer, Lesaux, & Snow, 2008). Second, the vast majority of ELLs—again including U.S.- and foreign-born students—were enrolled in sheltered English immersion programs in the fall of 2004. Consistent with the English-only instructional context in Massachusetts, 72% of U.S.-born ELLs ($n = 1,571$) and 82% of foreign-born ELLs ($n = 1,250$) statewide were enrolled in a sheltered English immersion program at the start of ninth grade.

Data Analysis

To address my research questions, investigating how ELLs' academic language proficiency changed over time between ninth and 12th grade, my first task was to employ exploratory data analysis to justify the model specification. Accordingly, in Figure 1, I present empirical growth plots—a summary of how a subsample of students' individual MEPA scores change over time, by generational status—whose visual inspection suggests that all ELLs started off in high school with relatively low MEPA scores and that their individual change in academic language proficiency generally increased over the duration of high school and then began to slow down typically around their 3rd year of high school or 11th grade. Further, I superimposed curvilinear trajectories on the empirical growth plots that highlight a curvilinear, not linear, pattern of change in academic English. Presenting the sets of plots separately by generational status highlights that the foreign-born students generally demonstrated lower levels of academic proficiency at the start of high school, compared with U.S.-born peers, but they appeared to develop academic English skills at a faster rate. Examination of a collection of observed individual academic English language trajectories and observed means, by generational status—which I present in Figure 2—further supports the choice of the quadratic specification. In addition, comparing the initial status and steepness of the growth trajectories in the two panels also suggests that foreign-born ELLs began ninth grade with lower academic English proficiency, compared with their U.S.-born peers, but grew at a faster pace.

In Table 3, I present the results of a taxonomy of fitted multilevel models for the linear, quadratic, and cubic unconditional growth Models 1, 2, and 3, respectively. General linear hypothesis (GLH) tests confirmed that the inclusion of the quadratic specification of time in Model 2 is an improvement over the linear specification of time presented in Model 1, GLH test: $\chi^2(1) = 44.4$; $p < .05$, and the cubic polynomial presented in Model 3, GLH test: $\chi^2(1) = -1.8$; $p > .05$. In summary, the statistical results of the models presented in Table 3 confirmed that a quadratic specification of time best summarized ELLs' academic English trajectories.

I fit a taxonomy of multilevel models for change in my person–period data set using the composite random intercepts and slopes specification of the multilevel model, with the quadratic specification of time. The complete Level-1, Level-2, and Level-3 specifications of the final model that best answers Research Question 1 are summarized in Appendix D.

It is important to account for the fact that the 3,702 students in the sample were nested within 193 schools and, thus, the

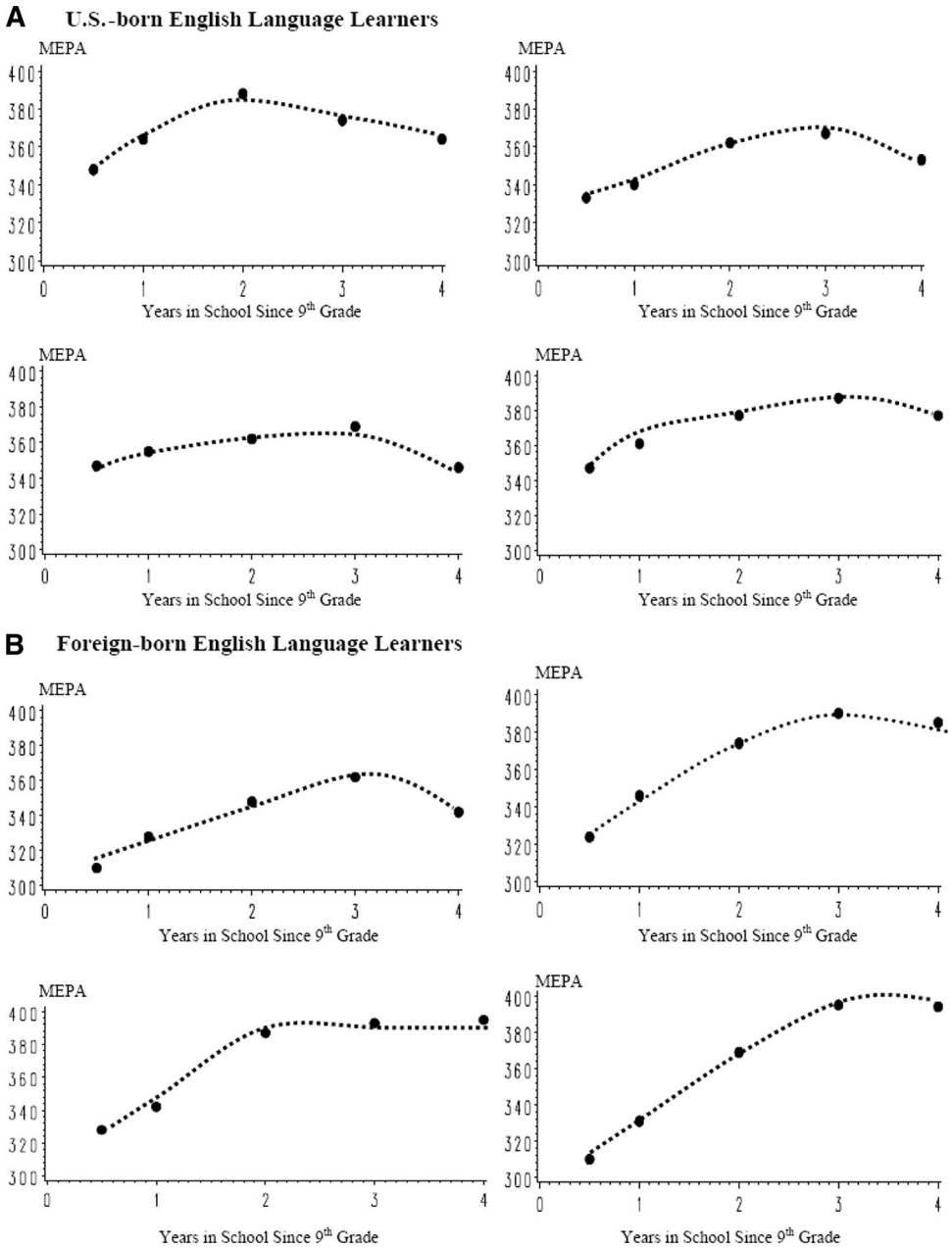

Figure 1. Empirical growth plots summarizing how English language learners' (ELLs) academic English proficiency performance changes over time for a random sample of U.S.-born students (A; $n = 4$) and foreign-born students (B; $n = 4$), with curvilinear trajectories superimposed. MEPA = *Massachusetts English Proficiency Assessment.*

random effects of school is included in the intercept in all models presented in the analysis. While the sole purpose of including the random effect of school was to account for the clustering of students within schools, I note that its inclusion in only the intercept is a limitation of the data. Clustering of ELLs across a smaller number of schools in Massachusetts is consistent with nationwide trends indicating that ELLs tend to be highly clustered in urban schools (Kasinitz, Mollenkopf, & Holdaway, 2008; Orfield, 2001; Ruiz-de-Velasco, Fix, & Chu Clewell, 2000; Suárez-Orozco, Suárez-Orozco, & Todorova, 2008).

Second, to examine differences in initial rate of change and initial academic English proficiency status by students' generational status (Research Question 2), I fitted the Level-1, Level-2, and Level-3 specifications of the final model into which I introduced my time-invariant question predictor U.S.-born at Level-2 (see Appendix D for complete Level-1, Level-2, and Level-3 specifications of the final model that best answers Research Question 2).

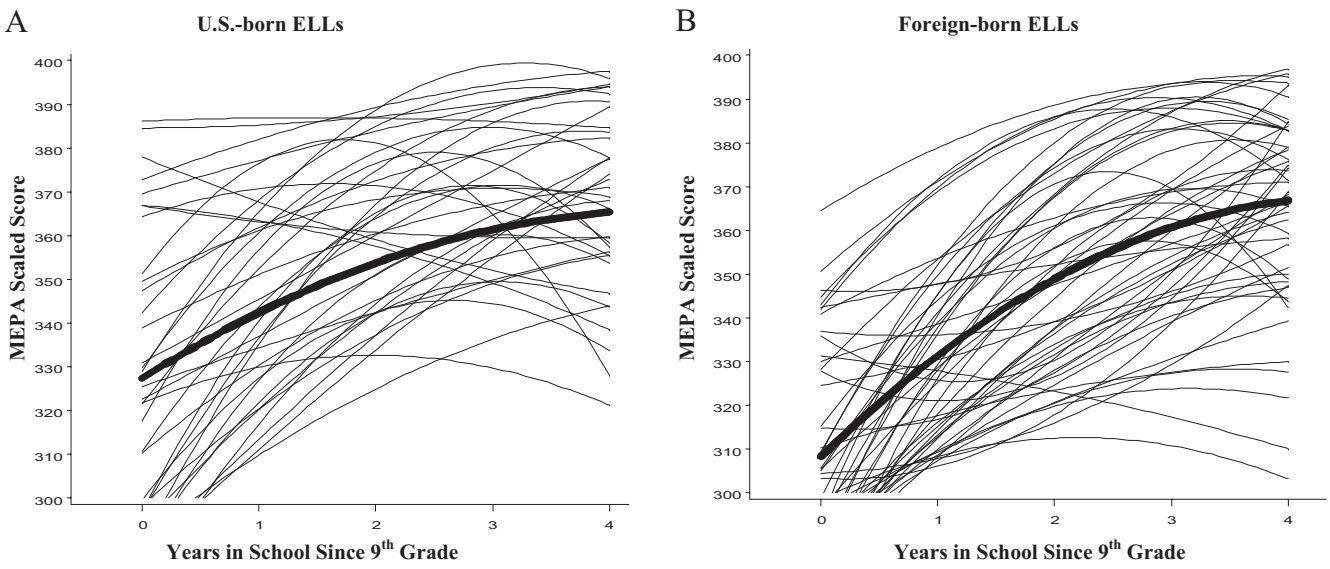

Figure 2. Fitted academic English language trajectories based on the *Massachusetts English Proficiency Assessment* (MEPA) scaled scores presented separately for a sample of 100 English language learners (ELLs) by generational status. A: Presents the collection of quadratic trajectories for foreign-born ELLs ($n = 58$). B: Presents the collection of fitted quadratic trajectories for U.S.-born ELLs ($n = 42$). Both panels also present an average change trajectory for the group.

Missing Data

Given that ELLs are at increased risk for dropping out of high school (Ruiz-de-Velasco, Fix, & Chu Clewell, 2000), it is not surprising that substantial numbers of students in the sample were missing values on the outcome on one or more occasions; of the original 3,702 students present in Fall 2004, nearly 28% ($n =$ 1,036) were missing outcome values by their junior year of high school. However, one of the strengths of the multilevel model for change is that it can accommodate missing outcome values by letting each student in the data set contribute whatever their empirical growth records permit to the estimation of their academic English proficiency growth trajectories. In short, sample attrition in and of itself does not threaten my ability to fit the hypothesized multilevel model for change. However, missingness offers a potentially more serious threat to the validity of the findings, when data are not missing at random (MAR; Singer & Willett, 2003).

What complicates the missing data problem in the present analysis is that the students in my sample could have been missing values on the outcome for two reasons, which could bias the results in opposite directions. First, because my sample is comprised entirely of ELLs, if a student was reclassified or redesignated subsequently as fluent English proficient (R-FEP), also commonly referred to in schools and districts as former limited English proficient (FLEP), they would no longer be required to take the annual MEPA assessment and would therefore have missing values on the outcome by design. In my view, students who had missing outcome data on the basis of reclassification ($n = 2{,}264$) do not pose a serious threat to the ecological validity of my findings, given that the objective of my study is to model academic English proficiency trajectories for ELLs in the current policy environment, and thus, it seems appropriate to exclude students when they are no longer classified as ELLs.

The second reason that students may be missing outcome data is that the student drops out of high school or disappears from the data set. This is of greater concern because if there is a link between the probability of disappearance from the data set and academic English proficiency, then my estimates of students' academic language trajectories may also be biased, although the direction of the bias is less clear. In order to examine the nature of the sample attrition and the direction of the bias, I conducted several sensitivity tests in which I compared the baseline characteristics of students whose values of the outcome were missing at Spring 2007 and Spring 2008 to those of students with no missing outcome values during the respective periods. I conducted these analyses for the entire sample and then again after excluding R-FEP students. I comment on these results in the Threats to Validity section.

Results

Research Question 1: ELLs' Academic English Proficiency Improved Over High School, but Improvement Slowed During Senior Year

To answer my first research question, I estimated the academic English proficiency trajectories from ninth through 12th grade for the 2004 cohort of ninth grade ELLs. Because my analyses indicated that the trajectories were curvilinear and had a quadratic functional form over time, my findings suggest that the average ELL in this cohort began high school while just beginning to develop proficiency in academic English, initially acquiring academic language at .6 of a standard deviation (*SD* at Fall 2004

Table 3
Taxonomy of Fitted Multilevel Models for Change Representing Massachusetts' Ninth–12th Grade Students' Performance on the Massachusetts' English Proficiency Assessment (MEPA) as a Function of Time (Years in School Since Ninth Grade)

Variable	Parameters	Model 1: Unconditional growth model, linear specification	Model 2: Unconditional growth model, quadratic specification	Model 3: Unconditional growth model, cubic specification
Fixed effects				
Initial status, π_{0i}				
Intercept	β_{000}	343.71***	335.46***	335.50***
Rate of change, π_{1i}				
Intercept	β_{100}	10.603***	19.590***	19.512***
Acceleration, π_{2i}, π_{3i}				
Intercept	β_{200}		−2.032***	−1.988
Intercept	β_{300}			−0.007
Variance components				
Level 1				
Within-person	σ^2_ε	211.69***	211.76***	211.79***
Level 2				
In initial status	σ^2_0	435.05***	336.94***	336.87***
In rate of change	σ^2_1	183.99***	114.42***	114.40***
In acceleration	σ^2_2	6.65***	3.19***	3.19***
Level 3				
Between-schools variance	σ^2_3	561.74***	561.49***	561.48***
Goodness of fit				
Deviance		83,237.1	83,192.7	83,194.5
AIC		83,253.1	83,208.7	83,210.5
BIC		83,278.8	83,234.4	83,236.2
GLH tests			Compare M1 H_0: $\beta_{200} = 0$; $\chi^2(1) = 44.4$, $p < .05$; Reject H_0	Compare M2 H_0: $\beta_{300} = 0$; $\chi^2(1) = -1.8$, $p > .05$; Fail to reject H_0

Note. N students = 3,702; N schools = 193. AIC = Akaike information criterion; BIC = Bayesian information criterion; GLH = general linear hypothesis; M = model; H_0 = the null hypothesis.
*** $p < .001$.

baseline = 32.02) per year, until their growth began to decelerate during the last year of high school. Furthermore, based on the initial estimated rate of growth in academic proficiency, an average student who persisted through high school in this cohort of ELLs reached the minimum score recommended for transition into mainstream classes (375 scaled score points) around the end of 11th grade (2.85 years after the start of ninth grade).[1] To clarify the complex nature of this change over time, I first present plots of fitted growth trajectories for prototypical cases and then map these findings onto their corresponding parameters in a taxonomy of fitted multilevel models for change. In Figure 3, I illustrate that the average ELL in the 2004 cohort began ninth grade with an overall low level of academic English proficiency and those who persisted attained sufficient academic English proficiency for placement in mainstream instructional classrooms at the end of 11th grade. The solid black line in the figure represents the average growth trajectory for all ELLs. First, notice the low starting point of the solid line, especially compared with the dotted line with which I indicate the minimum MEPA score recommended for reclassification into mainstream classrooms (MDESE, 2005; Table A1 in the Appendix). In Table 4, I present a taxonomy of fitted multilevel models for change representing Massachusetts' ninth–12th grade students' performance on the MEPA over time. Note that the fitted unconditional growth model, in the first column, Model 1, shows that at the start of high school the estimated true initial score on the MEPA assessment across all ELLs in the 2004 cohort ($\hat{\beta}_{000}$) was about 335 points ($p < .0001$). Based on the Massachusetts Department of Elementary and Secondary Education–determined performance levels, this score corresponds to an early intermediate level of academic English proficiency.

From this low start, the average ELLs' academic proficiency then grew over time, but not fast enough for most students to reach the minimum criterion to be considered for mainstream classes until near the end of their high school trajectories. Returning to Figure 3, based on this fitted trajectory, a student would reach academic English proficiency near the end of 11th grade (Year = 2.85). Notice that the dotted horizontal line—corresponding to the minimum MEPA scaled score recommended for reclassification—intersects the solid black line at roughly the end of 11th grade. However, it is worth noting that ELLs' academic English proficiency grew until declining at the beginning of 12th grade. Based on the parameter estimates in Table 4, Model 1, I conclude that ELLs' initial rate of change in academic English proficiency performance increased during high school ($\hat{\beta}_{100} = 19.590$; $p < .001$; Fall 2004 baseline $SD = 32.02$), but the estimated rate of

[1] For the average ELL who remained in the sample by 11th grade, I solved the quadratic equation representing the fitted MEPA trajectory for all students in order to project that these students scored in the transitioning category at the end of their 3rd year since the start of high school.

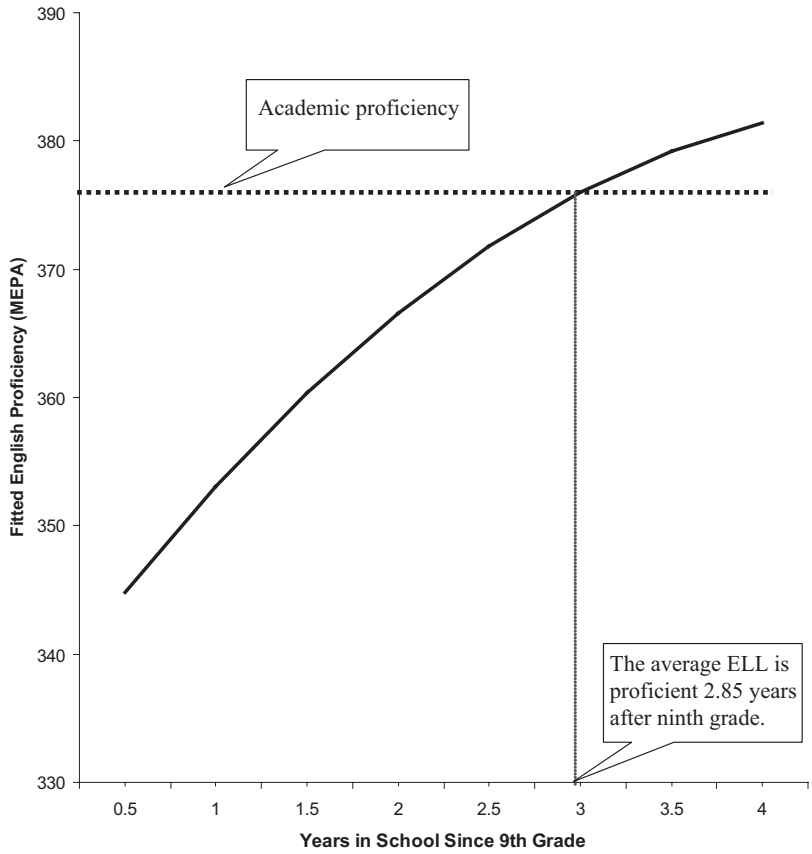

Figure 3. Fitted growth trajectories in performance on the *Massachusetts English Proficiency Assessment* (MEPA; from Model 1 [all English-language learners]). *n* students = 3,702; *n* schools = 193.

growth in academic English proficiency decelerated over time ($\hat{\beta}_{200} = -2.033$, $p < .0001$).

Research Question 2: Foreign-Born ELLs Caught up to Their U.S.-Born ELL Classmates by the End of High School, but Overall Proficiency Remained Low

In addressing Research Question 2, I asked whether the trajectories in academic English language proficiency were steeper for U.S.-born ELLs or for their foreign-born ELL classmates. I found that U.S.-born ELLs exhibited higher levels of academic English proficiency at the beginning of ninth grade than their foreign-born ELL peers. However, foreign-born ELLs had faster rates of growth than did U.S.-born ELLs and, by the end of high school, had caught up to their native-born classmates on measures of academic English proficiency.

Figure 4 shows that the fitted academic English proficiency for a prototypical U.S.-born ELL is higher than the corresponding trajectory for a prototypical foreign-born ELL at the beginning of ninth grade. I conducted a GLH test to confirm that U.S.-born ELLs scored higher than their foreign-born ELL classmates on measures of academic English proficiency in the fall of ninth grade ($\chi^2 = 75.13$; $p < .0001$). In Model 3, in Table 4, I include terms that describe the effect of a student's generational status, U.S.-born, on interindividual differences in intercept, initial rate of change, and curvature and show that the estimated true initial status for a foreign-born ELL (when U.S.-born = 0) was about 328 points ($p < .0001$). U.S.-born ELLs' estimated true initial academic English proficiency status was about .4 of a standard deviation or about 14 scaled points higher than that of their foreign-born ELL classmates ($\hat{\beta}_{010} = 13.568$; $p < .0001$; SD = 32.02). Further, Table 5 summarizes predicted differences in initial status between U.S. and foreign-born ELLs expressed as effect sizes. Table 5 indicates that U.S.-born ELLs began high school scoring one quarter of a standard deviation higher, on average, compared with their foreign-born ELL classmates.

On average, foreign-born ELLs caught up rapidly with their native-born ELL classmates (see Figure 4). Likewise, in Model 3, in Table 4, I show that U.S.-born ELLs had slower initial rates of change in academic English proficiency than did their foreign-born ELL peers ($\hat{\beta}_{100} = -7.131$; $p < .0001$) and an acceleration that was slightly higher ($\hat{\beta}_{210} = 0.893$; $p < .0001$). U.S.-born ELLs' pace of initial growth in academic proficiency was about .2 of a standard deviation lower than that of their foreign-born ELL classmates. However, I conducted a GLH test to confirm that at Year 4, there was no statistically significant difference in academic English proficiency performance between foreign and U.S.-born ELLs ($\chi^2 = 0.16$; $p = .6847$). In Figure 4, I illustrate that U.S. and

Table 4
Taxonomy of Fitted Multilevel Models for Change Representing Massachusetts' Ninth–12th Grade Students' Performance on the Massachusetts' English Proficiency Assessment (MEPA) as a Function of Time (Years in School Since 9th Grade) and Generational Status (U.S.-Born Versus Foreign-Born)

Variable	Parameters	Model 1: Unconditional growth model	Model 2: Including U.S.-born	Model 3: Including U.S.-born and time
Fixed effects				
Initial status, π_{0i}				
Intercept	β_{000}	335.46***	331.65***	327.54***
U.S.-born	β_{010}		6.299***	13.568***
Rate of change, π_{1i}				
Intercept	β_{100}	19.590***	19.801***	23.765***
U.S.-born	β_{110}			−7.131***
Acceleration, π_{2i}				
Intercept	β_{200}	−2.033***	−2.077***	−2.574***
U.S.-born	β_{210}			0.893**
Variance components				
Level 1				
Within-person	σ^2_ε	211.76***	211.79***	209.15***
Level 2				
In initial status	σ^2_0	336.94***	306.55***	280.62***
In rate of change	σ^2_1	114.42***	114.17***	88.83***
In acceleration	σ^2_2	3.19***	3.18***	2.65***
Level 3				
Between-schools variance	σ^2_3	561.49***	555.17***	557.24***
Goodness of fit				
Deviance		83,192.7	83,157.7	83,072
AIC		83,208.7	83,173.7	83,088
BIC		83,234.4	83,199.4	83,113.7
GLH tests			Compare M1 H_0: $\beta_{010} = 0$; $\chi^2(1) = 35$, $p < .05$; Reject H_0	Compare M2 H_0: $\beta_{110} = 0$; $\chi^2(3) = 85.7$, $p < .05$; Reject H_0

Note. N students = 3,702; N schools = 193. AIC = Akaike information criterion; BIC = Bayesian information criterion; GLH = general linear hypothesis; M = model; H_0 = the null hypothesis.
** $p < .01$. *** $p < .001$.

foreign-born ELLs demonstrated basic proficiency on the MEPA (375) near the end of their 3rd year of high school.

Threats to Validity

Students who were missing MEPA outcome data in their junior and senior years (Spring 2007 and 2008, respectively) had higher MEPA scores at baseline, on average (although still corresponding only to the early intermediate level according to official MEPA guidelines), had higher than beginner level academic proficiency at the beginning of ninth grade, were more likely to be U.S.-born and Spanish-speaking, and were slightly less likely to be from low-income families, compared with those students who were not missing MEPA scores during these two periods.

In Table 6, I present the results of my sensitivity analyses in which I examine the nature and impact of sample attrition. First, the second and third columns of the table compare baseline characteristics for the entire sample of students who were missing values of the outcome for Spring 2007 and Spring 2008 and for those who were not missing data during these respective periods. Students with missing values of the outcome were higher MEPA scorers (although again, still scoring at the early intermediate level on average), U.S.-born, and Spanish speaking. I found the same pattern when I examined baseline characteristics for students missing values of the outcome in Spring 2007 or Spring 2008 alone compared to those who were not missing test score data for the respective periods (not displayed in Table 6). If students with missing MEPA data tended to have higher MEPA scores, of course, including them in the central analysis would have elevated the growth trajectories presented here. These patterns remained stable even after excluding the R-FEP students from the analysis.

The last two columns in Table 6 show that the average baseline MEPA score after excluding the R-FEP students is reduced (332), compared with estimates of the MEPA baseline scores across the whole sample (ranging from 346–348). However, while the non-R-FEP students with missing outcome values scored slightly higher at baseline than did the students with no missing values, they still scored in the early intermediate range at baseline. The large decline in average scores after excluding the R-FEP students suggests that they were in part driving the association between higher MEPA scores and the probability of missing MEPA outcome data in the whole sample. However, even after controlling for R-FEP placement, more U.S.-born, Spanish-speaking students were missing outcome values, implying that this subset of ELLs are at the greatest risk of dropping out of high school or leaving the state. Finally, given that the average non-R-FEP student with missing MEPA data scored only slightly higher than the average

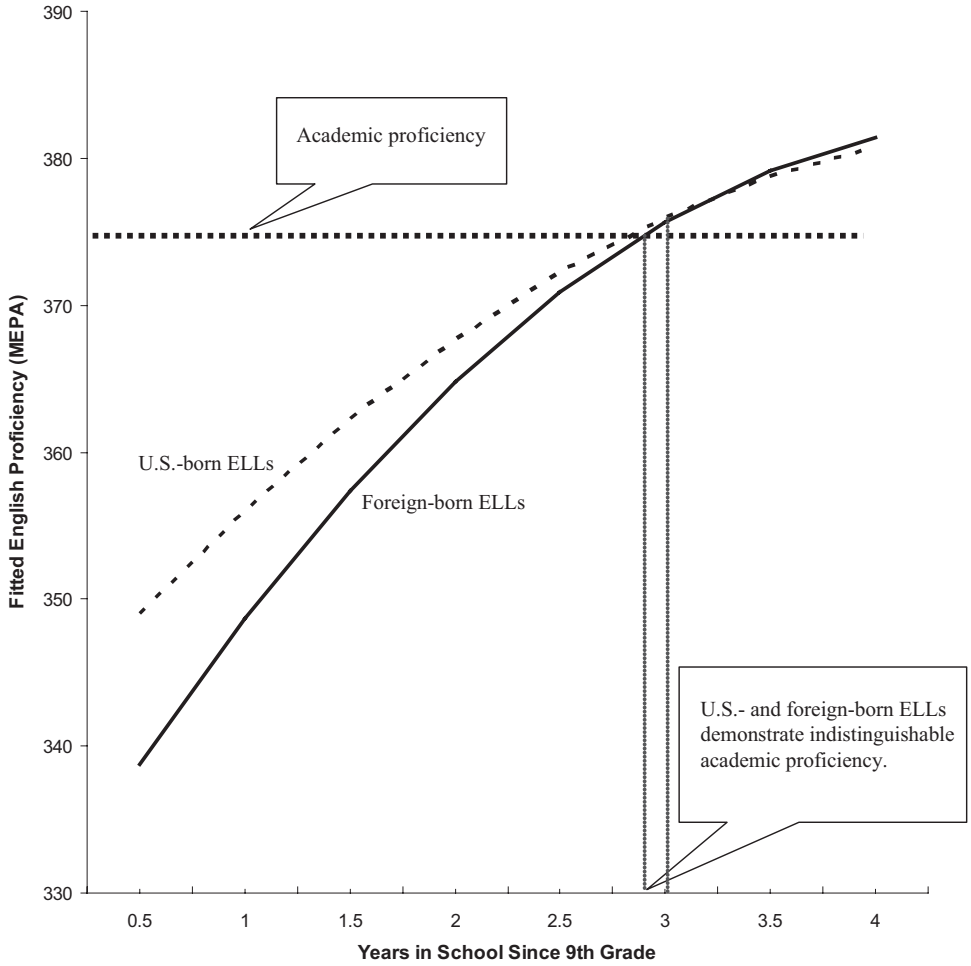

Figure 4. Fitted growth trajectories in performance on the *Massachusetts English Proficiency Assessment* (MEPA; from Model 3 [U.S. and foreign-born English language learners]). *n* students = 3,702; *n* schools = 193. ELLs = English language learners.

non-R-FEP student with MEPA scores, it does not appear that the missing outcome data biased the growth trajectory estimates presented in this study.

Discussion

Three important findings emerge from the present study. First, the average ELL started off high school while just beginning to develop academic English proficiency with MEPA scores that correspond to the early intermediate level of proficiency (see Table A1 in Appendix A). However, even though those who persisted through high school continued to show improvement in their academic English, these ELLs typically did not acquire a basic minimum English proficiency required to be successful in mainstream classrooms until finishing their penultimate year of high school. My second finding was that at high school entrance, U.S.-born ELLs had higher levels of academic English proficiency than did their foreign-born ELL peers (although still scoring at the early intermediate level), but foreign-born ELLs developed their academic proficiency skills at a faster rate so that by the end of high school they had caught up to their U.S.-born peers on measures of academic English proficiency. However, both groups remained at low levels of academic proficiency throughout most of high school and, thus, at a major academic disadvantage. Further, U.S.-born ELLs are likely to have spent 9 or more years in U.S. schools without developing sufficient academic language needed to perform grade-level academic work in English successfully. Finally, these findings indicate that there is important heterogeneity among high school ELLs, but on the basis of these students' intermediate level of academic English, both U.S.-born and foreign-born adolescent ELLs may face bleak secondary and postsecondary outcomes.

Despite Relative Growth, Low Levels of Academic English Bode Poorly for ELLs

The average high school ELL in my sample started off ninth grade while just beginning to develop academic English proficiency and initially acquired academic English at a rate of .4 of a standard deviation per year (fall baseline *SD* = 32.02). More

Table 5
Predicted Massachusetts English Proficiency Assessment (MEPA) Scaled Scores and Predicted Effect Sizes in Academic English Proficiency for Adolescent U.S.-Born and Foreign-Born English Language Learners (ELLs)

Time period	U.S.-born ELLs	Foreign-born ELLs	Effect size (Cohen's *d*)
Fall 2004	350.99	338.23	+0.40
Spring 2005	356.55	348.85	+0.25
Spring 2006	366.04	365.73	+0.01
Spring 2007	373.39	376.82	−0.13
Spring 2008	378.60	382.12	−0.14

Note. Adolescent U.S.-born $N = 1{,}590$; foreign-born $N = 1{,}258$. Estimates of MEPA scaled scores are based on predicted values of the outcome variable obtained from fitted Model 3 in Table 4. Standard deviations used to calculate effect sizes were based on the sample standard deviation of the outcome variable at each respective time period: $SD_{\text{Fall '04}} = 32.02$; $SD_{\text{Spring '05}} = 30.49$; $SD_{\text{Spring '06}} = 29.04$; $SD_{\text{Spring '07}} = 26.37$; $SD_{\text{Spring '08}} = 24.92$. All effect sizes were calculated with U.S.-born students as the reference, such that positive signs indicate higher U.S.-born performance and negative signs indicate higher foreign-born student performance.

concretely, on the basis of the state-determined performance indicators, the average ELL began high school able to read, write, and speak using common words and phrases in English, but with frequent errors, able to read and comprehend below-grade level text, and able to understand basic spoken vocabulary in English (see Table A1 in the Appendix). Despite initial steady growth, improvement in academic English proficiency for ELLs slowed during the last year of high school, and the average ELL who persisted through high school did not reach a minimum level of academic English proficiency considered sufficient to participate in mainstream classrooms until the end of their third year of high school. However, the finding that the average ELL reached proficiency by the end of junior year should be interpreted cautiously, given the substantial number of students in this cohort who are missing MEPA data at later waves and who are potentially among the many high school ELLs who drop out of school (Suárez-Orozco & Suárez-Orozco, 2001; White & Kaufman, 1997). Thus, the present study likely underestimates the average time to proficiency. Further, even for those who do reach the minimum bar denoting academic English proficiency on the academic language proficiency assessment, there is little empirical evidence that a minimum proficiency score will confer academic success in the mainstream classroom (Mahoney & MacSwan, 2005). In effect, language proficiency assessments represent a low bar for measuring ELLs' academic outcomes, and scoring at the proficient level is only one step toward the possibility of academic success.

Adolescent ELLs Are at High Risk for Educational Failure

The finding that the average ELL started off high school while just beginning to develop academic proficiency in one state serves as a case example for examining an overall portrait of risk for adolescent second language learners. Beginning secondary school with below grade-level academic English may limit the educational opportunities for the students in the present sample and, in particular, for those students who may have spent many years in U.S. schools without developing sufficient academic English to perform routine work in mainstream classrooms independently. Specifically, these students are tasked with "double the work" in order to catch up with their peers: Because these students began high school performing at an early intermediate level on the MEPA, they must continue to develop academic English proficiency and content knowledge simultaneously (Short & Fitzsimmons, 2007). Further, these learners face bleak secondary and postsecondary outcomes because continued enrollment in specialized language learning services at the secondary level may not give them access to college-preparatory curricula or sufficient course credits for on-time graduation. Moreover, because the average student in this sample scored at the intermediate level of English proficiency by the end of their high school trajectories, they are at risk of not passing mandatory high school exit exams (MDESE, 2010); in addition to academic language, these tests also contain test language, a specific register of academic language with particular grammatical structures and expressions (e.g., *which of these best describes*, as cited in Stevens, Butler, & Castellon-

Table 6
Sample Means and Differences Between the Means of Selected Baseline Student-Level Demographic Characteristics for ELLs Who Were Missing Massachusetts English Proficiency Assessment (MEPA) Data in Spring 2007 and Spring 2008 for the Entire Sample and Excluding All R-FEP Students

Variable	Whole sample, including R-FEP students		Excluding R-FEP students	
Baseline student characteristic	Missing Spring 2007 and 2008 MEPA	Not missing Spring 2007 and 2008 MEPA	Missing Spring 2007 and 2008 MEPA	Not missing Spring 2007 and 2008 MEPA
n	2,515	1,187	1,322	942
Fall 2004 MEPA scaled score	348.30***	331.10	331.9**	328.3
Beginner level	0.27	0.50***	0.41	0.54***
U.S.-born	0.64***	0.49	0.63***	0.47
Low income	0.83	0.87*	0.80	0.87***
Spanish-speaking	0.59*	0.55	0.66***	0.55

Note. Entire sample $n = 3{,}702$; excluding all R-FEP students $n = 2{,}264$. The following variables are dichotomous and thus their mean can be interpreted as a proportion: beginner level academic proficiency, US.-born, low income, and Spanish-speaking. R-FEP = reclassified fluent English proficient.
* $p < .05$. ** $p < .01$. *** $p < .001$. (Indicates that differences in means are significant at the following levels, where asterisks denote the greater mean.)

Wellington, 2000). If obtaining a high school diploma does not seem feasible to these students, they may see dropping out of high school as a viable option. Additionally, because the average ELL in this sample does not reach a minimum level of academic English proficiency—one indicator of their potential to participate fully in mainstream classrooms and access grade-level content-area instruction—until the end of 11th grade, it is unlikely that they will have been able to master grade-level academic courses that fulfill meaningful graduation requirements.

Long-term ELLs—defined in many school districts as those students who remain designated as ELL after 6 or more years in U.S. classrooms (New York City Department of Education, 2000)—are at particular risk for educational failure because they typically continue to perform several years below grade level in academic English skills such as reading, comprehension, and writing, despite having spent many years receiving specialized language learning services (Ruiz-de-Velasco, Fix, & Chu Clewell, 2000). The case of the long-term ELL in the United States highlights the importance of longitudinal analysis as a promising approach to monitor the academic growth, both in the societal language and content-area performance for the at-risk population of adolescent second language learners.

Foreign-Born ELLs Catch up to U.S.-Born Long-Term ELL Peers

In the fall of their ninth grade, U.S.-born ELLs demonstrated higher levels of academic-English proficiency than did their ELL classmates born outside of the U.S., but foreign-born ELLs developed academic language at a faster initial pace than did their native-born ELL peers. On the basis of this faster initial rate of growth, by the end of high school, foreign and U.S.-born ELLs demonstrated academic English proficiency that was indistinguishable from one another. Based on many years spent in U.S. schools, it is not surprising that U.S.-born ELLs had already acquired some academic English at high school entrance. This finding is consistent with prior work indicating that time in the United States is associated with improved reading skills in English (e.g., Kao & Tienda, 1995; Rumbaut, 1995). Further, second and higher generation ELLs likely do not have to struggle to acquire an entirely new language, compared with their first-generation peers (Suárez-Orozco et al., 2008).

However, though U.S.-born ELLs seem to develop a minimum level of academic English through their time spent in U.S. schools, they appear to lose ground in their academic English-language development, compared with their foreign-born ELL peers. Further, assuming that U.S.-born ELLs have attended U.S. schools since kindergarten, the present estimates of the time to develop academic English proficiency to be successful in the mainstream secondary school classroom—2.8 years since ninth grade or 11.8 school years since their entry into U.S. schools—exceed Hakuta and colleagues' (2000) 4–7 year estimates and are more in line with their projected 10 year estimates.

The present study's estimates of the time it takes these learners to develop academic English proficiency, may signal that U.S.-born ELLs have reached a "ceiling" in their academic language development. Specifically, ELLs who persisted until 12th grade remained at an intermediate level of English proficiency (see Table A1; Appendix A). Similar to the finding in the present study, other states and districts across the United States have reported that large numbers of ELLs remain at the intermediate level of English proficiency for many years. Researchers working closely with the English Language Learner Task Force in Arizona—one of the states with the largest proportion of ELL students nationwide (Kindler, 2002)—have reported that up to 60% of students in their districts reached an intermediate level of academic English proficiency and then stopped making progress (Clark, 2009). Similarly, 6% of New York's total student population received specialized services for 7 or more years without developing sufficient academic English proficiency to participate fully in mainstream classrooms (Newell & Smith, 1999).

The finding that 60% of ninth grade ELLs were U.S.-born is consistent with nationwide trends indicating that large percentages of middle and high school ELLs are U.S.-born (Batalova, Fix, & Murray, 2007; Flores, Painter, Harlow-Nash, & Pachon, 2009; Goldenberg, 2008; MDESE, 2009)—meaning that they have spent anywhere from 6–12 years in U.S. schools without acquiring sufficient academic English to be successful in mainstream classrooms. The length of time that ELLs spend in an instructional support program can have a big impact on their academic achievement (Ragan & Lesaux, 2006), and evidence from other U.S. states indicates that large numbers of ELLs may be remaining too long in language learning programs (e.g., California; Wong-Fillmore & Snow, 2000). In the present sample, large proportions of ELLs spent their entire high school careers in language learning programs (see Table C1 in Appendix C).

Foreign-born ELLs may demonstrate more rapid progress in academic English development simply because they begin at a lower level, yet the mechanisms explaining how foreign-born ELLs ultimately "catch up" to their native-born peers (albeit to an overall low bar of academic success) are less evident and serve to generate hypotheses for future research. It is difficult to determine if the estimated time to grade-level English proficiency for foreign-born ELLs were within the average range based on estimates from prior work because there is no reliable data on how long these students have attended U.S. schools. For those foreign-born ELLs who immigrated to the United States in middle school, the present estimates of time to academic English proficiency may indicate that these learners were relatively efficient in developing academic English, compared with their U.S.-born ELL peers. The relatively faster rate of growth in academic English language development of foreign-born students compared with their U.S.-born ELL classmates might be explained by the recently arrived immigrants' higher levels of academic motivation and educational aspirations (Rumbaut, 1995), better overall health and psychological well-being (NRC, 1998), more developed first-language skills (for earlier theories, see Cummins, 1981b; Cummins, 1979), or greater family financial and educational background resources (Suárez-Orozco et al., 2008). However, despite foreign-born ELLs' relative growth compared with that of their U.S.-born peers, the bottom line is that both groups remain at risk for academic failure on the basis of their academic language proficiency scores.

Implications and Directions for Future Research

An important implication of the present study is that large numbers of both U.S.-born and foreign-born ELLs entered high school with low levels of academic English proficiency, putting

them at increased risk of dropping out of school, failing to pass high-stakes exit examinations, and failing to compete in the new global, language-based economy. Further, my study highlights a growing national trend—that a majority (60%) of high school ELLs in one U.S. state were born in the U.S.—implying that large numbers of these students have spent 9 or more years in U.S. schools without developing a minimum level of academic language needed to perform mainstream academic work in English. Moreover, those ELLs who do reach the minimum criteria for reclassification as fluent English proficient do not do so until late in their high school trajectories and are not guaranteed to be successful in mainstream classrooms academically.

Future empirical work should prioritize the development of instructional interventions to equip adolescent ELLs with the academic language skills necessary to complete secondary and postsecondary school, with particular attention to the heterogeneity among the population of older ELLs. The research base on successful academic language interventions for adolescent ELLs is limited (Francis & Vaughn, 2009), however several recent, empirically rigorous intervention studies targeting academic language development in literacy and content areas in the middle grades (sixth through eighth grade) signal promising gains in academic language for ELLs and their classmates from linguistically diverse backgrounds (e.g., August, Branum-Martin, Cardenas-Hagan, & Francis, 2009; Lesaux, Kieffer, Faller, & Kelly, 2010; Snow, Lawrence, & White, 2009; Townsend & Collins, 2009; Vaughn et al., 2009). Although these interventions demonstrated significant improvement in academic language development for middle grade students, not all were designed specifically for ELLs. However, based on the program descriptions detailed in each of these studies, the interventions draw on similar skills required of ELL students to perform at the transitioning level (and beyond) on the MEPA and could potentially be adapted for ELLs. Specifically, at the transitioning level, students must recognize academic words and comprehend moderately difficult grade-level text, write short compositions using standard English conventions, use complex grammar and sentence structures, and understand prolonged oral communication with little or no need for clarification (see Appendix A). However, none of these programs were implemented with high school students, and none specifically addressed the needs of long-term ELLs.

The collection of more fine-grained, nationally representative longitudinal data on all second language learners could help researchers develop and adapt more targeted interventions (paired with rigorous evaluations) for students most at risk of failing to develop sufficient academic language for school success, including high school ELLs and long-term ELLs. Stakeholders should consider collecting demographic data on ELLs' educational background, including time spent in schools in the United States and abroad, degree of international and domestic mobility, generational status, parents' educational attainment, and students' degree of exposure to conversational and academic English outside of school. These data are routinely collected in many states in the form of locally administered home-language surveys but should be stored in state and national databases and longitudinally linked to student academic English proficiency and standardized assessment data to monitor language development over time for the students most at-risk for educational failure.

References

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness, 2,* 345–376. doi:10.1080/19345740903217623

Batalova, J., Fix, M., & Murray, J. (2007). *Measures of change: The demography and literacy of adolescent English language learners—A report to the Carnegie Corporation of New York.* Washington, DC: National Center on Immigration Integration Policy, Migration, Migration Policy Institute.

Belfield, C. R., & Levin, H. M. (2007). The education attainment gap: Who's affected, how much, and why it matters. In C. R. Belfield, H. M. Levin, & Brookings Institution (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 1–20). Washington, DC: Brookings Institution Press.

Buriel, R., & Cardoza, D. (1988). Sociocultural correlates of achievement among three generations of Mexican American high school seniors. *American Educational Research Journal, 25,* 177–192.

Cahill, D. P. (1987). Bilingual development of Italo-Australian children. *Australian Review of Applied Linguistics, 4,* 101–127.

Callahan, R., Wilkinson, L., Muller, C., & Frisco, M. (2009). ESL placement and schools: Effects on immigrant achievement. *Educational Policy, 23,* 355–384. doi:10.1177/0895904807310034

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.* Washington, DC: The Urban Institute.

Center on Education Policy. (2008). *State high school exit exams: Moving toward end-of-course exams.*

Clark, K. (2009). The case for structured English immersion. *Educational Leadership, 66,* 42–46.

Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21,* 617–641. doi:10.2307/3586986

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

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

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222–251.

Cummins, J. (1981a). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics, 11,* 132–149. doi:10.1093/applin/2.2.132

Cummins, J. (1981b). The role of primary language development in promoting educational success for language minority students. In D. P. Dolson (Eds.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Sacramento, CA: California State Department of Education.

Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review, 56,* 18–36.

Cummins, J. (2001). Empowering minority students: A framework for intervention. *Harvard Educational Review, 71,* 656–675.

Dolson, D. P. (1985). The effects of Spanish home language use on the scholastic performance of Hispanic pupils. *Journal of Multilingual and Multicultural Development, 6,* 135–155. doi:10.1080/01434632.1985.9994192

Driscoll, A. (1999). Risk of high school dropout among immigrant and native Hispanic youth. *International Migration Review, 33,* 857–875.

Fernandez, R. M., Paulsen, R., & Hirano-Nakanishi, M. (1989). Dropping out among Hispanic youth. *Social Science Research, 18,* 21–52. doi:10.1016/0049-089X(89)90002-1

Fernández-Kelly, M. P., & Schauffler, R. (1994). Divided fates: Immigrant children in a restructured U.S. economy. *International Migration Review, 28,* 662–689.

Flores, E., Painter, G., Harlow-Nash, Z., & Pachon, H. (2009). *Qué pasa? Are English language learning students remaining in English learning classes too long?* [Policy brief]. Los Angeles, CA: Tomas Rivera Policy Institute.

Francis, D. J., & Vaughn, S. (2009). Effective practices for English language learners in the middle grades: Introduction to the special issue of Journal of Research on Educational Effectiveness. *Journal of Research on Educational Effectiveness, 2,* 289–296. doi:10.1080/19345740903217664

García, E. G., McKoon, G., & August, D. (2006). Synthesis: Language and literacy assessment. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 583–596). Mahwah, NJ: Erlbaum.

Glick, J. E., & Hohmann-Marriott, B. (2007). Academic performance of young children in immigrant families: The significance of race, ethnicity, and national origins. *International Migration Review, 41,* 371–402. doi:10.1111/j.1747-7379.2007.00072.x

Glick, J. E., & White, M. J. (2003). The academic trajectories of immigrant youths: Analysis within and across cohorts. *Demography, 40,* 759–783. doi:10.1353/dem.2003.0034

Goldenberg, C. (2008). Improving achievement for English language learners. In S. B. Neuman (Ed.), *Educating the other America: Top experts tackle poverty, literacy, and achievement in our schools* (pp. 139–162). Baltimore, MD: Brookes.

Goldenberg, C., Rueda, R. S., & August, D. (2006). Sociocultural influences on the literacy attainment of language minority children and youth. In D. August & T. Shanahan (Ed.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth* (pp. 249–267). Mahwah, NJ: Erlbaum.

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

Hancock, D. R. (2002). The effects of native language books on the preliteracy skill development of language minority kindergarteners. *Journal of Research in Childhood Education, 17,* 62–68. doi:10.1080/02568540209594999

Ima, K., & Rumbaut, R. G. (1989). Southeast Asian refugees in American schools: A comparison of fluent-English-proficient and limited-English-proficient students. *Topics in Language Disorders, 9,* 54–75. doi:10.1097/00011363-198906000-00008

Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing.* Lanham, MD: Rowman & Littlefield.

Kao, G., & Tienda, M. (1995). Optimism and achievement: The educational performance of immigrant youth. *Social Science Quarterly, 76,* 1–19.

Kasinitz, P., Mollenkopf, M. W., & Holdaway, J. (2008). Culture matters. In P. Kasinitz, J. H. Mollenkopf, M. C. Waters, & J. Holdaway (Eds.), *Inheriting the city: The children of immigrants come of age* (pp. 241–272) Cambridge, MA: Harvard University Press.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100,* 851–868. doi:10.1037/0022-0663.100.4.851

Kieffer, M. J., Lesaux, N. K., & Snow, C. E. (2008). Promises and pitfalls: Implications of NCLB for identifying, assessing, and educating English language learners. In G. L. Sunderman (Ed.), *Holding NCLB accountable: Achieving accountability, equity, & school reform* (pp. 57–74). Thousand Oaks, CA: Corwin Press.

Kindler, A. L. (2002). *Survey of the states' limited English proficient students and available educational programs and services 2000–2001 summary report*. Retrieved from National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs website: http://www.ncela.gwu.edu/policy/states/reports/seareports/0001/sea0001.pdf

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for sixth graders in urban middle schools. *Reading Research Quarterly, 45,* 196–228. doi:10.1598/RRQ.45.2.3

Mahoney, K. S., & MacSwan, J. (2005). Reexamining identification and reclassification of English language learners: A critical discussion of select state practices. *Bilingual Research Journal, 29,* 31–42. doi:10.1080/15235882.2005.10162822

Massachusetts Department of Elementary and Secondary Education. (2005). *Massachusetts English Proficiency Assessment* (Technical report). Retrieved from http://iservices.measuredprogress.org/files/MCAS/MEPA_Report_Final.pdf

Massachusetts Department of Elementary and Secondary Education. (2009). *Student Information Management System data handbook* (Version 2.1). Retrieved from http://www.doe.mass.edu/infoservices/data/sims/

Massachusetts Department of Elementary and Secondary Education. (2010). *Requirements for participation of students with limited English proficiency in MCAS and MEPA.* Retrieved from www.doe.mass.edu/mcas/participation/lep.doc

National Center for Education Statistics. (2004). *Language minority learners and their labor market indicators: Recent trends.* Washington, DC: United States Department of Education.

National Research Council. (1998). *From generation to generation: The health and well-being of children in immigrant families.* Washington, DC: National Academy Press.

New York City Board of Education Division of Assessment and Accountability. (2000). *English language learner (ELL) subcommittee research studies progress report*. Retrieved from http://schools.nyc.gov/daa/reports/ELL_Research_Studies.pdf

Newell, J., & Smith, J. (1999, March). *Academic success for long-term ESL students*. Paper presented at the 33rd Annual Meeting of the Teachers of English to Speakers of Other Languages, New York, NY.

Ogbu, J., & Matute-Bianchi, M. (1986). Understanding sociocultural factors: Knowledge, identity, and school adjustment. In C. F. Leyba (Eds.), *California State Department of Education. Beyond language: Social and cultural factors in schooling language minority students* (pp. 73–142). Los Angeles, CA: Evaluation, Dissemination and Assessment Center, California State University.

Ogbu, J. U. (1987). Variability in minority school performance: A problem in search of an explanation. *Anthropology and Education Quarterly, 18,* 312–334. doi:10.1525/aeq.1987.18.4.04x0022v

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health, and happiness from compulsory schooling. *Journal of Political Economy, 91,* 2213–2229.

Orfield, G. (2001). *Schools more separate: Consequences of a decade of resegregation.* Cambridge, MA: Harvard Civil Rights Project. Retrieved from http://www.civilrightsproject.ucla.edu/research/deseg/separate_schools01.php

Portes, A., & Rumbaut, R. G. (1990). *Immigrant America: A portrait.* Berkeley, CA: University of California Press.

Portes, A., & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation.* Berkeley, CA: University of California Press.

Portes, A., & Schauffler, R. (1996). Language and the second generation: Bilingualism yesterday and today. In A. Portes (Ed.), *The new second generation* (pp. 246). New York, NY: Russell Sage Foundation.

Ragan, A., & Lesaux, N. K. (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. Retrieved from http://epaa.asu.edu/ojs/article/view/91/217

Ruiz-de-Velasco, J., Fix, M., & Chu Clewell, B. (2000). *Overlooked and underserved: Immigrant students in U.S. secondary schools.* Washington, DC: The Urban Institute.

Rumbaut, R. (1995). The new Californians: Comparative research findings on the educational progress of immigrant children. In R. G. Rumbaut & W. A. Cornelius (Eds.), *California's immigrant children: Theory, research, and implications for educational policy.* San Diego, CA: Center for U. S.–Mexican Studies, University of California, San Diego.

Rumberger, R. W., & Lamb, S. P. (2003). The early employment and further education experiences of high school dropouts: A comparative study of the United States and Australia. *Economics of Education Review, 22,* 353–366. doi:10.1016/S0272-7757(02)00038-9

Scarcella, R. (2003). *Academic English: A conceptual framework* (University of California Linguistic Minority Research Institute Technical Report 2003–1). Retrieved from http://lmri.ucsb.edu/publications/03_scarcella.pdf.

Short, D. J., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners.* New York, NY: Carnegie Corporation of New York.

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

Snow, C. E., & Kim, Y. S. (2007). Large problem spaces: The challenges of vocabulary for English language learners. In A. M. R. Wagner & K. K. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 123–139). New York, NY: Guilford Press.

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2,* 325–344. doi:10.1080/19345740903167042

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In N. Torrance & D. R. Olson (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge, England: Cambridge University Press.

Stevens, R., Butler, F. A., & Castellon-Wellington, M. (2000). *Academic language and content assessment: Measuring the progress of ELLs* (CSE Technical Report No. 552). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Suárez-Orozco, C., Suárez-Orozco, M., & Todorova, I. (2008). *Learning a new land.* Cambridge, MA: Harvard University Press.

Suárez-Orozco, C. M., & Suárez-Orozco, M. (1996). *Transformations: Immigration, family life, and achievement motivation among Latino adolescents.* Stanford, CA: Stanford University Press.

Suárez-Orozco, C. & Suárez-Orozco, M. M. (2001). *Children of immigration.* Cambridge, MA: Harvard University Press.

Thomas, R. M. (2005). *High-stakes testing: Coping with collateral damage.* Mahwah, NJ: Erlbaum.

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing, 22,* 993–1019. doi:10.1007/s11145-008-9141-y

United States Department of Education. (2009). *Fact sheet: NCLB provisions ensure flexibility and accountability for limited English proficient students*. Retrieved from http://www.ed.gov/nclb/accountability/schools/factsheet-english.html

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research on Educational Effectiveness, 2,* 297–324. doi:10.1080/19345740903167018

Velez, W. (1989). High school attrition among Hispanic and non-Hispanic White youths. *Sociology of Education, 62,* 119–133. doi:10.2307/2112844

White, M. J., & Kaufman, G. (1997). Language usage, social capital, and school completion among immigrants and native-born ethnic groups. *Social Science Quarterly, 78,* 385–398.

Wong-Fillmore, L., & Snow, C. E. (2000). *What teachers need to know about language.* Washington, DC: ERIC Clearinghouse on Language and Linguistics.

(*Appendices follow*)

Appendix A

Table A1
Massachusetts Department of Elementary and Secondary Education Performance Level Descriptors on the 2004–2008 Massachusetts English Proficiency Assessment (Ninth–12th Grade Span)

Performance level	Scaled score	Reading tasks	Writing tasks	Speaking tasks	Listening tasks
Beginning	300–324	Student recognizes simple written words and phrases.	Student writes basic words or phrases, with frequent errors.	Student speaks using basic words or phrases, with frequent errors.	Student understands basic spoken vocabulary or phrases.
Early intermediate	325–342	Student recognizes simple written words and phrases; reading comprehension is below grade level.	Student writes short paragraphs with limited control of standard English conventions.	Student speaks using common words and simple phrases, but word choice often incorrect.	Student understands basic spoken vocabulary and phrases, with constant need for clarification.
Intermediate	343–374	Student recognizes common written words, some academic words, and comprehends simple grade-level text.	Student writes short, simple compositions with partial control of standard English conventions.	Student speaks using common words and phrases, basic and complex grammar and sentence structure, but with errors.	Student understands oral communication with some need for clarification.
Transitioning	375–400	Student recognizes most common and academic words, and can comprehend moderately difficult grade-level texts.	Student writes short compositions, demonstrating control of standard English conventions.	Student speaks using appropriate and correct words, phrases, and expressions, and basic and complex grammar and sentence structures.	Student understands prolonged oral communication with little or no need for clarification.

Note. *N* students = 3,702; *N* schools = 193. Adapted from the 2005 *Massachusetts English Proficiency Assessment* (MEPA), by Massachusetts Department of Elementary and Secondary Education, 2005, Retrieved from http://iservices.measuredprogress.org/files/MCAS/MEPA_Report_Final.pdf.

(*Appendices continue*)

Appendix B

Table B1
Sample Frequencies and Percentages of Select Demographic Characteristics of Ninth-12th Grade Test Takers of the Massachusetts English Proficiency Assessment (MEPA), by Wave of Data Collection

Variable	Fall 2004	Spring 2005	Spring 2006	Spring 2007	Spring 2008
Federal designation					
Limited English proficient	3,702	3,483	2,493	1,629	1,029
%	100	94.19	78.27	38.87	48.08
Low income	2,873	2,907	2,369	1,858	1,487
%	77.6	78.61	74.38	69.72	69.49
Program status					
Sheltered English immersion	2,821	2,758	2,186	1,276	810
%	76.20	74.58	68.63	47.88	37.85
Bilingual program	258	238	145	96	25
%	6.97	6.44	4.55	3.60	1.17
Opted out of services	198	75	66	195	136
%	5.35	2.03	2.07	7.32	6.36
Not enrolled in language program	425	627	788	1,098	1,169
%	11.48	16.96	24.74	41.20	54.63
Language group[a]					
Spanish	2,144	—	—	—	—
%	57.91	—	—	—	—
Portuguese	367	—	—	—	—
%	9.91	—	—	—	—
Other[b]	300	—	—	—	—
%	8.1	—	—	—	—
Haitian Creole	190	—	—	—	—
%	5.13	—	—	—	—
Cape Verdean	159	—	—	—	—
%	4.29	—	—	—	—
Khmer	152	—	—	—	—
%	4.11	—	—	—	—
U.S.-born[c]	2,186	—	—	—	—
%	59.05	—	—	—	—

Note. *N* students = 3,702; *N* schools = 193; *N* observations = 18,510.
[a] I treated demographic characteristics such as language group as time-invariant; therefore, I only list values for Fall 2004. The dashes in subsequent columns indicate that the language group variables were measured only at baseline. [b] The Other category includes 49 students whom the Department of Education listed as a speaking a language not listed in their data collection codebook, in addition to 251 students who were members of language groups with less than 20 speakers. [c] In addition to U.S.-born students, this group also includes students born to U.S. citizens who are Embassy or military personnel, students born abroad to U.S. citizens, and students who are foreign exchange students.

(*Appendices continue*)

Appendix C

Table C1
Most Common Language Learning Program Profiles for the 2004 Ninth Grade ELL Cohort

Program profile	U.S.-born ELL	Foreign-born ELL	All ELLs
4 years no program (*N*)	117	—	125
1 year sheltered English immersion, then exit (*N*)	151	127	278
2 years sheltered English immersion, then exit (*N*)	142	116	258
3 years sheltered English immersion, then exit	108	101	209
4 years sheltered English immersion (no exit) (*N*)	303	299	602

Note. *N* schools = 193; *N* observations = 18,510. While I chose the most common program profiles among the cohort of 3,702 English language learners (ELLs), these profiles represent 66% of the ELLs in the 2004 ninth grade ELL cohort (n = 2,452). A dash indicates numbers that were too small to report by Massachusetts Department of Elementary and Secondary Education policy.

Appendix D

Statistical Model Specifications by Research Question

Research Question 1. What is the shape of the growth trajectories in academic English proficiency for English-language learners (ELLs) between ninth and 12th grade?

In the Level-1, Level-2 and Level-3 submodels below, $MEPA_{ijk}$ is student *j's* value of their English proficiency score on occasion *k*, attending school *i*:
Level-1 model:

$$MEPA_{ijk} = \pi_{0ij} + \pi_{1ij}YEAR_{ijk} + \pi_{2ij}YEAR^2_{ijk} + \varepsilon_{ijk}, \textit{ where}$$

$$\varepsilon_{ijk} \sim N(0,\sigma^2_\varepsilon).$$

Level-2 model:

$$\pi_{0ij} = \gamma_{00i} + \zeta_{0ij},$$

$$\pi_{1ij} = \gamma_{10i} + \zeta_{1ij}, \textit{ and}$$

$$\pi_{2ij} = \gamma_{20i} + \zeta_{2ij}, \textit{ where}$$

$$\begin{bmatrix}\zeta_{0ij}\\ \zeta_{1ij}\\ \zeta_{2ij}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_0 & \sigma_{01} & \sigma_{02}\\ \sigma_{10} & \sigma^2_1 & \sigma_{12}\\ \sigma_{20} & \sigma_{21} & \sigma^2_2\end{bmatrix}\right).$$

Level-3 model:

$$\gamma_{00i} = \beta_{000} + u_i, \textit{ where } u_i \sim N(0,\sigma^2_u),$$

$$\gamma_{10i} = \beta_{100}, \textit{ and}$$

$$\gamma_{20i} = \beta_{200}.$$

In order to represent the multilevel model for change more parsimoniously, I collapsed the three submodels above algebraically into the following composite multilevel model (Singer & Willett, 2003):

$$MEPA_{ijk} = \beta_{000} + \beta_{100}YEAR_{ijk} + \beta_{200}YEAR^2_{ijk} + [\varepsilon_{ijk} + \zeta_{0ij} + \zeta_{1ij}YEAR_{ij} + \zeta_{2ij}YEAR^2_{ij} + u_i], \textit{ where}$$

$$\varepsilon_{ijk} \sim N(0, \sigma^2_\varepsilon),$$

$$\begin{bmatrix}\zeta_{0ij}\\ \zeta_{1ij}\\ \zeta_{2ij}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_0 & \sigma_{01} & \sigma_{02}\\ \sigma_{10} & \sigma^2_1 & \sigma_{12}\\ \sigma_{20} & \sigma_{21} & \sigma^2_2\end{bmatrix}\right), \textit{ and}$$

$$u_i \sim N(0,\sigma^2_u). \quad (1)$$

Parameter β_{000} represented the population-average true initial status of academic English proficiency (in the fall of ninth grade), parameter β_{100} represented the population-average true initial rate of change, and parameter β_{200} represented the population-average true acceleration in academic English acquisition. Random effect ε_{ijk} was a Level-1 residual for student *j,* on occasion *k*, attending school *i*. Random effects ζ_{0ij}, ζ_{1ij} and ζ_{2ij} were Level 2 residuals for intercept, slope, and initial acceleration, respectively. Random effect u_i was a Level 3 residual for intercept. These residuals were hypothesized to be drawn independently and identically from their respective normal distributions, each with mean of zero, and unknown variances.

Research Question 2. Are the trajectories in academic English proficiency steeper for U.S.-born children of immigrants—those who have plausibly spent more time in U.S. classrooms—or for first-generation immigrants who have had more time to develop literacy in their first language?

(*Appendices continue*)

Level-1 model:

$$MEPA_{ijk} = \pi_{0ij} + \pi_{1ij}YEAR_{ijk} + \pi_{2ij}YEAR^2_{ijk} + \varepsilon_{ijk}, \textit{ where}$$

$$\varepsilon_{ijk} \sim N(0,\sigma^2_\varepsilon).$$

Level-2 model:

$$\pi_{0ij} = \gamma_{00i} + \gamma_{01i}USBORN_{ij} + \zeta_{0ij},$$

$$\pi_{1ij} = \gamma_{10i} + \gamma_{11i}USBORN_{ij} + \zeta_{1ij}, \textit{ and}$$

$$\pi_{2ij} = \gamma_{20i} + \gamma_{21i}USBORN_{ij} + \zeta_{2ij}, \textit{ where}$$

$$\begin{bmatrix}\zeta_{0ij}\\ \zeta_{1ij}\\ \zeta_{2ij}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_0 & \sigma_{01} & \sigma_{02}\\ \sigma_{10} & \sigma^2_1 & \sigma_{12}\\ \sigma_{20} & \sigma_{21} & \sigma^2_2\end{bmatrix}\right).$$

Level-3 model:

$$\gamma_{00i} = \beta_{000} + u_i, \textit{ where } u_i \sim N(0, \sigma^2_u),$$

$$\gamma_{01i} = \beta_{010},$$

$$\gamma_{10i} = \beta_{100},$$

$$\gamma_{11i} = \beta_{110},$$

$$\gamma_{20i} = \beta_{200}, \textit{ and}$$

$$\gamma_{21i} = \beta_{210}.$$

As above, I represented the Level-1, Level-2 and Level-3 submodels algebraically into the following composite multilevel model:

$$MEPA_{ijk} = \beta_{000} + \beta_{010}USBORN_{ij} + \beta_{100}YEAR_{ij} + \beta_{110}USBORN_{ij} \times YEAR_{ijk} + \beta_{200}YEAR^2_{ijk} + \beta_{210}USBORN_{ijk} \times YEAR^2_{ijk} + [\varepsilon_{ijk} + \zeta_{0ij} + \zeta_{1ij}YEAR_{ij} + \zeta_{2ij}YEAR^2_{ij} + u_i], \textit{ where}$$

$$\varepsilon_{ijk} \sim N(0,\sigma^2_\varepsilon),$$

$$\begin{bmatrix}\zeta_{0ij}\\ \zeta_{1ij}\\ \zeta_{2ij}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_0 & \sigma_{01} & \sigma_{02}\\ \sigma_{10} & \sigma^2_1 & \sigma_{12}\\ \sigma_{20} & \sigma_{21} & \sigma^2_2\end{bmatrix}\right), \textit{ and}$$

$$u_i \sim N(0,\sigma^2_u). \quad (2)$$

Parameter β_{000} represents the population-average true initial status in academic English proficiency for foreign-born ELLs (when U.S.-born = 0), β_{010} represents the population-average difference in true initial status between U.S.- and foreign-born ELLs, parameters β_{110}, and β_{210} represent population-average differences in initial instantaneous rates of growth and overall acceleration between U.S.- and foreign-born ELLs. The random effects and assumptions on the residuals are identical to those detailed in the composite specification of Research Question 1.

Received March 11, 2010
Revision received September 6, 2011
Accepted September 13, 2011 ■