ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-300

Dual Language Education and Academic Growth

Angela Johnson
NWEA

This paper reports math and reading academic achievement and growth in grades 2 to 8 for Hispanic participants and nonparticipants of a Spanish-English dual language program. I apply a piecewise multilevel growth model to administrative data from a large school district that enrolls a substantial English Learner student population. Dual language participants started 2nd grade with lower achievement than nonparticipants. In math, dual language participants grew faster than nonparticipants during each school year in grades 2 to 5 but lost more learning during subsequent summers. Thus, despite growing faster in the beginning, dual language students did not learn more than their peers in the long run, and the gap between dual language students and the national average was not closing. In reading, dual language participants grew slightly more slowly during school years but lost less learning during the summers, closing the gap between themselves and the national average. These findings suggest that programs aimed at addressing achievement gaps need to consider summer as well as school-year learning for historically-underserved student populations.

VERSION: September 2020

Suggested citation: Johnson, Angela. (2020). Dual Language Education and Academic Growth. (EdWorkingPaper: 20-300). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/7zth-7373

Dual Language Education and Academic Growth

Angela Johnson
NWEA

Abstract
This paper reports math and reading academic achievement and growth in grades 2 to 8 for Hispanic participants and nonparticipants of a Spanish-English dual language program. I apply a piecewise multilevel growth model to administrative data from a large school district that enrolls a substantial English Learner student population. Dual language participants started 2nd grade with lower achievement than nonparticipants. In math, dual language participants grew faster than nonparticipants during each school year in grades 2 to 5 but lost more learning during subsequent summers. Thus, despite growing faster in the beginning, dual language students did not learn more than their peers in the long run, and the gap between dual language student and the national average was not closing. In reading, dual language participants grew slightly more slowly during school years but lost less learning during the summers, closing the gap between themselves and the national average. These findings suggest that programs aimed at addressing achievement gaps need to consider summer as well as school-year learning for historically-underserved student populations.

Key Words
Dual language education, English Learners, Education policy

Acknowledgements
The author is grateful for continued support from the anonymous school district and valuable feedback from Claude Goldenberg, Nate Jensen, Megan Kuhfeld, and participants of NWEA Ideas in Research.

Dual Language Education and Academic Growth

Dual language education—an inclusive category for programs that provide academic content instruction in both English and a partner language—aims to foster the development of bilingualism, biliteracy, and academic skills in all school subjects. In the past few decades, a variety of program models have been implemented throughout the US and continue to expand. Dual language programs are currently implemented across 42 states and the District of Columbia; further, the programs are gaining prominence in policy contexts new to adopting dual language education, such as Delaware and New York City (Dual Language Schools, 2020).

Early waves of correlational research suggest that participation in dual language education is associated with higher achievement (Collier & Thomas, 2017). Recent studies leveraged more comprehensive sets of baseline characteristics to compare the achievement of students in dual language to other programs and found that dual language students improve their achievement ranks more than other students (Steele et al., 2017; Valentino & Reardon, 2015). A major limitation to these studies is their lack of ability to model within-student growth. Thus, we lack evidence on the relation between dual language education and growth trajectories within and across grades.

This paper addresses this critical gap in the research. Applying a piecewise multilevel model to rich assessment data on a large sample of Hispanic students (N=27,661), I compare the math and reading growth rates of participants to nonparticipants of a dual language program. This study is conducted in a research-practice partnership with a large suburban district in the Midwest. It makes three important contributions.

First, this study leverages vertically-scaled measures to report the first estimates of within- and across-year absolute growth in the dual language education literature. I observe

300,000 test scores each in math and reading, measured in fall, winter, and spring of each academic year from 2nd to 8th grade, at up to 20 time points per student. These unique data allow me to estimate growth rates in each grade, as well as loss rates in summers after each grade, disaggregating seasonal patterns of learning unfound in dual language education research. Second, the study context is novel to the body of large-scale research studies on dual language education. Most extant studies addressed dual language programs in California, North Carolina, Oregon, Texas, and Utah. This study contributes new evidence from a policy context that was largely unexplored in previous work. Third, situated in a research-practice partnership, this study offers opportunities for continued program evaluation and improvement. I discuss lessons learned from the partnership that will inform dual language program design and implementation across the nation.

Dual Language Program Features

The central feature of dual language programs is the provision of literacy and content instruction in both English and a partner language such as Spanish for promoting bilingualism, biliteracy, grade-level academic achievement, and sociocultural competencies (Howard et al., 2018). Two-way dual language programs enroll a mix of native users of English and native users of the partner language, with the idea that the two groups of students would model languages for each other. One-way programs use both English and the partner language to teach native users of the partner language. Both two-way and one-way programs are intended to support the content learning of students who are developing English proficiency, to whom federal education policy refer as English Learners (ELs).

In dual language programs, ELs receive home language support in content learning for an extended duration. Transfer theory and underlying proficiency theory explain that EL students acquire literacy and academic concepts more easily in their home language than in English, and when they learn new concepts, the knowledge is added to their repertoire of skills and transferred to English. Ultimately, the students gain literacy and academic skills in both English and their home language (Goldenberg, 2008). However, learning to read in two languages takes more time than learning to read in one. Thus, we might expect EL students in dual language programs to outperform their EL peers in other instructional services in reading and academic content not immediately but in the long run (Valentino & Reardon, 2015).

Instructional practices, curriculum, and classroom context may also explain how dual language programs reduce the achievement gap between ELs and English-proficient students. Dual language teachers, many of whom are bilingual, may be more likely to engage in culturally-responsive teaching than teachers in monolingual programs; students would also benefit from dual language curricula that effectively incorporate authentic, rigorous, and engaging materials from students' home language and culture (Thomas & Collier, 2012). Leveraging its bilingual staff, dual language programs might also be better than other programs at family engagement, which contributes to better student outcomes (Kraft & Dougherty, 2013; Kraft & Rogers, 2015). Family involvement can help students maintain or build on the skills they acquired in the classroom. But for students who use a language other than English at home, family involvement is often hindered by linguistic barriers (Vera et al., 2012). Dual language instruction enables families to use the home language to help students with homework, reading, and application of skills acquired in school to life-wide situations. With increased family involvement and opportunities for out-of-school learning, we might expect students in dual language programs to experience more academic growth every year, including out-of-school periods such as summer, compared to students who are taught only in English.

In addition, dual language programs can elevate the social status of the partner language and boost ELs' academic expectations and success (Valdés, 2017). In other words, dual language programs can reframe students' status from "English Learner" to "native users of the partner language," undoing negative effects of the EL label on students' outcomes (Umansky, 2016). In the case of two-way dual language programs, the placement of emergent and native English users in the same classroom can reduce linguistic and academic segregation. As a result, EL students in dual language programs can be exposed to higher-quality academic discourse than they would in EL-only classrooms. As the population of immigrant students in the US continues to grow, dual language programs offer a potentially effective approach to addressing their academic needs.

However, dual language program design and implementation, and consequently program quality, vary greatly by context (Howard et al., 2018). A range of factors crucial to program quality and student achievement differ across programs, such as the proportion of instruction allocated to each language (e.g., 90:10 or 50:50), the academic subjects taught in each language, the language proficiencies of its teachers, and curriculum materials (Valdés, 2017). Short-term programs, known as transitional, developmental, or early-exit, provide instruction in the partner language for no more than two or three years. These programs prioritize on helping ELs develop English proficiency and grade-level academic competencies over maintaining or developing proficiency or literacy in their home language (Valentino & Reardon, 2015). Long-term programs continue to use the partner language to teach academic subjects in middle or even high school, with the aim to develop proficiency and literacy in both English and the partner language. Among programs that span elementary to middle grades, the extent of partner language use also varies. Many programs phase out partner language instruction during late elementary school. As a result, students may develop interpersonal communication proficiency in the partner language but have limited opportunity to develop academic partner language skills or apply partner language proficiency in the learning of academic subjects.

Dual Language Participation and Academic Achievement

A mature line of descriptive studies has documented positive correlations between participation in dual language programs and academic achievement status (see Bialystok, 2018 and Howard et al., 2018 for reviews). Generally, research has found that longer program duration (at least six years of partner language instruction) is associated with better downstream achievement outcomes for ELs. For instance, across a large body of work, Collier and Thomas used multivariate regression to compare the achievement of dual language students with students in other programs, such as ESL and short-term home language education. In a recent review, they summarized this work by stating that they found dual language programs completed closed the achievement gap between ELs and English-proficient students, with effect sizes twice as large as those of other programs (2017). In a series of studies focused on schools in California, Lindholm-Leary and colleagues reported that Latino students in dual language programs performed better than Latino students in mainstream classes (e.g., Lindholm-Leary & Block, 2010; Lindholm-Leary & Hernández, 2011). However, many of the early studies either insufficiently addressed selection bias into dual language programs or used test scores that violated the assumption of vertical scaling.

Two recent studies leveraged rigorous experimental and quasi-experimental designs to provide robust estimates on the relation between dual language program participation and academic achievement. Steele and colleagues (2017) conducted a randomized experiment in Portland, Oregon, using a lottery to assign kindergarten students to a dual language immersion program. They found positive intent-to-treat effects on reading achievement in 5th (.13 standard deviations (SDs)) and 8th grade (.22 SDs) but no effect on math achievement. Valentino and Reardon (2015) applied multilevel modeling to achievement data from a district in California to examine the rate of change in standardized achievement measures across grade levels. Controlling for an extensive list of covariates, including parent program preference, they compared the spring California Standards Test (CST) scores of ELs who enrolled in dual language immersion to students in three other language programs. The study found that in terms of SD units of reading achievement scores, ELs in dual language immersion started the lowest among all programs in 2nd grade but experienced the most gains between 2nd and 7th grade.

Taken together, extant research on dual language programs provided evidence on students' achievement status. However, the studies faced key limitations in modeling academic *growth* over time. In most study contexts, the test was not vertically-scaled; as a result, scale scores could not be compared across grade levels. Valentino and Reardon (2015) addressed this by converting spring test scores to z-scores in order to model students' growth relative to grade-level peers. But no study has been able to estimate the association between dual language program participation and within-student growth. This means that when a student moves up in rank, we assume but cannot be sure of the direction or amount of academic growth. In other words, the rank approach simply presents a rank order but lacks the information to describe magnitude or meaning between any two points of interest. When we see a student move up the ranks, her new position does not confirm learning. Maybe her peers grew, and she grew more; or perhaps she stayed the same while her peers lost learning; she might have even lost some learning but moved up the ranks because her peers lost more. As such, analysis of relative status does not capture within-student growth patterns contributing to the achievement gaps.

In addition, the extant research relies on scores from annual standardized testing. As a result, we lack information about within-year growth between fall and spring and about summer learning growth or loss. Variations in patterns of learning by grade level and season are well-documented in recent research (Kuhfeld, 2019; von Hippel & Hamrock, 2019). Learning gains during the school year tend to be followed by stagnation or loss during the summer, and inferences on the effectiveness of teachers and school programs appear to be sensitive to the inclusion of summers (Atteberry & Mangan, 2020). This means that estimating a single effect size for each year masks important differences in seasonal learning rates.

Seasonal learning patterns for students who use another language at home—and participation in dual language programs—merit investigation because of the students' unique home contexts. Language barriers hinder family involvement (Vera et al., 2012). When schools provide academic materials in English for students to practice at home, Spanish-speaking families may not be able to help the students. Dual language programs have the potential to increase family involvement, which creates more out-of-school learning opportunities. For example, if families read to students in Spanish during summer vacation, students would continue to add to their repertoire of literacy skills, preparing them to acquire more advanced concepts in both Spanish and English in the fall. In this case, students would grow or at least maintain their skills during summers. However, if families do not pursue reading and other achievement-enhancing activities on their own regardless of language, then students might flatline or decline over the summer, and having Spanish instruction in school would not make a difference. Research is needed to first identify which pattern holds and then explore how dual language program features, such as teacher-family interactions and summer learning support, shape students' seasonal learning.

In exploring program efficacy, it is important to examine within-year growth to track seasonal learning. If dual language instruction is indeed more effective at closing achievement gaps than other programs, we should see dual language students making more learning gains than other students between the first and last day of each school year. Similarly, we need to interrogate what happens to student learning while they are out of school during summer break. If dual language programs are also more effective than other programs at providing students with extended learning opportunities in the summer, through family engagement or other mechanisms, then we would see dual language students making gains or losing less learning than other students. However, it is also possible that dual language programs are more effective during the year when students are physically in class but are no more (or less) effective than other programs at supporting students while they are out-of-school. In this case, we would want to identify any gaps in summer learning rates so that schools can revise program designs to help students maintain or augment their school-year learning in the summer. This is not possible by only examining spring-to-spring achievement status, as previous research has done.

Current Study

This study addresses these important gaps in the literature by estimating the association between participation in a dual language program and students' achievement growth within and across years. I leverage rich longitudinal data from a diverse district in the Midwest with a long history of serving a large Spanish-speaking student population. A district-wide curriculum change in 2011-12 resulted in the replacement of transitional bilingual education with long-term dual language instruction as the primary support program for Spanish-speaking ELs. Controlling for student- and school-level covariates, I compare achievement and growth for Hispanic students who participated in the long-term dual language program to Hispanic students who did not participate. My research questions are:

1. How does Hispanic dual language participants' achievement status compare to nonparticipants'?
2. How does Hispanic dual language participants' achievement growth within and across grades compare to nonparticipants'?

Study Context

This study is conducted in a research-practice partnership with a large school district in the Midwest. The district enrolls about 40,000 students in total. Slightly more than half the students are Hispanic; another quarter are White; and the rest of the students are members of other racial/ethnic groups. Just under two-thirds of the students come from low-income families. The district has a long history of serving students from diverse linguistic backgrounds, with current ELs comprising over 30% of its total student population. About 80% of the ELs use Spanish at home.

Spanish-English Dual Language Instruction

Spanish-English bilingual education was first implemented in the district in the 1970s and expanded over the years. For decades, a transitional bilingual education model made Spanish instruction available to students in kindergarten through 2nd grade in most elementary schools. Starting in the 2011-12 school year, Spanish-English dual language instruction was expanded across the district, such that any Spanish EL student enrolling in 2nd grade or below in 2011-12, as well as all subsequent cohorts, would be able to receive dual language instruction from elementary to high school. Students who are native and fluent users of English are also eligible to participate. The remainder of this paper will refer to this expanded, long-term Spanish-English dual language instruction program as “DLP.”

In 2011-12, all Spanish-speaking EL students in the district became eligible to participate in a DLP by default, though students’ families have the right to opt out of participation. Since DLP started, it has been the primary support service for Spanish-speaking ELs in the district. Other support services, such as transitional (short-term) bilingual, ESL, and sheltered instruction, remain available to ELs in the district. Spanish-speaking ELs who opt out of DLP can choose to receive or decline these alternative services. More than 85% Spanish-speaking ever-ELs who enrolled in the district in or after 2011-12 enrolled in DLP during at least one school year. Once enrolled in DLP, ELs can (and most do) stay in the program even after gaining fluent English proficiency and reclassifying. In other words, exiting EL status does not lead to removal of services in this context. For all students who choose to remain in DLP, their instructional programs and services do not change as a result of EL reclassification.

Program Features

DLP offers Spanish-English dual language instruction between kindergarten and 12th grade. Kindergarten instruction is delivered 80% in Spanish and 20% in English (80:20), followed by 70:30 and 60:40 in 1st and 2nd grade, respectively, and 50:50 between 3rd and 6th grade. In 7th and 8th grade (middle school), students take language arts in both Spanish and English, social studies in Spanish, and math and science in English. From 9th to 12th grade, students take Spanish language arts and at least one additional core subject course delivered in Spanish every year. In addition, all 9th grade DLP students are required to take the Advanced Placement Spanish Language exam. Spanish-speaking students who are dually eligible for EL and special education (SPED) services are co-taught by dual language and SPED specialists.

Data and Sample

The data for this study come from the district's administrative records between school years 2000-01 and 2018-19. Variables include students' (a) demographic characteristics (e.g., gender, race/ethnicity, home language); (b) eligibility to receive EL, free or reduced-price lunch (FRPL), and SPED services and participation; (c) course enrollment, which identifies DLP participation; and (d) academic assessment scores.

Assessment Scores

. This study focuses on NWEA's MAP Growth assessment. MAP Growth is a computer adaptive test, which provides a precise measurement even for students above or below grade level. The assessments are aligned to the state's content standards. Each test takes approximately 40 to 60 minutes to administer and typically takes place three times per academic year—in the fall, winter, and spring. The assessment is vertically scaled to allow for the estimation of gains across time. Test scores are reported on the RIT scale, where RIT is a linear transformation of the logit scale units of the Rasch item response theory model. In this district, students took the MAP Growth math and reading assessments from the fall of 2nd grade to the winter of 8th grade between 2004-05 and 2018-19. Eighth-grade spring testing is optional in the district, therefore test scores in that term are excluded.

Student and School Covariates

I observe a comprehensive set of student demographic characteristics and school-level student composition variables. I include key student- and school-level covariates in my analyses (described in the Analysis section) because prior research (e.g., Valentino & Reardon, 2015) has shown them to be strongly correlated with student achievement and used them as statistical controls. For the purpose of multilevel growth modeling, student- and school-level covariates are treated as constant across time. Annual support service eligibility flags are used to generate ever-eligible indicators at the student level. Each student thus has an ever-EL, an ever-FRPL, and an ever-SPED variable that has the same value across all time points in the data. I also include an indicator for refusing EL service at any time before the fall of 2^{nd} grade because students whose families refused service may be systematically different from students whose families consistently accepted EL service. Using student-level data, I calculate the percentages of students enrolled at each school who are eligible for EL and FRPL services. Students are nested within the first school attended in the district, and school-level characteristics are treated as constant across time as required by the multilevel model.

Sample

I retain only Hispanic students in the analytic sample since the Spanish-English DLP was designed primarily to serve Spanish-speaking students and the vast majority of participants have been students whose reported race/ethnicity is Hispanic. In the descriptive analyses outlined in the next section, I compare the academic growth of DLP participants (students in post-program cohorts who enrolled) to nonparticipants (a pooled group composed of students in pre-program cohorts who did not have access to DLP *and* students in post-program cohorts who could have enrolled but did not). Nonparticipants in post-DLP cohorts shared time-specific schooling experiences with participants but may have differed in systematic ways. The unobserved characteristics that led some Hispanic students to enroll in DLP and others to opt-out may also be driving differences in achievement growth. Pre-DLP cohorts, on the other hand, may share key characteristics with participants, though their schooling took place years apart. Thus, all cohorts who enrolled in 2^{nd} through 8^{th} grade both before and after DLP was implemented were included to strengthen the comparability of participants and nonparticipants in the main analysis. I also test the sensitivity of my findings to restricting the sample to only students in post-DLP cohorts and report them in the supplemental materials.

I observe achievement at up to 20 time points for each student, for a total of 341,448 math test scores and 341,333 reading test scores nested within 27,661 students within 51 schools. Appendix Table A1 shows the number of students tested by cohort and grade. Since only four post-DLP cohorts have 7th and 8th grade test scores, I interpret findings pertaining to middle school growth as suggestive rather than definitive evidence. Table 1 shows the summary statistics for the pooled sample, students who enrolled in DLP at least one year ("DL"), and students who never enrolled in DLP ("non-DL"). About 71% of the students in the pooled sample were ever eligible for EL services, 91% ever eligible for FRPL services, and 14% ever eligible for SPED services. DL students were had higher rates of ever-EL (94% vs 58%) and ever-FRPL (96% vs 88%) than non-DL students.

[Table 1 here]

Analysis

I examine the math and reading trajectories of DL and non-DL students in three ways. Prior literature primarily relied on z-scores to compare students' spring achievement across grades (e.g., Valentino & Reardon, 2015). Following this line of research, my first approach plots spring achievement z-scores, calculated using MAP Growth norms, for DL and non-DL students between 2nd and 7th grade (Thum & Kuhfeld, 2020). Then, I leverage vertically-scaled test scores from the fall, winter, and spring to interrogate growth within and across grades. First, I plot the mean achievement by DLP participation for each test term between the fall of 2nd grade and the winter of 8th grade. My plots illustrate changes in achievement score not only between but also within grades.

Finally, I use a piecewise multilevel growth model to estimate growth (e.g., von Hippel, Workman, & Downey, 2018). Exposure to school in each grade and each summer varied. Students were not tested on the first and last days of school each year; even within school, students' test dates varied depending on the availability of electronic devices used for testing. An important advantage of the piecewise multilevel model is its ability to account for variation in test administration dates within the school year and allow for separate growth terms in each school year and summer (e.g., Quinn et al., 2016). By separately specifying growth terms for each school year, we can test whether any differences in growth rates between DL and non-DL students expand, stay the same, or diminish across grade levels.

The model accounts for variations in test dates and estimates students' academic growth as a linear function of their "months of exposure" to each school year and summer. Months of exposure is calculated based on school start and end dates and the test administration dates (see Appendix A3 for details). For example, a student testing at the end of August in 3rd grade may have 9.7 months of exposure to 2nd grade, 2.3 months exposure to summer following 2nd grade, and one week of exposure to 3rd grade.

At level 1, I model achievement conditional on exposure to school during the academic year for each grade level (e.g., $G2_{ij}$ = Grade 2 academic year) and exposure to summer after each grade level (e.g., $S2_{ij}$ = summer after Grade 2). Appendix A3 details the calculation of each of the level-1 predictors ($G2_{ij}$ through $G8_{ij}$).

Level 1 (time (*t*) within student (*i*) within school (*j*)):

$$y_{tij} = \pi_{0ij} + \pi_{1ij}G2_{ij} + \pi_{2ij}S2_{ij} + \pi_{3ij}G3_{ij} + \pi_{4ij}S3_{ij} + \pi_{5ij}G4_{ij} + \pi_{6ij}S4_{ij} + \pi_{7ij}G5_{ij} + \pi_{8ij}S5_{ij} + \pi_{9ij}G6_{ij} + \pi_{10ij}S6_{ij} + \pi_{11ij}G7_{ij} + \pi_{12ij}S7_{ij} + \pi_{13ij}G8_{ij} + e_{tij} \quad (1)$$

As von Hippel and colleagues (2018) state, this model "implicitly extrapolates beyond the test dates to the scores that would have been achieved on the first and last day of the school year" (p. 335). The intercept (π_{0ij}) therefore is the predicted score for student *i* in school *j* testing on the first day of 2nd grade, even if the student tested in the third week of the school year. The slopes ($\pi_{1ij}, \ldots, \pi_{13ij}$) are the monthly learning rates of student *i* during each school year and summer. Each test score y_{tij} is viewed as a linear function of the months that student *i* in school *j* has been exposed to 2nd grade ($G2_{ij}$), 3rd grade ($G3_{ij}$), etc., through 8th grade ($G8_{ij}$); and the number of months that the student has been exposed to the summers after 2nd ($S2_{ij}$) through 7th grade ($S7_{ij}$).

At level 2, I include a random intercept to allow students' starting achievement in fall of 2nd grade to vary by student; slopes are treated as fixed. I start with a null model that only includes the random intercept (Model 0). Then, I add the program participation indicator, *DL*, as a predictor, and an interaction term between *DL* and each time variable, which will yield estimates for additional growth DLP students experienced above and beyond the growth rate of non-DLP students (Model 1). After that, I run another model that additionally includes student-level pre-program covariates (Model 2). Equation (2) below shows my preferred specification. *DL* represents ever participation in DLP. *EL*, *FRPL*, and *SPED* are student-level covariates representing ever being eligible for each service during the students' entire time in the district; *female* is an indicator for the student's biological sex reported as female; *refuse* is an indicator for ever refusing EL service before the fall of 2nd grade. For a subsample (about 30% of the students), I also observe Fountas & Pinnell Literacy Gradient (2020) scores measured prior to the fall of 2nd grade. I include this pre-treatment literacy measure as an additional covariate in a sensitivity check.

Level 2 (student (i) within school (j)): (2)

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}DL_{ij} + \beta_{02j}EL_{ij} + \beta_{03j}FRPL_{ij} + \beta_{04j}SPED_{ij}$$

$$+ \beta_{05j}female_{ij} + \beta_{06j}refuse_{ij} + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}DL_{ij}$$

$$\vdots$$

$$\pi_{13ij} = \beta_{130j} + \beta_{131j}DL_{ij}$$

At level 3, a random intercept allows starting achievement in fall of 2nd grade to vary by school; slopes are treated as fixed. Again, I start with a model with only the random intercept. Then, I add *%EL*, and *%FRPL*, which are school-level covariates for the first school in which the student enrolled (Model 3). Equation (3) below shows the preferred specification that includes these covariates.

Level 3 (school (j)): (3)

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(\%EL_j) + \gamma_{002}(\%FRPL_j) + u_{00j}$$

$$\beta_{10j} = \gamma_{100}$$

$$\vdots$$

$$\beta_{130j} = \gamma_{1300}$$

Variance component specification:

$$e_{tij} \sim \mathrm{N}(0, \sigma^2_{tij}),\ \ r_{ij} \sim \mathrm{MVN}(0, T_{St}),\ \ u_j \sim \mathrm{MVN}(0, T_{Sch}).$$

Models are estimated using HLM 8.0 software (Raudenbush et al., 2019). I apply the 3-level model to the full sample of Hispanic students. Then, I restrict the sample to students who have ever been classified as EL (N=19,632) to compare the growth rates of ever-ELs who have and have not participated in DLP. In this EL-only analysis, I drop the student-level predictor *EL* from the model.

Findings

RQ1: How does Hispanic dual language participants' achievement status compare to nonparticipants'?

Figure 1 shows Hispanic students' achievement z-score movement between 2nd and 7th grade for math and reading. In math, DLP students scored in the fall of 2nd grade about .3 SD below the national average and non-DL students scored slightly above DLP students. By the end of 7th grade, the gap between DLP students and the national average widened to about .45 SD, while the gap between non-DLP students and the national average shrank to about .1 SD. This is because the two groups of students had different z-score trajectories. Between the springs of 2nd and 6th grade, DLP students' math z-scores had only small fluctuations but mostly remained flat; then, between the spring of 6th and 7th grade, DLP students dropped in rank to about .45 SD below the national average. Non-DLP students, on the other hand, moved quite a bit toward the national average between 2nd and 3rd grade (by about .1 SD) and between 5th and 6th grade (by about .15 SD), though they fell or stagnated between 3rd and 5th and between 6th and 7th grade.

In reading, the gap between DLP students and the national average was about .65 SD in 2nd grade and .5 SD in 7th grade; between non-DLP students, the gap was just under .4 SD in 2nd grade and just above .2 SD in 7th grade. Both groups closed their respective gaps by about .15 SD in those five years, following trajectories that were mostly parallel.

It is important to note that these movements in z-scores provide information about DLP students' academic positions relative to non-DLP students and the national average but cannot speak to the progress students themselves are making within and across years.

[Figure 1 here]

RQ2: How does Hispanic dual language participants' achievement growth within and across grades compare to nonparticipants'?

Achievement Means

Figure 2 shows growth in math and reading achievement between the fall of 2nd grade and winter of 8th grade. The horizontal axis shows the 20 test terms, with labels and national average markers for the winter test term of each year (Thum & Kuhfeld, 2020), which the district uses for school accountability.

In math, non-DLP students scored at or slightly below the national average each winter, while DLP students scored slightly below non-DLP students. DLP students started the fall of 2nd grade scoring slightly lower than non-DLP students. Between the fall and spring of each grade between 2 and 5, DLP students grew more, closing the gap between themselves and non-DLP peers; however, DLP students lost more learning over the summers in these years (shown by the larger dips in the red line) than non-DLP students, so that the gap is somewhat restored by fall. In grades 6 to 8, the growth trends for DLP and non-DLP students are mostly parallel, suggesting similar academic-year growth and summer loss rates.

In reading, both non-DLP and DLP students scored consistently below the national average in each winter; however, the gaps between both groups and the national average shrank as students moved to higher grades. Similar to math, DLP students started the fall of 2nd grade scoring below non-DLP students in reading. In each subsequent academic year and summer, the growth trajectories of DLP and non-DLP students were parallel, indicating that the two groups had similar academic-year growth and summer loss rates. It is also noteworthy that neither DLP nor non-DLP students seemed to lose much learning during each summer, as suggested by the flat segments in the graph.

[Figure 2 here]

Monthly Growth Estimates

Monthly growth rates for DLP students, non-DLP students, and the national average (calculated using MAP Growth norms from Thum & Kuhfeld, 2020) in each academic year and subsequent summer are presented in Figure 3. Estimates are from the preferred model specification which includes student and school covariates (Table 2, Columns 3 and 6). The black markers at the tip of each bar represents 95% confidence intervals for the estimates.

[Figure 3 here]

In math, between grades 2 and 5, DLP students grew significantly more than non-DLP students during each academic year; but they also had greater monthly loss during each summer. For instance, in the 2nd grade year, DLP students grew about 2 RIT per month, while non-DLP students grew 1.7 RIT per month. Over the school year, 2nd grade DLP students would close the gap between themselves and their non-DLP peers by 2.7 RIT. In the subsequent summer, DLP students lost more per month, almost exactly undoing the amount they caught up over the year. In 6th grade, there was no difference between the two groups in terms of growth or summer loss. From 7th to 8th grade, the pattern reverses: DLP students grew significantly slower than non-DLP students but also experienced significantly less monthly during the summer after 7th grade, though the differences were small in practical terms (e.g., about 0.1 RIT per month in 7th grade year). Across all years, both DLP and non-DLP students in the district had similar or lower summer loss rates compared to the national average.

The results for reading are quite different. DLP students grew slightly faster (by about an additional .1 RIT per month) during 2nd grade but lost significantly more (by about an additional .2 RIT per month) in the subsequent summer. From 3rd to 6th grade, DLP students tended to grow slower during the year but lose less over the summer compared to non-DLP students. In 7th and 8th grade, growth rates did not differ significantly between the two groups. With the exception of 2nd and 8th grade, students in the district tended to grow slightly slower than the national average. At the same time, summer loss rates in the district were significantly and substantively lower than the national average in grades 2 to 5.

Regression results corresponding to Figure 3 are shown in Table 2. Null models (Appendix Table A2) yield student intraclass correlations of .51 and .53 for math and reading, and the school intraclass correlation is about .07 for both subjects. This suggests that between-student variations explain about half the total variation in student achievement at the start of 2nd grade, and between-school variations explain about 7% of the total variation. I therefore add student and school covariates to account for these differences. Columns (1) (math) and (4) (reading) show results from a model that includes ever DLP participation as a predictor and interaction terms between DLP and each time predictor (exposure to 2nd grade academic year through exposure to 8th grade academic year). The intercept represents student achievement on the first day of 2nd grade. The average achievement of non-DL students was 171.3 RIT for math and 163.9 for reading. In comparison, DL students scored a significant 3.07 RIT and 4.57 RIT lower in math and reading, respectively. Columns (2) and (5) show results from adding sex and educational service eligibility and participation. The intercepts (179.2 and 172.4) represent starting RIT for male, non-DL students who were never eligible to receive EL, SPED, or lunch services. Being female and being eligible to receive EL, SPED, and lunch services are each associated with significantly lower starting achievement in math and reading. Holding all other factors constant, having refused EL service in kindergarten or 1st grade is associated with lower achievement in math but not in reading. When school-level percentage of EL and FRPL enrollment are added to the model (columns 3 and 6), higher percentages of FRPL eligibility is associated with lower starting achievement in math and reading; but the percentage of EL eligibility is positively associated with achievement, significant only for reading. Growth estimates in all academic years and summers are stable in both magnitude and statistical significance across the three models for both math and reading. For instance, non-DL students grew 1.68 RIT during 2nd grade. As represented by the coefficient for the interaction term, DL students grew an additional .29 RIT, which is statistically significant. Results from columns (3) and (6) correspond to those reported in Figure 3.

[Table 2 here]

EL Growth Estimates

Table 3 shows the findings from multilevel growth models applied to only Hispanic ever-ELs. Again, estimates are stable across models in both magnitude and significance. I interpret results from columns (3) and (6), which are the preferred specification that includes student and school covariates. EL growth estimates for both math and reading are very similar to the full sample. In math, ELs in DLP grew significantly more than ELs never in DLP during each academic year in grades 2 to 5, but ELs in DLP also lost more learning during summers after grades 2 to 5 compared to non-DL ELs. The pattern reverses in 7th and 8th grade. In reading, ELs in DLP tended to grow less than ELs never in DLP during each grade between 2 and 7, but DLP ELs also lost less learning during summers.

[Table 3 here]

Sensitivity Checks

The series of models I estimated, including and excluding student- and school-level covariates as controls, yielded very similar results. Appendix A4-A6 reports findings from two sensitivity checks. I ran the same piecewise 3-level models to estimate month growth rates for Hispanic DLP and non-DLP students who attended school after DLP was implemented and took MAP Growth assessments between 2011-12 and 2018-19 ("post-DLP cohorts"). Results are very similar to those for the full sample including pre-DLP cohorts (Figure A4 and Table A5). For a subsample (about 30% of the full sample) for whom I observe Fountas and Pinnell literacy assessment data, I repeated the analysis including their baseline literacy score measured prior to the fall of 2nd grade; estimates similar to the full sample (Table A6).

Discussion

This study leverages rich longitudinal data to estimate growth within and across grade levels for Hispanic participants and nonparticipants of a long-term DLP. Achievement z-score movements between grades suggest that in reading, both DLP and non-DLP students moved closer to the national average; in math, however, non-DLP students closed most of the math achievement gap between themselves and the national average while the gap between DLP students and the national average widened.

At first glance, my findings on students' achievement relative grade-level peers to seem to contradict some of the results from earlier studies. For instance, Valentino and Reardon (2015) found that longer-term dual language ELs outperformed ELs in English-only and short-term transitional or developmental bilingual programs in 5th grade and reached the state average in 6th grade in ELA achievement rank. Similarly, Steele et al. (2017) found that dual language immersion students outperformed their peers on reading at the end of 5th grade. In my data, however, DLP students did not completely close the reading achievement gap between themselves and their peers (either non-DLP students in the same district or the national average) even at the end of 7th grade. In math, Valentino and Reardon (2015) and Steele et al. (2017) found that dual language students did better or at least no worse than students in other programs in terms of achievement relative to their state average across grades; but in my data, DLP students did not get closer to the national average as non-DLP students did.

A closer look at seasonal learning explains how the achievement z-score patterns in my data developed. DLP and non-DLP students exhibited different growth patterns during academic years and summers breaks, at least between the start of 2nd and start of 6th grade. In math, both DLP and non-DLP students grew faster compared to the national average, and DLP students grew faster than non-DLP students. In the summers, however, non-DLP students generally lost learning at a rate similar to the national average, while DLP students lost learning at a significantly higher rate. The difference in summer loss rates would explain why DLP students were not approaching the national average across grades as the non-DLP students were. In reading, both DLP and non-DLP students grew slightly less than the national average during the academic years; however, both groups (DLP students in particular) experienced significantly less summer loss than the national average, which would explain their approaching the national achievement average across grades.

The analysis of seasonal learning patterns in this study contributes significantly to the extant literature on dual language education. Prior research tended to find participation in dual language instruction to have positive associations with (or impact on) downstream achievement relative to their state average or improvements in z-scores across grades, but none was able to decompose these general patterns. This study shows that academic-year growth and summer maintenance are two key issues that merit consideration by my partner district and dual language programs across the nation. In designing programs to support vulnerable student populations, policymakers and educators should consider ways to maximize learning not only within academic years but across the entire time span between the fall of kindergarten to the spring of 12^{th} grade. Summers can and should be leveraged to reinforce learning gained over the previous year.

In my research partner district, Hispanic DLP students experienced little summer loss in reading in each of the early grades. The districts might reflect on its successful practices that led to minimal reading summer loss and share with other programs. In math, however, DLP students, who grew more during the school year, experienced significantly larger summer loss in math compared to Hispanic non-DLP students across 2^{nd} to 5^{th} grade. This is consistent with patterns that previous research has observed for students across the nation: students who grow more during the school year compared to their peers also generally tend to lose more over the summer (Kuhfeld, 2019). The district aims to break from this national trend and is considering aspects of DLP that could be improved to help students sustain their learning gains. Many current district efforts focus on supporting students' summer reading, and they seem to be paying off, as reflected by the very low reading summer loss rates. In contrast, no organized curriculum or materials are provided for summer math learning. A viable first step would be to make bilingual print or online materials available for students to work on at home or organizing regular review sessions with bilingual support from adults in the community. Upon seeing these findings, the district's EL and math departments have initiated discussions around potential approaches to address math summer loss.

Limitations

The findings of this study should be interpreted in light of a few limitations. First, the data for this study come from a single district with a long history of serving a large Hispanic student population, the majority of whom were eligible for EL service at some point. Its Spanish-English DLP is unique in many ways, including content instruction in Spanish across multiple subjects through middle school and into high school. Due to the uniqueness of the programs and the study context, my findings may not generalize well to other districts with different student demographics or different approaches to program design or implementation.

Second, the results of this study are descriptive and do not support causal inference. Some factors do alleviate concerns for selection bias to a certain extent (e.g., program eligibility applied to entire cohorts of students, and among Hispanic EL students in post-program cohorts participation was very high). Still, the differences between participants and nonparticipants of dual language instruction should not be construed as causal estimates of program impact.

Third, this study was not able to identify the specific mechanisms driving the differences between DLP and non-DLP students. For instance, the factors contributing to larger school-year gains and larger summer losses for math remain unclear. More research is needed to explain these differences in growth rates and to identify effective strategies for addressing the disparities.

Finally, this study focuses on math and English reading achievement and does not look at other key outcomes such as Spanish proficiency and literacy, EL reclassification, or social emotional development. These outcomes are also crucial to students' academic and overall wellbeing. The relation between DLP and Spanish proficiency and literacy especially merit rigorous research, since bilingualism and biliteracy—potential benefits unique to long-term dual language instruction and unfound in other EL service programs—are at the center of DLP's mission and purpose. In follow-up studies, the research partnership plans to examine additional outcomes, including Spanish reading achievement and growth across grades and participation rates and outcomes on 9th grade students' Advanced Placement Spanish Language exams.

Conclusion

Enabled by a research-practice partnership with a large school district, this study compares the academic growth of Hispanic students who did and did not participate in Spanish-English dual language instruction. Prior research was only able to compare achievement status or explore changes over time in achievement rank relative to other students. Building on this body of research, I show that seasonal patterns of learning differ between participants and nonparticipants of dual language instruction. As such, this study is the first to demonstrates *how* achievement gaps between program participants and nonparticipants develop within and across grades.

As dual language programs continue to develop across the nation, program implementation and evaluation should consider academic growth, in addition to static achievement, as a key outcome. In order to maximize learning throughout the student's academic trajectory, progress needs to be monitored through regular formal or informal assessments both within and across grades. This requires collaboration between teachers across subjects and grade levels, as well as unwavering support from school- and district-level leadership. Students' families and adults in the community are also indispensable resources for maintaining learning and growth during out-of-school time, such as summer breaks.

In the spirit of Castaneda v. Pickard (1981), as with all other programs that serve ELs and other vulnerable student populations, dual language programs should be evaluated, formally or informally, on a regular basis. Program implementation needs to be interrogated for equity in opportunities to learn and key outcomes like academic growth. As Valdés (2018) pointed out, dual language educators are charged with the critical task of managing the resources that diverse groups of students bring from their homes and communities. To continuously evaluate and improve dual language programs in an iterative fashion, experts on literacy, content areas, language acquisition, and data analysis all need to work together. This research partnership is an example of a collaborative effort to study and improve dual language education. The partnership is extending this study by examining the relations between dual language program participation and other student outcomes, such as high school course-taking. The great strength of research-practice partnerships is its ability to inform policy and instruction. I hope this study will catalyze research collaborations for improving learning opportunities and outcomes for diverse student populations.

References

Atteberry, A., & Mangan, D. (2020). The sensitivity of teacher value-added scores to the use of fall or spring test scores. *Educational Researcher*, *49*(5), 335-349. https://doi.org/10.3102/0013189X20922993

Bialystok, E. (2018). Bilingual education for young children: review of the effects and consequences, *International Journal of Bilingual Education and Bilingualism, 21*(6), 666-679. DOI: 10.1080/13670050.2016.1203859

Castañeda v. Pickard, 648 F.2d 989. 5th Cir. (1981).

Collier, V. P. & Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale, longitudinal research. *Annual Review of Applied Linguistics (ARAL)* 2017.

Dual Language Schools. (2020). https://duallanguageschools.org/

Fountas and Pinnell. (2020). Literacy Gradient. Retrieved from https://www.fountasandpinnell.com/assessment/

Goldenberg, C. (2008). Teaching English Language Learners: What the research does—and does not—say. *American Educator*, *32*(2), 8-44.

Howard, E. R., Lindholm-Leary, K. J., Rogers, D., Olague, N., Medina, J., Kennedy, B., Sugarman, J., & Christian, D. (2018). *Guiding Principles for Dual Language Education* (3rd ed.). Washington, DC: Center for Applied Linguistics.

Kraft, M. A., & Dougherty, S. M. (2013). The effect of teacher-family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness*, *6*(3), 199-222. https://doi.org/10.1080/19345747.2012.743636.

Kraft, M. A., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review*, *47*, 49-63. https://doi.org/10.1016/j.econedurev.2015.04.001

Kuhfeld, M. (2019). Surprising new evidence on summer learning loss. *Phi Delta Kappan, 101* (1), 25-29.

Lindholm-Leary, K. & Block, N. (2010). Achievement in predominantly low SES/Hispanic dual language schools. *International Journal of Bilingual Education and Bilingualism*, *13*(1), 43-60. DOI: 10.1080/13670050902777546

Lindholm-Leary, K. & Hernández, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: native English speakers, fluent English/previous ELLs, and current ELLs. *Journal of Multilingual and Multicultural Development*, *32*(6), 531-545. DOI: 10.1080/01434632.2011.611596

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher, 45,* 443–453.

Raudenbush, S.W., Bryk, A.S, Cheong, Y.F. & Congdon, R. (2019). HLM 8 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion programs on student achievement: Evidence from lottery data. *American Educational Research Journal*, *54*(1S), 282S-306S.

Thomas, W. P. & Collier, V. P. (2012). *Dual language education for a transformed world*. Albuquerque: Fuerte Press.

Thum, Y. M. & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth norms for students and school achievement status and growth*. NWEA Research Report. Portland, OR: NWEA.

Umansky, I. (2016). "To Be or Not to Be EL: An Examination of the Impact of Classifying Students as English Learners." *Educational Evaluation and Policy Analysis,* 38(4), 714-737.

Valdés, G. (2017). AERA and the study of non-English languages in American education. *American Educational Research Journal, 54*(1S), 307S-301S.

Valdés, G. (2018). Analyzing the curricularization of language in twoway immersion education: Restating two cautionary notes. *Bilingual Research Journal, 41*(4), 388-412. DOI: 10.1080/15235882.2018.1539886

Valentino, R. A. & Reardon, S. F. (2015). Effectiveness of four instructional programs designed to serve English Learners: Variation by ethnicity and initial English proficiency. *Educational Evaluation and Policy Analysis, 37*(4), 612-637.

Vera, E. M., Israel, M. S., Coyle, L., Cross, J., Knight-Lynn, L., Moallem, I., Bartucci, G., & Goldberger, N. (2012). Exploring the educational involvement of parents of English Learners. *School Community Journal, 22*(2), 183-202.

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6,* 43–80.

von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of “Are schools the great equalizer?” *Sociology of Education, 91*(4), 323-357.

Table 1. Sample Characteristics

	DL (N=9,980)	Non-DL (N=17,681)	Total (N=27,661)
Female	0.49	0.49	0.49
Ever-EL	0.94	0.58	0.71
Ever-FRPL	0.96	0.88	0.91
Ever-SPED	0.14	0.13	0.14
Refused EL Service K or G1	0.05	0.14	0.11

Notes: Sample includes students whose reported race/ethnicity was Hispanic and who took MAP Growth math or reading assessments between 2004-05 and 2018-19. DLP represents students who ever enrolled in the dual language program. Non-DL represents students who never enrolled in the dual language program. Ever-EL indicates having ever been eligible for EL services while enrolled in the district. Ever-FRPL indicates having ever been eligible for free or reduced-price lunch services. Ever-SPED indicates having ever been eligible for Special Education services. Refused ELL Service K or G1 indicates having been eligible for but refused to receive EL service during kindergarten or 1st grade.

Table 2. Estimated Monthly Growth Rates for Full Sample

	Math			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Starting RIT	171.337***	179.162***	179.147***	163.855***	172.433***	172.418***
	(0.556)	(0.587)	(0.634)	(0.835)	(0.798)	(0.718)
Gr 2 AY	1.683***	1.683***	1.683***	1.827***	1.826***	1.826***
	(0.028)	(0.028)	(0.028)	(0.030)	(0.030)	(0.030)
DL x Gr 2 AY	**0.287*** **	**0.287*** **	**0.287*** **	**0.100** **	**0.100** **	**0.100** **
	(0.059)	(0.058)	(0.058)	(0.048)	(0.048)	(0.048)
Gr 2 Summer	-0.627***	-0.624***	-0.624***	-0.275***	-0.269***	-0.269***
	(0.069)	(0.069)	(0.069)	(0.062)	(0.062)	(0.062)
DL x Gr 2 Summer	**-0.978*** **	**-0.979*** **	**-0.979*** **	**-0.201** **	**-0.204** **	**-0.204** **
	(0.118)	(0.117)	(0.117)	(0.098)	(0.097)	(0.097)
Gr 3 AY	1.524***	1.523***	1.523***	1.368***	1.367***	1.367***
	(0.020)	(0.020)	(0.020)	(0.026)	(0.026)	(0.026)
DL x Gr 3 AY	**0.095*** **	**0.095*** **	**0.095*** **	**-0.163*** **	**-0.161*** **	**-0.161*** **
	(0.032)	(0.032)	(0.032)	(0.034)	(0.035)	(0.035)
Gr 3 Summer	-0.942***	-0.940***	-0.940***	-0.374***	-0.371***	-0.371***
	(0.048)	(0.048)	(0.048)	(0.057)	(0.057)	(0.057)
DL x Gr 3 Summer	**-0.462*** **	**-0.464*** **	**-0.464*** **	**0.302*** **	**0.296*** **	**0.296*** **
	(0.080)	(0.080)	(0.080)	(0.092)	(0.092)	(0.092)
Gr 4 AY	1.219***	1.218***	1.218***	0.964***	0.963***	0.963***
	(0.020)	(0.020)	(0.020)	(0.018)	(0.018)	(0.018)
DL x Gr 4 AY	**0.165*** **	**0.166*** **	**0.166*** **	**-0.041**	**-0.039**	**-0.039**
	(0.035)	(0.035)	(0.035)	(0.032)	(0.032)	(0.032)
Gr 4 Summer	-0.807***	-0.807***	-0.807***	-0.256***	-0.255***	-0.255***
	(0.054)	(0.054)	(0.054)	(0.052)	(0.052)	(0.052)
DL x Gr 4 Summer	**-0.590*** **	**-0.592*** **	**-0.592*** **	**0.196** **	**0.191** **	**0.191** **
	(0.102)	(0.102)	(0.102)	(0.076)	(0.077)	(0.077)
Gr 5 AY	1.066***	1.066***	1.066***	0.766***	0.766***	0.766***
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
DL x Gr 5 AY	**0.105*** **	**0.105*** **	**0.105*** **	**-0.065** **	**-0.064** **	**-0.064** **
	(0.033)	(0.033)	(0.033)	(0.027)	(0.027)	(0.027)
Gr 5 Summer	-0.767***	-0.769***	-0.769***	0.000	-0.003	-0.003
	(0.045)	(0.044)	(0.044)	(0.037)	(0.037)	(0.037)
DL x Gr 5 Summer	**-1.272*** **	**-1.271*** **	**-1.271*** **	**0.205*** **	**0.204*** **	**0.204*** **
	(0.099)	(0.099)	(0.099)	(0.073)	(0.074)	(0.074)
Gr 6 AY	0.888***	0.889***	0.889***	0.614***	0.615***	0.615***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
DL x Gr 6 AY	**0.024**	**0.023**	**0.023**	**-0.083*** **	**-0.082*** **	**-0.082*** **
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Gr 6 Summer	-1.254***	-1.254***	-1.254***	-0.664***	-0.665***	-0.665***
	(0.040)	(0.040)	(0.040)	(0.044)	(0.044)	(0.044)
DL x Gr 6 Summer	**-0.016**	**-0.015**	**-0.015**	**0.299*** **	**0.298*** **	**0.298*** **
	(0.079)	(0.079)	(0.079)	(0.082)	(0.082)	(0.082)
Gr 7 AY	0.678***	0.678***	0.678***	0.450***	0.450***	0.451***
	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)	(0.015)
DL x Gr 7 AY	**-0.119*** **	**-0.119*** **	**-0.119*** **	**-0.048**	**-0.048**	**-0.048**
	(0.019)	(0.019)	(0.019)	(0.030)	(0.030)	(0.030)

Gr 7 Summer	-0.678***	-0.677***	-0.677***	-0.073	-0.071	-0.071
	(0.034)	(0.034)	(0.034)	(0.054)	(0.054)	(0.054)
DL x Gr 7 Summer	**0.406*** **	**0.404*** **	**0.404*** **	**0.041**	**0.038**	**0.038**
	(0.064)	(0.063)	(0.063)	(0.081)	(0.081)	(0.081)
Gr 8 AY	0.832***	0.832***	0.832***	0.662***	0.661***	0.661***
	(0.026)	(0.026)	(0.026)	(0.036)	(0.036)	(0.036)
DL x Gr 8 AY	**-0.208*** **	**-0.207*** **	**-0.207*** **	**0.083**	**0.084**	**0.084**
	(0.049)	(0.049)	(0.049)	(0.087)	(0.086)	(0.086)
Ever-DL	-3.066***	-1.592***	-1.585***	-4.573***	-1.827***	-1.821***
	(0.411)	(0.386)	(0.387)	(0.453)	(0.352)	(0.352)
Female		-1.534***	-1.533***		1.212***	1.213***
		(0.134)	(0.134)		(0.142)	(0.142)
Ever-FRPL		-3.918***	-3.870***		-4.276***	-4.236***
		(0.328)	(0.328)		(0.363)	(0.366)
Ever-EL		-4.408***	-4.386***		-7.859***	-7.841***
		(0.358)	(0.356)		(0.446)	(0.445)
Ever-SPED		-10.533***	-10.535***		-12.310***	-12.311***
		(0.325)	(0.326)		(0.312)	(0.313)
Refused EL Service		-1.226***	-1.234***		0.402	0.395
		(0.395)	(0.396)		(0.456)	(0.456)
School % FRPL			-6.601***			-8.706***
			(2.535)			(2.182)
School % EL			3.508			7.320**
			(4.634)			(3.397)
Student Covariates	no	yes	yes	no	yes	yes
School Covariates	no	no	yes	no	no	yes
Test Scores	341448	341448	341448	341333	341333	341333
Students	27661	27661	27661	27657	27657	27657
Schools	51	51	51	51	51	51

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes Hispanic students who took MAP Growth assessments between 2004-05 and 2018-19. Each column is a separate 3-level piecewise growth model regression with RIT score as the dependent variable. Starting RIT is the intercept and represents test scores on the first day of 2nd grade. Ever-DL is an indicator for ever enrolling in the dual language program. Female, Ever-FRPL, Ever-EL, Ever-SPED, and refused EL service are student-level covariates. Refused EL service represents having refused EL service in either kindergarten or 1st grade. School % FRPL and school % EL are school-level covariates. Interaction effects representing differential growth for dual language participants are bolded.

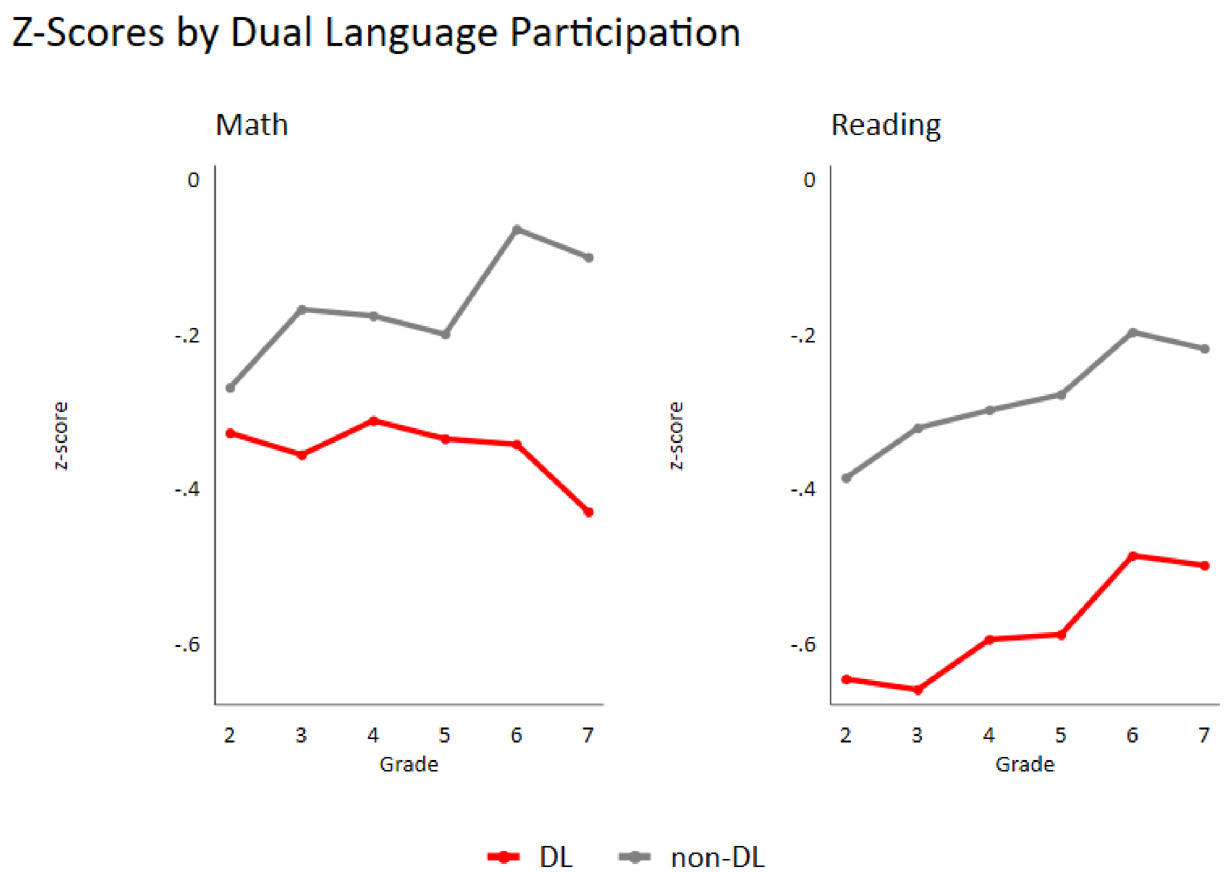
Table 3. Estimated Monthly Growth Rates for Ever-ELs

	Math			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Starting RIT	168.163***	174.137***	174.161***	157.927***	163.647***	163.673***
	(0.813)	(0.875)	(1.326)	(1.099)	(1.281)	(1.159)
Gr 2 AY	1.665***	1.663***	1.663***	1.852***	1.849***	1.849***
	(0.037)	(0.037)	(0.037)	(0.045)	(0.045)	(0.045)
DL x Gr 2 AY	**0.310*** **	**0.310*** **	**0.310*** **	**0.065**	**0.067**	**0.067**
	(0.065)	(0.065)	(0.065)	(0.059)	(0.059)	(0.059)
Gr 2 Summer	-0.688***	-0.688***	-0.688***	-0.275***	-0.272***	-0.272***
	(0.086)	(0.085)	(0.085)	(0.078)	(0.078)	(0.078)
DL x Gr 2 Summer	**-0.977*** **	**-0.975*** **	**-0.975*** **	**-0.210** **	**-0.211** **	**-0.211** **
	(0.125)	(0.124)	(0.124)	(0.102)	(0.101)	(0.101)
Gr 3 AY	1.542***	1.541***	1.541***	1.441***	1.440***	1.440***
	(0.023)	(0.023)	(0.023)	(0.033)	(0.033)	(0.033)
DL x Gr 3 AY	**0.093** **	**0.094** **	**0.094** **	**-0.227*** **	**-0.225*** **	**-0.225*** **
	(0.037)	(0.037)	(0.037)	(0.041)	(0.041)	(0.041)
Gr 3 Summer	-1.058***	-1.057***	-1.057***	-0.495***	-0.493***	-0.493***
	(0.057)	(0.057)	(0.057)	(0.068)	(0.068)	(0.068)
DL x Gr 3 Summer	**-0.399*** **	**-0.400*** **	**-0.400*** **	**0.394*** **	**0.390*** **	**0.391*** **
	(0.088)	(0.088)	(0.088)	(0.099)	(0.099)	(0.099)
Gr 4 AY	1.255***	1.255***	1.255***	1.023***	1.023***	1.023***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
DL x Gr 4 AY	**0.145*** **	**0.145*** **	**0.145*** **	**-0.084** **	**-0.084** **	**-0.084** **
	(0.037)	(0.037)	(0.037)	(0.035)	(0.035)	(0.035)
Gr 4 Summer	-0.919***	-0.919***	-0.919***	-0.317***	-0.317***	-0.317***
	(0.059)	(0.059)	(0.059)	(0.062)	(0.062)	(0.062)
DL x Gr 4 Summer	**-0.521*** **	**-0.521*** **	**-0.521*** **	**0.223*** **	**0.221** **	**0.221** **
	(0.107)	(0.107)	(0.107)	(0.086)	(0.087)	(0.087)
Gr 5 AY	1.095***	1.095***	1.095***	0.816***	0.816***	0.816***
	(0.020)	(0.020)	(0.020)	(0.018)	(0.018)	(0.018)
DL x Gr 5 AY	**0.075** **	**0.075** **	**0.075** **	**-0.103*** **	**-0.103*** **	**-0.103*** **
	(0.035)	(0.035)	(0.035)	(0.028)	(0.028)	(0.028)
Gr 5 Summer	-0.789***	-0.790***	-0.790***	-0.076*	-0.078*	-0.078*
	(0.054)	(0.053)	(0.053)	(0.040)	(0.041)	(0.041)
DL x Gr 5 Summer	**-1.249*** **	**-1.249*** **	**-1.249*** **	**0.260*** **	**0.259*** **	**0.259*** **
	(0.107)	(0.107)	(0.107)	(0.079)	(0.079)	(0.079)
Gr 6 AY	0.881***	0.882***	0.882***	0.650***	0.651***	0.651***
	(0.015)	(0.015)	(0.015)	(0.018)	(0.018)	(0.018)
DL x Gr 6 AY	**0.035**	**0.034**	**0.034**	**-0.113*** **	**-0.113*** **	**-0.113*** **
	(0.022)	(0.022)	(0.022)	(0.019)	(0.019)	(0.019)
Gr 6 Summer	-1.215***	-1.216***	-1.216***	-0.725***	-0.727***	-0.727***
	(0.045)	(0.045)	(0.045)	(0.055)	(0.055)	(0.055)
DL x Gr 6 Summer	**-0.086**	**-0.085**	**-0.085**	**0.339*** **	**0.339*** **	**0.339*** **
	(0.084)	(0.084)	(0.084)	(0.075)	(0.075)	(0.075)
Gr 7 AY	0.661***	0.662***	0.662***	0.486***	0.486***	0.486***
	(0.014)	(0.014)	(0.014)	(0.019)	(0.019)	(0.019)
DL x Gr 7 AY	**-0.114*** **	**-0.114*** **	**-0.114*** **	**-0.074** **	**-0.074** **	**-0.074** **
	(0.019)	(0.019)	(0.019)	(0.034)	(0.034)	(0.034)

Gr 7 Summer	-0.687***	-0.686***	-0.686***	-0.140**	-0.139**	-0.139**
	(0.041)	(0.041)	(0.041)	(0.061)	(0.061)	(0.061)
DL x Gr 7 Summer	**0.435*** **	**0.434*** **	**0.434*** **	**0.114**	**0.112**	**0.112**
	(0.071)	(0.070)	(0.070)	(0.082)	(0.082)	(0.082)
Gr 8 AY	0.839***	0.838***	0.838***	0.678***	0.677***	0.677***
	(0.032)	(0.032)	(0.032)	(0.047)	(0.047)	(0.047)
DL x Gr 8 AY	**-0.219*** **	**-0.218*** **	**-0.218*** **	**0.050**	**0.051**	**0.051**
	(0.057)	(0.057)	(0.057)	(0.096)	(0.096)	(0.096)
Ever-DL	-1.980***	-2.213***	-2.208***	-1.448***	-1.402***	-1.397***
	(0.408)	(0.400)	(0.399)	(0.393)	(0.381)	(0.380)
Female		-1.561***	-1.560***		1.034***	1.034***
		(0.139)	(0.139)		(0.162)	(0.162)
Ever-FRPL		-3.529***	-3.491***		-4.662***	-4.625***
		(0.498)	(0.500)		(0.575)	(0.577)
Ever-SPED		-10.581***	-10.581***		-12.275***	-12.275***
		(0.427)	(0.428)		(0.404)	(0.404)
Refused EL Service		-1.406***	-1.417***		0.052	0.042
		(0.412)	(0.419)		(0.472)	(0.473)
School % FRPL			-7.263			-14.223***
			(8.568)			(3.896)
School % EL			10.026			17.605***
			(15.507)			(6.380)
Student Covariates	no	yes	yes	no	yes	yes
School Covariates	no	no	yes	no	no	yes
Test Scores	247202	247202	247202	246667	246667	246667
Students	19632	19632	19632	19627	19627	19627
Schools	51	51	51	51	51	51

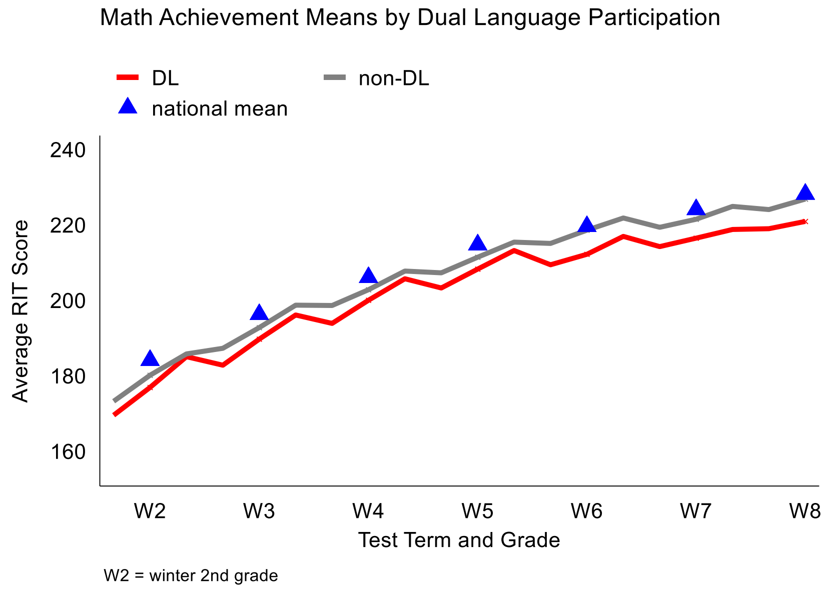
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes Hispanic students who were ever eligible for EL service and took MAP Growth assessments between 2004-05 and 2018-19. Each column is a separate 3-level piecewise growth model regression with RIT score as the dependent variable. Starting RIT is the intercept and represents test scores on the first day of 2nd grade. Ever-DL is an indicator for ever enrolling in the dual language program. Female, Ever-FRPL, Ever-SPED, and refused EL service are student-level covariates. Refused EL service represents having refused EL service in either kindergarten or 1st grade. School % FRPL and school % EL are school-level covariates. Interaction effects representing differential growth for dual language participants are bolded.

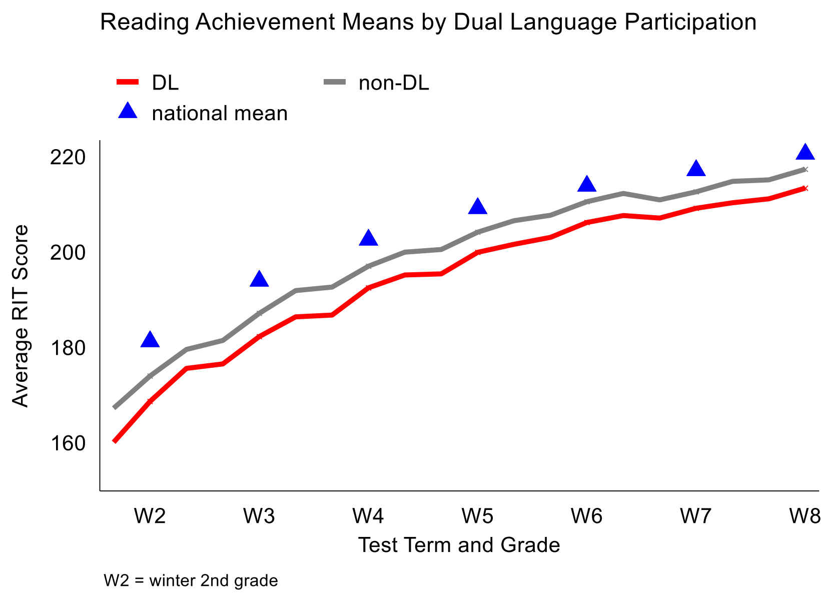
Figure 1. Achievement Z-Score Movements (Spring-to-Spring) Across Grades



Notes: DL represents students who ever enrolled in the dual language program. Non-DL represents students who never enrolled in the dual language program. Z-scores are calculated using spring test scores and NWEA MAP Growth Norms 2020 for each grade and subject.

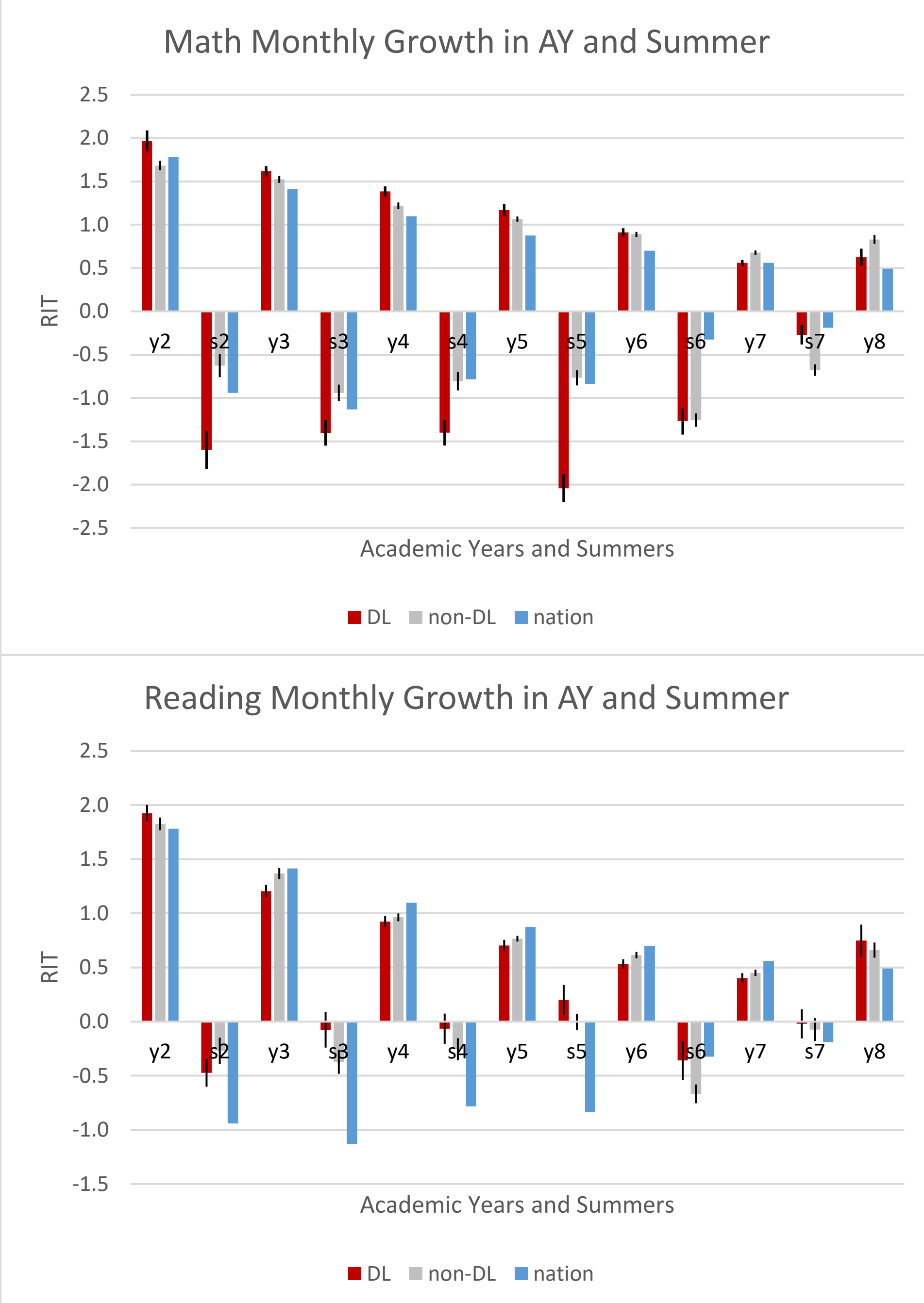
Figure 2. Math and Reading Mean Achievement Scores





Notes: DL represents students who ever enrolled in the dual language program. Non-DL represents students who never enrolled in the dual language program. Trends represent unconditional means for each group of students at each test grade and term.

Figure 3. Estimated Monthly Growth Rates in School Year and Loss Rates in Summer



Notes: DL represents students who ever enrolled in the dual language program. Non-DL represents students who never enrolled in the dual language program. Monthly growth rates are estimated using 3-level piecewise growth models and correspond to Model (3) in Table 2. Sample includes Hispanic students who took MAP Growth assessments between 2004-05 and 2018-19.

Appendix Table A1. Number of Students by 1st Grade Cohort and Grade Tested

	Grades Tested							
1st Grade Cohort (Year of Fall Term)	2	3	4	5	6	7	8	
2003	122	674	1,114	1,363	1,414	1,407	1,390	
2004	668	1,286	1,500	1,582	1,580	1,565	1,543	
2005	1,416	1,765	1,826	1,837	1,848	1,862	1,794	
2006	1,474	1,547	1,575	1,562	1,570	1,552	1,501	
2007	1,556	1,577	1,570	1,565	1,552	1,534	1,514	
2008	1,487	1,522	1,518	1,495	1,479	1,444	1,476	
2009	1,598	1,608	1,605	1,588	1,551	1,567	1,534	Non-DL Cohorts
2010	1,676	1,684	1,649	1,653	1,685	1,680	1,660	DL Cohorts
2011	1,642	1,616	1,591	1,628	1,643	1,638	1,611	
2012	1,636	1,619	1,649	1,665	1,678	1,670	1,572	
2013	1,597	1,636	1,633	1,629	1,635	1,513		
2014	1,492	1,509	1,521	1,527	1,432			
2015	1,600	1,627	1,605	1,064				
2016	1,434	1,416	1,008					
2017	1,515	1,047						

Table A2. Null Model Estimates

	Math		Reading	
	(1)	(2)	(3)	(4)
	full sample	Ever-ELs only	full sample	Ever-ELs only
Starting RIT	205.262***	200.500***	197.367***	191.271***
	(0.839)	(0.578)	(0.830)	(0.397)
Test Scores	341448	247202	341333	246667
Students	27661	19632	27657	19627
Schools	51	51	51	51
Level-1 RE SD	15.38	15.45	15.22	15.64
Level-2 RE SD	14.58	14.01	15.38	14.73
Level-3 RE SD	6.08	3.44	5.83	2.440
Level-2 ICC	0.513	0.466	0.539	0.477
Level-3 ICC	0.076	0.027	0.068	0.013

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Models include student- and school- random effects for the intercept term and no additional predictors. RE SDs are random effect standard deviations. ICCs are intraclass correlations. Full sample (columns 1 and 3) includes Hispanic students who took MAP Growth assessments between 2004-05 and 2018-19. Ever-ELs only is a subsample of the Hispanic students and includes students who were ever eligible to receive EL services while enrolled in the district.

Appendix A3. Calculating months of exposure to school

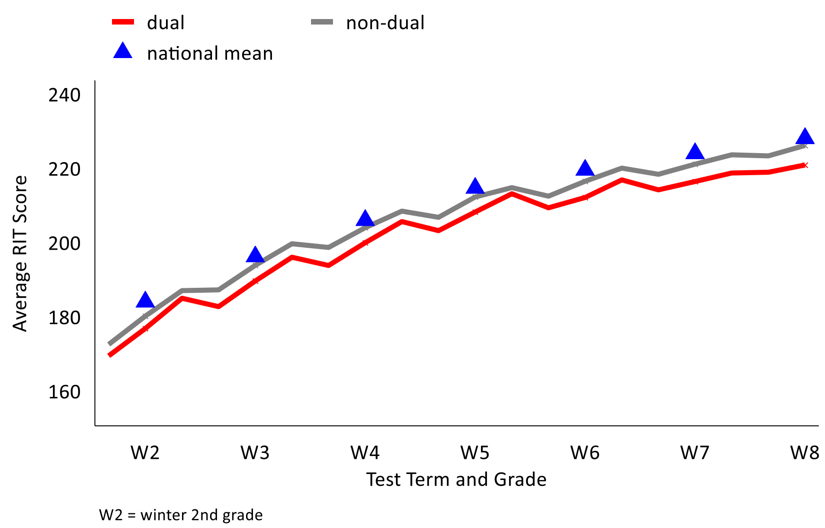
To set up the design matrix for this seasonal learning model, I calculate three sets of time variables: (a) number of months in school prior to testing, (b) total number of months spent in school across the whole school year, and (c) months of summer vacation. Time before testing was calculated as the difference between the school start date and test administration date for each student. The total number of months in school is calculated as the end date subtracted by the school start date, divided by 30.25 days per month. The months of summer vacation is the fall school start date subtracted by the prior year spring end date, divided by 30.25 days per month. For example, if a student tests in the fall of 3rd grade, they have been exposed to all of 2nd grade, a couple months of summer vacation after 2nd grade, and one or two months of 3rd grade. Since they have not been exposed to another summer vacation or 4th grade, the values for those predictors are set to zero.

Table A3. Monthly Exposure Rates for a Hypothetical Student Testing in 2nd and 3rd Grade

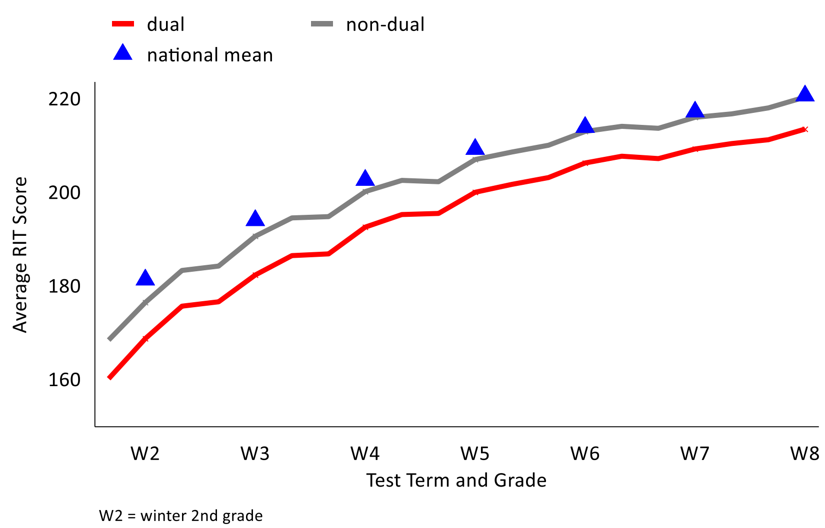
				Monthly Exposure Design Matrix					
Grade/Term	School Start Date	School End Date	Test date	Int.	G2	Sum2	G3	Sum3	G4
Fall 2nd	8/20/2015	6/12/2016	9/1/2015	1.00	0.39	0.00	0.00	0.00	0.00
Winter 2nd	8/20/2015	6/12/2016	12/1/2015	1.00	3.39	0.00	0.00	0.00	0.00
Spring 2nd	8/20/2015	6/12/2016	5/1/2016	1.00	8.23	0.00	0.00	0.00	0.00
Fall 3rd	8/19/2016	6/11/2017	9/15/2016	1.00	9.82	2.25	0.89	0.00	0.00
Winter 3rd	8/19/2016	6/11/2017	11/20/2016	1.00	9.82	2.25	3.11	0.00	0.00
Spring 3rd	8/19/2016	6/11/2017	4/1/2017	1.00	9.82	2.25	7.26	0.00	0.00

Figure A4. Math and Reading Achievement Trends for Post-Dual Language Program Cohorts

MAP Growth Math by dual language participation



MAP Growth Reading by dual language participation



Notes: Dual represents students who ever enrolled in the dual language program. Non-dual represents students who never enrolled in the dual language program. Trends represent unconditional means for each group of students at each test grade and term. Sample includes Hispanic students who took MAP Growth assessments between 2011-12 and 2018-19.

Table A5. Monthly Growth Rates for Post-Dual Language Program Cohorts

	Math			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Starting RIT	171.609***	179.503***	179.473***	167.584***	175.961***	175.938***
	(0.620)	(0.664)	(0.604)	(0.839)	(0.768)	(1.085)
Gr 2 AY	1.788***	1.785***	1.784***	1.839***	1.833***	1.833***
	(0.033)	(0.033)	(0.033)	(0.027)	(0.027)	(0.027)
DL x Gr 2 AY	0.182***	0.184***	0.184***	0.089*	0.093*	0.093*
	(0.063)	(0.063)	(0.063)	(0.049)	(0.049)	(0.049)
Gr 2 Summer	-0.489***	-0.474***	-0.473***	-0.183**	-0.160*	-0.160*
	(0.106)	(0.105)	(0.105)	(0.088)	(0.087)	(0.087)
DL x Gr 2 Summer	-1.115***	-1.126***	-1.126***	-0.293**	-0.312**	-0.312**
	(0.148)	(0.148)	(0.148)	(0.126)	(0.125)	(0.125)
Gr 3 AY	1.441***	1.439***	1.439***	1.148***	1.146***	1.146***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
DL x Gr 3 AY	0.177***	0.178***	0.178***	0.057	0.059*	0.059*
	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Gr 3 Summer	-0.872***	-0.865***	-0.865***	-0.141	-0.133	-0.133
	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)
DL x Gr 3 Summer	-0.531***	-0.538***	-0.538***	0.070	0.059	0.058
	(0.104)	(0.105)	(0.105)	(0.116)	(0.116)	(0.116)
Gr 4 AY	1.125***	1.123***	1.123***	0.840***	0.838***	0.837***
	(0.031)	(0.031)	(0.031)	(0.025)	(0.025)	(0.025)
DL x Gr 4 AY	0.259***	0.261***	0.261***	0.083**	0.086**	0.086**
	(0.036)	(0.036)	(0.036)	(0.038)	(0.038)	(0.038)
Gr 4 Summer	-0.758***	-0.754***	-0.754***	-0.177**	-0.173**	-0.173**
	(0.076)	(0.075)	(0.075)	(0.086)	(0.085)	(0.085)
DL x Gr 4 Summer	-0.642***	-0.647***	-0.647***	0.118	0.108	0.108
	(0.110)	(0.110)	(0.110)	(0.094)	(0.094)	(0.094)
Gr 5 AY	0.891***	0.891***	0.891***	0.613***	0.613***	0.613***
	(0.028)	(0.028)	(0.028)	(0.031)	(0.031)	(0.031)
DL x Gr 5 AY	0.282***	0.282***	0.282***	0.089**	0.089**	0.089**
	(0.038)	(0.038)	(0.038)	(0.041)	(0.041)	(0.041)
Gr 5 Summer	-1.427***	-1.431***	-1.431***	0.252***	0.246***	0.246***
	(0.082)	(0.082)	(0.082)	(0.095)	(0.095)	(0.095)
DL x Gr 5 Summer	-0.618***	-0.615***	-0.615***	-0.048	-0.046	-0.047
	(0.102)	(0.102)	(0.102)	(0.113)	(0.113)	(0.113)
Gr 6 AY	0.878***	0.878***	0.878***	0.489***	0.490***	0.490***
	(0.029)	(0.029)	(0.029)	(0.023)	(0.023)	(0.023)
DL x Gr 6 AY	0.034	0.035	0.035	0.042	0.043	0.043
	(0.035)	(0.035)	(0.035)	(0.037)	(0.037)	(0.037)
Gr 6 Summer	-0.957***	-0.957***	-0.958***	-0.356***	-0.356***	-0.357***
	(0.079)	(0.079)	(0.079)	(0.064)	(0.064)	(0.064)
DL x Gr 6 Summer	-0.312***	-0.312***	-0.311***	-0.006	-0.009	-0.009
	(0.103)	(0.103)	(0.103)	(0.108)	(0.109)	(0.109)
Gr 7 AY	0.613***	0.613***	0.613***	0.315***	0.316***	0.316***
	(0.035)	(0.035)	(0.035)	(0.026)	(0.026)	(0.026)
DL x Gr 7 AY	-0.054	-0.055	-0.055	0.086**	0.085**	0.085**
	(0.043)	(0.043)	(0.043)	(0.038)	(0.038)	(0.038)

Gr 7 Summer	-0.410***	-0.411***	-0.412***	0.239**	0.238**	0.238**
	(0.089)	(0.089)	(0.089)	(0.102)	(0.102)	(0.102)
DL x Gr 7 Summer	0.139	0.138	0.139	-0.263**	-0.264**	-0.264**
	(0.107)	(0.106)	(0.106)	(0.113)	(0.113)	(0.113)
Gr 8 AY	0.741***	0.740***	0.740***	0.662***	0.663***	0.663***
	(0.061)	(0.061)	(0.061)	(0.078)	(0.078)	(0.078)
DL x Gr 8 AY	-0.114	-0.112	-0.112	0.081	0.080	0.081
	(0.081)	(0.081)	(0.081)	(0.104)	(0.104)	(0.104)
Ever-DL	-3.067***	-0.414	-0.328	-8.384***	-3.790***	-3.713***
	(0.574)	(0.569)	(0.572)	(0.594)	(0.603)	(0.597)
Female		-1.634***	-1.627***		1.054***	1.060***
		(0.151)	(0.152)		(0.171)	(0.172)
Ever-FRPL		-3.955***	-3.843***		-4.455***	-4.356***
		(0.308)	(0.309)		(0.428)	(0.436)
Ever-EL		-5.284***	-5.269***		-8.512***	-8.500***
		(0.414)	(0.412)		(0.459)	(0.454)
Ever-SPED		-11.421***	-11.417***		-14.580***	-14.576***
		(0.352)	(0.352)		(0.350)	(0.350)
Refused EL Service		0.832**	0.843**		2.159***	2.171***
		(0.350)	(0.349)		(0.402)	(0.398)
School % FRPL			-8.362**			-11.799
			(3.349)			(7.216)
School % EL			4.613			9.340
			(7.576)			(16.336)
Student Covariates	no	yes	yes	no	yes	yes
School Covariates	no	no	yes	no	no	yes
Test Scores	156212	156212	156212	155459	155459	155459
Students	14576	14576	14576	14576	14576	14576
Schools	50	50	50	50	50	50

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes Hispanic students who took MAP Growth assessments between 2011-12 and 2018-19. Each column is a separate 3-level piecewise growth model regression with RIT score as the dependent variable. Starting RIT is the intercept and represents test scores on the first day of 2nd grade. Ever-DL is an indicator for ever enrolling in the dual language program. Female, Ever-FRPL, Ever-EL, Ever-SPED, and refused EL service are student-level covariates. Refused EL service represents having refused EL service in either kindergarten or 1st grade. School % FRPL and school % EL are school-level covariates. Interaction effects representing differential growth for dual language participants are bolded.

Table A6. Monthly Growth Estimates for Subsample with F&P Scores

	Math				Reading			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Starting RIT	171.736***	179.854***	171.104***	171.006***	167.028***	175.466***	164.403***	164.272***
	(0.573)	(0.643)	(0.869)	(0.814)	(0.695)	(0.756)	(1.008)	(0.987)
Gr 2 AY	1.884***	1.882***	1.881***	1.881***	1.986***	1.984***	1.982***	1.982***
	(0.029)	(0.029)	(0.030)	(0.030)	(0.025)	(0.025)	(0.025)	(0.025)
DL x Gr 2 AY	0.120*	0.121*	0.120*	0.120*	-0.024	-0.024	-0.024	-0.024
	(0.071)	(0.071)	(0.071)	(0.071)	(0.060)	(0.060)	(0.060)	(0.060)
Gr 2 Summer	-0.747***	-0.737***	-0.734***	-0.734***	-0.473***	-0.462***	-0.458***	-0.457***
	(0.095)	(0.095)	(0.096)	(0.096)	(0.090)	(0.089)	(0.089)	(0.089)
DL x Gr 2 Summer	-1.048***	-1.055***	-1.055***	-1.056***	-0.064	-0.069	-0.068	-0.068
	(0.131)	(0.130)	(0.130)	(0.131)	(0.131)	(0.130)	(0.130)	(0.130)
Gr 3 AY	1.486***	1.485***	1.482***	1.483***	1.247***	1.247***	1.245***	1.245***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.034)	(0.034)	(0.034)	(0.034)
DL x Gr 3 AY	0.181***	0.182***	0.184***	0.184***	-0.005	-0.005	-0.003	-0.003
	(0.034)	(0.034)	(0.035)	(0.035)	(0.036)	(0.037)	(0.037)	(0.037)
Gr 3 Summer	-0.931***	-0.927***	-0.923***	-0.923***	-0.352***	-0.351***	-0.346***	-0.346***
	(0.095)	(0.095)	(0.096)	(0.096)	(0.098)	(0.098)	(0.099)	(0.099)
DL x Gr 3 Summer	-0.724***	-0.727***	-0.730***	-0.730***	0.277**	0.275**	0.271**	0.272**
	(0.106)	(0.105)	(0.105)	(0.105)	(0.132)	(0.132)	(0.133)	(0.133)
Gr 4 AY	1.158***	1.157***	1.156***	1.156***	0.882***	0.881***	0.880***	0.880***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.026)	(0.026)	(0.026)	(0.026)
DL x Gr 4 AY	0.303***	0.304***	0.304***	0.304***	0.024	0.025	0.025	0.025
	(0.045)	(0.045)	(0.045)	(0.045)	(0.041)	(0.041)	(0.041)	(0.041)
Gr 4 Summer	-0.802***	-0.799***	-0.799***	-0.798***	-0.263***	-0.260***	-0.261***	-0.260***
	(0.090)	(0.090)	(0.091)	(0.091)	(0.087)	(0.087)	(0.087)	(0.087)
DL x Gr 4 Summer	-0.711***	-0.712***	-0.711***	-0.712***	0.161	0.158	0.160	0.160
	(0.125)	(0.125)	(0.125)	(0.125)	(0.102)	(0.101)	(0.101)	(0.101)
Gr 5 AY	0.977***	0.977***	0.977***	0.977***	0.678***	0.678***	0.678***	0.678***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.032)	(0.032)	(0.032)	(0.032)
DL x Gr 5 AY	0.314***	0.313***	0.314***	0.314***	0.047	0.046	0.046	0.046
	(0.050)	(0.051)	(0.051)	(0.051)	(0.041)	(0.041)	(0.041)	(0.041)

Gr 5 Summer	-1.433***	-1.433***	-1.435***	-1.435***	0.095	0.094	0.092	0.092
	(0.080)	(0.080)	(0.079)	(0.079)	(0.095)	(0.095)	(0.095)	(0.095)
DL x Gr 5 Summer	-0.910***	-0.908***	-0.907***	-0.907***	0.127	0.129	0.131	0.131
	(0.120)	(0.120)	(0.119)	(0.119)	(0.102)	(0.102)	(0.102)	(0.102)
Gr 6 AY	0.907***	0.907***	0.907***	0.907***	0.531***	0.531***	0.530***	0.530***
	(0.034)	(0.034)	(0.034)	(0.034)	(0.032)	(0.032)	(0.032)	(0.032)
DL x Gr 6 AY	0.134***	0.134***	0.134***	0.134***	-0.032	-0.032	-0.032	-0.032
	(0.042)	(0.042)	(0.042)	(0.042)	(0.048)	(0.048)	(0.048)	(0.048)
Gr 6 Summer	-1.149***	-1.149***	-1.150***	-1.150***	-0.352***	-0.353***	-0.353***	-0.354***
	(0.083)	(0.083)	(0.083)	(0.083)	(0.086)	(0.086)	(0.086)	(0.086)
DL x Gr 6 Summer	-0.370***	-0.369***	-0.369***	-0.368***	0.088	0.090	0.091	0.091
	(0.109)	(0.109)	(0.109)	(0.109)	(0.114)	(0.114)	(0.114)	(0.114)
Gr 7 AY	0.605***	0.605***	0.604***	0.604***	0.286***	0.286***	0.285***	0.285***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.036)	(0.036)	(0.036)	(0.036)
DL x Gr 7 AY	-0.032	-0.033	-0.031	-0.031	0.085	0.084	0.087	0.086
	(0.042)	(0.042)	(0.042)	(0.042)	(0.071)	(0.071)	(0.071)	(0.071)
Gr 7 Summer	-0.472***	-0.472***	-0.471***	-0.471***	0.150	0.151	0.153	0.153
	(0.079)	(0.080)	(0.080)	(0.080)	(0.149)	(0.149)	(0.150)	(0.150)
DL x Gr 7 Summer	0.036	0.039	0.038	0.038	0.179	0.183	0.181	0.181
	(0.094)	(0.094)	(0.095)	(0.095)	(0.200)	(0.200)	(0.201)	(0.201)
Gr 8 AY	0.715***	0.716***	0.715***	0.715***	0.732***	0.732***	0.730***	0.730***
	(0.069)	(0.069)	(0.069)	(0.069)	(0.082)	(0.082)	(0.082)	(0.082)
DL x Gr 8 AY	0.452	0.443	0.441	0.441	-0.768***	-0.780***	-0.783***	-0.782***
	(0.302)	(0.300)	(0.301)	(0.300)	(0.267)	(0.264)	(0.265)	(0.264)
Ever-DL	-2.711***	-0.746	-0.408	-0.306	-6.515***	-1.910***	-1.510**	-1.366*
	(0.644)	(0.652)	(0.633)	(0.630)	(0.697)	(0.710)	(0.752)	(0.748)
Female		-2.236***	-2.396***	-2.389***		0.303	0.103	0.109
		(0.235)	(0.221)	(0.220)		(0.254)	(0.244)	(0.243)
Ever-FRPL		-4.220***	-3.326***	-3.152***		-4.531***	-3.406***	-3.194***
		(0.373)	(0.344)	(0.349)		(0.488)	(0.441)	(0.463)
Ever-EL		-4.076***	-4.208***	-4.171***		-7.882***	-8.040***	-7.996***
		(0.439)	(0.370)	(0.375)		(0.513)	(0.453)	(0.456)
Ever-SPED		-10.533***	-7.557***	-7.560***		-14.136***	-10.372***	-10.374***
		(0.380)	(0.530)	(0.528)		(0.340)	(0.540)	(0.538)

Refused EL Service		0.197	0.740**	0.710**		2.681***	3.357***	3.324***
		(0.430)	(0.352)	(0.349)		(0.520)	(0.413)	(0.414)
F&P Score			0.519***	0.519***			0.657***	0.657***
			(0.049)	(0.049)			(0.060)	(0.060)
School % FRPL				-4.812***				-6.213***
				(1.134)				(1.220)
School % EL				0.096				3.523
				(2.012)				(2.258)
Student Covariates	no	yes	yes	yes	no	yes	yes	yes
School Covariates	no	no	no	yes	no	no	no	yes
Test Scores	99046	99046	99046	99046	98574	98574	98574	98574
Students	9207	9207	9207	9207	9207	9207	9207	9207
Schools	42	42	42	42	42	42	42	42

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes Hispanic students who took MAP Growth assessments between 2011-12 and 2018-19 and took Fountas & Pinnell Literacy assessments in kindergarten or 1st grade.. Each column is a separate 3-level piecewise growth model regression with RIT score as the dependent variable. Starting RIT is the intercept and represents test scores on the first day of 2nd grade. Ever-DL is an indicator for ever enrolling in the dual language program. Female, Ever-FRPL, Ever-EL, Ever-SPED, and refused EL service are student-level covariates. Refused EL service represents having refused EL service in either kindergarten or 1st grade. F&P score is the highest F&P literacy gradient recorded in kindergarten or 1st grade. School % FRPL and school % EL are school-level covariates. Interaction effects representing differential growth for dual language participants are bolded.