Achievement and Growth for English Learners

Angela Johnson
NWEA

Abstract

This study reports achievement and growth from kindergarten to 4th grade for three groups of English Learners (ELs): (a) ever-ELs; (b) ELs consistently eligible for service; and (c) EL and Special Education dually-identified students. All three EL groups had lower test scores than never-ELs throughout K-4. In math, ELs grew more than never-ELs during academic years but lost more during summers. In reading, ELs grew less than never-ELs in K-1 and grew more in later grades, but ELs also lost more during summers. These findings suggest summer support is required to help ELs maintain and develop academic skills.

Acknowledgements

The author is grateful for helpful feedback from Nate Jensen, Megan Kuhfeld, and James Soland.

Achievement and Growth for English Learners

English Learners (ELs) in US public schools are a diverse student population. Every one in ten students has been classified as ELs at some point during their K-12 schooling (U.S. Department of Education (DOE), 2018). In 2014-15, about 75% of ELs were Hispanic or Latino, 11% were Asian, and 6% were White (U.S. DOE, 2018). Approximately 80% of EL students are US-born, reflecting the greater multilingualism in the nation (García & Kleifgen, 2018).

EL students bring rich cultural and linguistic assets into their schools and in turn receive a variety of services and experiences. Some are educated in mainstream English monolingual classrooms, with or without language support; some spend a large fraction of their day in designated English language development (ELD) courses with other ELs (Gándara & Orfield, 2012); still others participate in dual language or bilingual programs, in which instruction is conducted in both their home language and in English (García & Kleifgen, 2018). These different models foster and/or leverage students' bilingual or multilingual skills to different extents and in some cases, not at all. In two-way dual language programs, ELs develop literacy and content knowledge in their home language as well as English while modeling their home language for native English speaker peers. In many English monolingual instructional environments, however, ELs' home languages are not developed or acknowledged, resulting in deficit-driven depictions of EL students as semilingual instead of bilingual (Flores et al., 2015).

Consistent with this language-deficit perspective, earlier research studies also tended to highlight EL underachievement by showing large gaps between the achievement of students currently receiving EL service and students who are native or fluent users of English (e.g., Carnoy & Garcia, 2017; Hemphill & Vanneman, 2011; Polat et al., 2016). These comparisons ignore two important measurement and reporting practices that first set up and then maintain these gaps. The first has to do with reporting achievement scores by current EL status. Every year new students with low English proficiency enter EL status and students with high English proficiency exit EL status; as a result, only contrasting the achievement of current ELs and native or fluent English users can lead to overestimation of gaps, minimize the achievement of multilingual students as a group and fuel deficit framing (Hopkins et al., 2013; Kieffer & Thompson, 2018; Saunders & Marcelletti, 2013). Second, achievement scores on English monolingual assessments are, by default, likely to be lower bounds of EL students' academic skills. When assessments are administered only in English, it is not surprising that students who are still developing English proficiency would score lower than students with native or fluent English proficiency who have the same content knowledge or skills. Linguistic complexity in the English language assessment poses a barrier to accurate measurement of ELs' actual skills or ability (Abedi & Levine, 2013). Thus, cross-sectional monolingual achievement scores on English monolingual assessments will always show gaps without illuminating the academic progress of either group. Moreover, simply reporting outcomes at a given point in time for students who are or are not identified as EL overlooks the fact that English proficiency develops over time. Student outcomes need to be reported and examined in more nuanced way to better understand short- and long-term achievement trends for students, including how achievement develops in relation to English proficiency. Instead of comparing cross-sectional achievement scores by EL status, policy and practice will be better informed by data on academic progress over time, disaggregated within the population of students who have ever been identified as ELs.

Examining Within-Year Growth

Examining ELs' within-student academic progress over time is important for two reasons. First, within-student growth, especially within-year growth measured with interim assessments, immediately inform instructional practice. During the year, teachers can use growth data for individual students to set goals and tailor instruction to meet their specific needs. Across years, disaggregated growth data for EL subgroups inform programs and policies at the school and district levels. Separate summer learning estimates reveal additional opportunities for reinforcing and supporting learning during out-of-school time. Second, growth data provide an important supplement to achievement status in school accountability and teacher evaluation. Compared to achievement measured at one point in time, students' academic trajectories are less strongly tied to underlying socioeconomic inequalities and more reflective of the effects that schools have on learning (Atteberry & McEachin, 2020; Reardon, 2019). Because of structural inequalities faced by students and their families, including the use of monolingual English assessments, ELs tend to enter school with lower achievement scores compared to their peers. Schools and teachers serving EL students are charged with the crucial task of helping ELs make progress toward mastery of grade-level academic content; but progress takes time, and EL students are likely to lag behind other students in achievement level for the first few years while they develop English proficiency. Evaluations based on achievement level alone will likely result in penalties to schools and teacher serving large populations of EL students, as well as propagate deficit framing around student achievement. Examining growth, on the other hand, will reward schools and educators for the progress they are helping students make despite the systemic inequities students and their communities face.

Disaggregating EL Data

Each subgroup within the larger ever-EL population has distinct educational needs, and the first step to targeting these various needs is identifying the subgroup's academic achievement trajectories. As shown in Figure 1 and detailed in the next section, multilingual students are students who use a language other than English at home, in school, or both. Multilingual students who are developing English proficiency become eligible to receive language support services after an EL classification process, usually involving an English proficiency assessment. ELs are assessed annually and when they reach fluent English proficiency, they exit EL status or reclassify. Former or reclassified ELs, as well as ELs who continue to receive services ("current-ELs"), constitute the ever-EL population because all of them were ELs at some point during their schooling. Students' strengths and needs vary considerably within the ever-EL population. Downstream academic outcomes, such as reading and math achievement during middle and high school, can differ greatly between reclassified ELs and ELs who continue to require language services. In addition to language support, some ELs are identified as also requiring Special Education (SPED) services. These "dually-identified" students, an especially important population, is protected by multiple strands of federal education policy. As Figure 1 shows, dually-identified students are a subgroup of ever-ELs. Though national and state data are not available, anecdotal evidence from school districts suggest that dually-identified students tend to take considerably longer to reclassify than ELs without SPED eligibility. Figure 1 thus shows the dually-identified subgroup to be largely overlapping with the current-EL group.

Extant research on EL academic achievement has relied heavily on cross-sectional data, examining between-cohort differences in EL achievement and comparing ELs to native and initially-fluent users of English (e.g., Carnoy & Garcia, 2017; Kieffer & Thompson, 2018). To my best knowledge, only three studies have used longitudinal student samples, following students over time to track their academic achievement (Johnson, 2020a; Soland & Sandilos, 2020; Valentino & Reardon, 2015). These longitudinal studies tended to use data from a single school district, producing findings that are unlikely to generalize to other education contexts. Overall, the literature has faced two major limitations in its ability to track academic achievement by EL status. First, few databases and research studies have tracked the achievement trajectories of both ever-ELs and current-ELs; and (b) no large-scale research has investigated the academic achievement of dually-identified students.

Improving upon previous research in scale and richness of findings, this study estimates achievement growth in math and reading for ever-EL students and highlights two important subgroups within this population: (a) ELs who continue to be eligible for language support; and (b) dually-identified students eligible for both EL and SPED services. I leverage a unique data set that contains longitudinal achievement measures for 8,200 ever-ELs across 915 schools in 31 states and the District of Columbia. My research questions are:

1. How do achievement levels at a given grade and time for EL subgroups (i.e., ever-ELs, ELs who continue to require service, and dually-identified students) compare to students who were never ELs?
2. How does achievement growth for ever-ELs, ELs who continue to require service, and dually-identified students compare to that of never-EL students?

Background

Beginning with the Bilingual Education Act of 1968, the federal government provided financial assistance to local educational agencies to develop "new and imaginative elementary and secondary school programs" to meet the special needs of "limited-English-speaking" students from low-income families (Section 702). However, the law did not specify instructional models or practices; as a result, innovation was up to individual schools and districts. Some developed bilingual programs and materials; many did not. In fact, many states had laws requiring English-only instruction (Ovando, 2003). In the decades that followed, reauthorizations of the Bilingual Education expanded eligibility to all students with limited English proficiency (Stewner-Manzanares, 1988), but the goal of support shifted toward developing English proficiency and assimilation into mainstream classrooms, and English-only services replaced many bilingual programs (Ovando, 2003). In the 2000s, No Child Left Behind and the Common Core State Standards further established English language arts skills as front and center (Garía & Kleifgen, 2018). Consequently, students from multilingual and multicultural backgrounds continue to be educated and assessed primarily in English, often without recognition for their literacy and academic skills in other languages.

EL Classification and Reclassification

Following two landmark Supreme Court decisions (*Lau v. Nichols*, 1974; *Castañeda v. Pickard*, 1981), school districts are required to identify students who have not gained fluent English language proficiency and provide services to enable their meaningful participation in educational activities. The processes for identifying ELs vary by local context (Linquanti et al., 2016). The two most common features used in the identification process are a home language survey and an English proficiency test. Students who primarily use a language other than English at home and do not pass the initial test are classified as ELs and become eligible to receive language services such as sheltered academic content courses and English Language Development (ELD) courses.

Students who are initially classified as ELs are tested annually until they demonstrate fluent English proficiency (and in some cases English Language Arts skills) to exit EL status. The tests and standards set for EL reclassification vary substantially across state and local educational agencies and over time. Many states use only English proficiency test scores; others, including California, additionally require reaching a cut score on the state's ELA standardized test. In some school districts consultations with teachers, parents, and other stakeholders may also play a role in the decision (Estrada & Wang, 2018). After reclassification, students are no longer eligible for EL services and are taught in mainstream classes with native English users.

No national data exist on overall EL reclassification rates. Data from some states show that between 25% and 50% of students who start kindergarten as ELs reclassify before 4th grade and 70% to 85% before 8th grade (Greenberg Motamedi et al., 2016; Kieffer & Parker, 2016; Thompson, 2017). Since only students with high English proficiency (and in some contexts, high reading or ELA achievement) reclassify, students currently receiving EL service at any given grade are necessarily those with lower English proficiency and lower reading or ELA test scores. For this reason, recent literature on EL achievement highlights the need to separately examine current-ELs still receiving service and ever-ELs, which additionally include students who have exited service. Focusing on current-ELs can better identify the needs of students who are developing English proficiency; looking at all ever-ELs provides a broader view of achievement and progress by the larger ever-EL group. As such, disaggregating data can help target services and address deficit-based narratives.

Dually-Identified ELs

About 14% of ELs are dually-identified to receive language and SPED services (U.S. DOE, 2018). Federal law requires that all children between the ages of 3 and 21 have access to a free and appropriate public education. According to the Individuals with Disabilities in Education Act (IDEA), students with special needs due to a disability are entitled to necessary educational accommodations in the least restrictive environment (Carnock & Silva, 2019). Students who are eligible for SPED as well as EL services have a unique intersection of requirements for educational support, making them one of the highest-need student populations in the education system. The challenge is that their needs are also the hardest to identify. Discerning between developing language proficiency and disability needs is a difficult task, especially in younger learners (Carnock & Silva, 2019). Research has shown that poor assessment design with weak psychometric properties can result in inappropriate identification (Macswan & Rolstad, 2006), such as favoring early EL identifications and delaying diagnoses for a disability identification (Burr, 2019) or overidentifying Latino/a students for learning disabilities solely due to low English language proficiency (Ortiz & Polyzoi, 1986). National data on the rates at which dually-identified students exit EL status are unavailable; some districts report that very few dually-identified ELs reclassify, especially in states that have ELA requirements in addition to English proficiency. Academically, dually-identified students are difficult to track since accountability systems do not require achievement data to be disaggregated within the EL and SPED-eligible populations (Albus et al., 2015). As a result, our understanding of dually-identified students' academic progress is extremely limited. Only one study of which I am aware estimated growth rates for dually-identified students, and the estimates were imprecise due to small sample size (Johnson & Barker, 2021).

Tracking Academic Progress for EL Subgroups

EL status matters because it determines access to language services and academic opportunities. Ideally, ELs should receive language support for as long as, but no longer than, they need them, and those services should support and not harm their academic progress. However, research has shown EL status to be associated with lower academic expectations from teachers and the students themselves (Kanno & Kangas, 2014), as well as a lower likelihood of taking general and advanced courses in core subjects (Estrada, 2014; Umansky, 2016a, 2018), which contributes to lower academic achievement (Umansky, 2016b). These findings call into question whether the services are helping or hindering ELs' academic progress and whether schools are distributing educational opportunities equitably.

Recent research has used district administrative data to track academic performance longitudinally from elementary to middle school. With data from middle schools in a district in California, Soland and Sandilos (2020) modeled ever-ELs' growth in academic achievement and self-efficacy in tandem. In terms of academic growth, they found that ever-ELs had lower achievement than their non-EL peers in 5th grade and that during middle school ever-ELs grew faster than non-ELs in reading but at a similar rate in math. Interestingly, Soland and Sandilos (2020) also found initial self-efficacy at the start of middle school to be a significant predictor of academic growth during middle school. Valentino and Reardon (2015) followed students in another California district who entered kindergarten as ELs from 2nd to 7th grade and estimated their linear achievement trajectories This study found that ELs who enrolled in dual language immersion increased in z-scores, or made gains in ELA achievement rank relative to the state average from 2nd to 7th grade, while ELs in three other language programs with short-term or no home language instruction fell in rank relative to the state average; in math, ELs in all programs started in the spring of 2nd grade with higher achievement than the state average but dropped in z-scores between 2nd and 7th grade. Johnson (2020a) used vertically-scaled achievement data from a district in the Midwest to compare achievement and growth for Hispanic students who did or did not participate in a dual language program. She found that in math, ELs in the dual language program grew more than non-dual language students during each school year between grades 2 and 5 but lost more during each summer between those grades. In reading, ELs in dual language grew less during each school year and lost less during the summers.

Current Study

This study builds on Johnson (2020a) to provide novel evidence on academic achievement and growth for ever-ELs and subgroups of high policy relevance. Using student characteristics and rich assessment data collected in the fall, winter, and spring from kindergarten to 4th grade, I estimate achievement growth for the pooled ever-EL group and subgroups who continue to need service or are dually-identified. This study makes two main contributions. It is the first to provide growth estimates, including separate year and summer learning rates, for a large ever-EL sample comprised of students from across the nation. It is also the first to disaggregate data and report estimates separately for dually-identified students.

Data

The data for this study come from the NWEA Growth Research Database (GRD). The data fields include MAP Growth assessment scores and school-reported student demographics and EL and SPED service eligibility indicators. School districts across the nation choose to administer MAP Growth assessments for various purposes, including monitoring student achievement and growth, staff evaluation, and school accountability. Data from the GRD cover a sizable portion (more than 20%) of the K-12 student population but are not nationally representative.

Districts that administer MAP Growth assessments provide students' gender and race/ethnicity and can choose to also provide students' eligibility for and participation in EL and SPED services. Since reporting service eligibility data is optional, only a subset of districts provided complete data in these fields. Having verified the number of students eligible for EL and SPED services in the district against the National Center for Education Statistics (NCES) Common Core of Data (CCD), I restrict my analysis to the districts that provided complete data.[1]

Sample

While the GRD includes private and international schools, I focus only on U.S. public schools in this study. A school is included in the sample if it is in a district that served any EL student and reported complete data on EL services. Appendix Table A1 in the online supplemental materials presents a comparison of summary statistics of the 915 schools in this study to all public schools serving kindergarten in the CCD. Compared to all public schools, schools in my sample were more likely to be urban, less likely to be rural, and served higher percentages of Black students and students eligible for free or reduced-price lunch (FRPL) and lower percentages of Hispanic and White students.

My sample includes more than 56,000 students who attended kindergarten and took at least one MAP Growth assessment in 2014-15. I follow this intact kindergarten cohort for five years to 2018-19, or the end of their 4th grade. The maximum number of terms (fall, winter, spring) in which students were assessed was 15. Due to differences in assessment policies across districts and states, as well as student attrition, not all students were assessed during all 15 terms. Appendix Table A2 shows the number of students assessed at each term and the total number of terms students were assessed. More than 70% of the students were assessed for eight or more terms; about 27% of the students were assessed for all 15 terms. Among ever-EL students, around 80% of the students were assessed in eight or more terms; 30% were assessed in all 15 terms. As described in the Analysis section, I include all students in the kindergarten cohort in the main analyses regardless of attrition. As a sensitivity check, I repeat the analyses for the subsample of students who were assessed in all 15 terms.

Table 1 shows summary statistics for students in the full sample. I include all students attending the 915 schools that reported complete data for EL and SPED services, regardless of individual students' EL status. Demographics for students who took the MAP Growth math assessments are very similar to students who took the reading assessments because almost all students took both subjects. The math sample is 49% female, 5% Asian, 23% Black, 19% Hispanic, and 45% White.

[Table 1 here]

Students eligible for EL service in at least one year between kindergarten and 4th grade are categorized as "ever-ELs" in the data regardless of the duration and timing of EL service. Ever-ELs comprise 15% of the math sample. About 44% of the ever-ELs in the math sample are girls, 16% are Asian, 6% Black, 61% Hispanic, and 10% White.

About half of ever-ELs were consistently flagged as ELs for every term in which they were assessed. In the context of this study, this subgroup is analogous to current-ELs—they were ELs during all the terms in which they contributed assessment scores to the data. An important caveat is that about 38% of the students were assessed in K-3 but not in 4th grade, so if they reclassified after attrition from the data, I would not observe their reclassification. To mark this distinction, I refer to these students as always-ELs (in the observed data) instead of current-ELs (throughout K-4). The gender and racial/ethnic compositions for the always-EL group were similar to ever-ELs' but with a higher percentage of Hispanic students and lower percentage of White students.

Dually-identified students, or ever-ELs additionally eligible for SPED services at any time between kindergarten and 4th grade, comprise 2% of the full sample and 13% of ever-ELs. Only 32% of the dually-identified students were girls; Asian students formed a smaller fraction of this subgroup compared to all ever-ELs, and Hispanic and White students formed a larger fraction. The data do not have a student-level measure of socioeconomic status (SES), so I am unable to compare SES between subgroups of students in my sample.

Never-ELs, or students who were never eligible for EL service (including multilingual students who had initial high English proficiency), had similar percentages of female, Asian, and Black students, but a lower percentage of Hispanic and higher percentage of White students compared to ever-ELs.

Measures of Achievement

Students were tested using the math and English reading MAP Growth assessments up to three times (fall, winter, and spring) during each school year. MAP Growth assessments are computerized, adaptive tests aligned to state content standards. Each test takes approximately 40 to 60 minutes to administer. Achievement scores are reported on the Rasch unIT (RIT) scale, where RIT is a linear transformation of the logit scale units of the Rasch item response theory model.[2] Test scores are vertically scaled to allow estimation of growth within and across grades.

Analysis

Comparison of Achievement Levels

To visualize achievement status for the student groups across time, I plot the mean achievement scores in the fall, winter, and spring of each grade for ever-ELs, always-ELs, dually-identified, and never-EL students. The plot also shows the national average from NWEA achievement norms for comparison (Thum & Kuhfeld, 2020). Separate comparisons of the national average to Asian, Black, Hispanic, and White ever-ELs and the same comparisons for always-ELs are reported in the Online Supplemental Materials.

Monthly Learning Rates

To estimate academic growth, I apply a piecewise multilevel growth model separately to ever-EL, always-EL, dually-identified, and never-EL students (e.g., von Hippel et al., 2018). One 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). I test whether any differences in growth rates between student groups expand, stay the same, or diminish across grade levels.

The model accounts for variations in test dates and estimates growth as a linear function of students' exposure to each school year and summer. 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. Therefore, exposure to instruction varied. I calculate months of exposure based on school start and end dates and the test administration dates (see Online Appendix B for details). For example, a student testing at the end of August in 1st grade may have 9.7 months of exposure to kindergarten, 2.3 months exposure to summer following kindergarten, and one week of exposure to 1st grade.

At level 1, I model achievement conditional on exposure to school during the academic year for each grade level (e.g., $G0_i$ = kindergarten academic year) and exposure to summer after each grade level (e.g., $S0_i$ = summer after kindergarten).

Level 1 (time (t) within student (i)):

$$y_{ti} = \pi_{0i} + \pi_{1i}G0_i + \pi_{2i}S0_i + \pi_{3i}G1_i + \pi_{4i}S1_i + \pi_{5i}G2_i + \pi_{6i}S2_i + \pi_{7i}G3_i + \pi_{8i}S3_i + \pi_{9i}G4_i + e_{ti} \quad (1)$$

The 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" (von Hippel et al., 2018, p. 335). The intercept (π_{0i}) is the predicted score for student *i* testing on the first day of kindergarten, regardless of how many instructional days elapsed. The slopes ($\pi_{1i}, \ldots, \pi_{9i}$) are the monthly learning rates of student *i* during each school year and summer. Each test score y_{ti} is viewed as a linear function of the number of months that student *i* has been exposed to kindergarten ($G0_i$), 1st grade ($G1_i$), etc., through 4th grade ($G4_i$); and the number of months that the student has been exposed to the summers after kindergarten ($S0_i$) through 3rd grade ($S3_i$).

At level 2, I include a random intercept to allow students' starting achievement in fall of kindergarten to vary by student; slopes are treated as fixed. I start with an unconditional model that only includes random intercept (Model 1), then I run a second model that includes student-level covariates *female*, *Asian*, *Black*, *Hispanic, and OtherRace* for the intercept (Model 2), with White, male students being the omitted category. I report results from Model 2, the preferred specification.

Level 2 (student (i)): (2)

$$\pi_{0i} = \beta_{00} + \beta_{01}female_i + \beta_{02}Asian_i + \beta_{03}Black_i + \beta_{04}Hispanic_i + \beta_{05}OtherRace_i + r_{0i}$$

$$\pi_{1i} = \beta_{10}$$

$$\vdots$$

$$\pi_{9i} = \beta_{90}$$

Variance component specification:

$$e_{ti} \sim \mathrm{N}(0, \sigma_{ti}^2),\ r_i \sim \mathrm{MVN}(0, T_{St}).$$

Models are estimated using HLM 8.0 software (Raudenbush et al., 2019). For each subject, I apply the 2-level model to ever-EL students, and then to always-EL, dually-identified, and never-EL students separately.

Sensitivity Checks

To interrogate the potential effect of student attrition on my findings, I test the sensitivity of my results to restricting the sample to only students who were assessed in all 15 terms between the fall of kindergarten and the spring of 4th grade. This subsample of students also had a complete set of EL and SPED flags, which means that students who were always-ELs in this subsample did not reclassify before the end of 4th grade.

Findings

Comparison of Achievement Levels

Figure 2 shows average achievement at each test term from the fall of kindergarten to the spring of 4th grade for ever-ELs, always-ELs, dually-identified students, and never-ELs. The corresponding means (in RIT points) are reported in Appendix Table A3. In both math and English reading and across time, never-ELs consistently had the highest test scores which were above the national average, followed by ever-ELs, always-ELs, then dually-identified students. In math, ever-ELs started kindergarten below the national average but surpassed it in 3rd grade. Always-ELs started kindergarten slightly below ever-ELs and stayed in a trajectory parallel to ever-ELs. Dually-identified students started kindergarten slightly below always-ELs but the gaps between themselves and other groups widened over time, with larger dips in the summers for dually-identified students than other groups. In English reading, the trajectories of the groups never crossed. Always-ELs started kindergarten slightly below ever-ELs, with the gap between the two groups growing larger over time. Dually-identified students consistently had the lowest scores, and the gaps between them and the other groups expanded over time, driven by larger summer learning loss.

[Figure 2 here]

Within the ever-EL category, visible differences exist across race/ethnicity groups (see Figure A1 in the Supplemental Materials). In both math and English reading, Asian and White ever-EL students scored similarly as or lower than the national average from kindergarten to 2nd grade and higher than the national average from 2nd to 4th grade. In math, Black and Hispanic ever-ELs started kindergarten well below the national average but the gap shrank substantially by the end of 4th grade. In English reading, Black consistently scored above Hispanic ever-ELs scored from kindergarten to 4th grade. Both groups started kindergarten slightly below the national average, with disparities expanding during the kindergarten year and shrinking in subsequent years.

Within the always-EL category, differences by race/ethnicity are also clear. As shown in Supplemental Figure A2, Asian always-ELs started kindergarten below or close to the national average but surpassed the national average in 2nd grade and continued to score higher through 4th grade. Black, Hispanic, and White always-ELs consistently scored similarly as one another and below the national average from kindergarten to 4th grade. In math, the disparities between Black, Hispanic, and White always-ELs and the national average shrank over time; in English reading, the disparities expanded over time.

Monthly Learning Rates

Figure 3 depicts monthly learning rates (in RIT points) and their 95% confidence intervals for each academic year and summer, estimated separately for each group of students using Model 2 (for coefficients and standard errors see Appendix Table A4). The top panel illustrates positive growth during the years; negative bars in the bottom panel represent summer learning loss. The four student groups had distinct growth patterns.

In math, the three EL groups tended to grow more during the year than never-ELs, and always-ELs had the highest growth rates. In the summers, ever-ELs had the lowest math learning loss rates; always-ELs and dually-identified students had the highest. In reading, the three EL groups had lower growth rates than never-ELs during the first two grades but higher growth rates during 3rd and 4th grade. Similar to math, ever-ELs also lost the least in reading during summers among the four student groups while dually-identified students lost the most.

[Figure 3 here]

Sensitivity Checks

Results for the subsample of students who tested in all 15 terms between the fall of kindergarten and the spring of 4th grade were similar to the full sample (see Appendix Tables A5 and A6). This restricted sample and the subgroups within it had slightly higher math and reading test scores in each test term compared to their full-sample counterparts. But the between-subgroup comparisons were qualitatively the same. In terms of achievement scores at each point in time, never-ELs were the highest and consistently higher than the national average, followed by ever-ELs, always-ELs, then dually-identified students. In terms of growth, the three EL groups tended to grow more in math than never-ELs during the academic years then lost more during the summers; in reading, the three EL groups grew less than never ELs during the first two academic years but more during later years, and the EL groups tended to lose more learning during the summers.

Discussion

This study examines academic achievement and growth for a unique EL sample and report two main findings. First, ever-EL as a group were closing in on the national average in math and English reading achievement by the end of 4th grade. Specifically, Asian and White ever-ELs surpassed the national average in 2nd grade; Black and Hispanic ELs continued to reduce the disparities between themselves and the national average in math during all five years. Second, students who continued to require EL support and students who additionally needed SPED services grew more than never-ELs during the academic years but also lost more learning during the summers.

These findings add to the body of recent evidence (e.g., Kieffer & Thompson, 2018; Valentino & Reardon, 2015) affirming ever-ELs' academic progress and potential. For instance, Valentino and Reardon (2015) showed, using spring-to-spring z-score changes, that ELs were able to rise in achievement rank against their grade-level peers in the state. My results on achievement and growth rates in the early grades complement their findings. Challenging the deficit narrative portraying ELs as underachieving, my findings show that the EL subgroups are capable of *growing more* during the school year than never-ELs.

Both this study and Valentino and Reardon (2015) point to the need to focus on academic progress over time as a key outcome, especially for EL students who are assessed in English. When all students are assessed in English, students who are developing English proficiency will likely score lower than their English-proficient peers, and their scores at any point in time may underestimate their current achievement level compared to if those students were assessed in their home language. Since the precise extent to which English assessment scores underestimate achievement for individual EL students is unknown, test scores at one point in time do not provide accurate or actionable evidence for educators or policymakers. Gains over time, on the other hand, is a more informative measure of academic development that reflects progress made by the students and their schools. In contexts in which all students must be assessed in English, measures of learning gains should be considered the key indicator of EL academic progress in both instruction and accountability.

More importantly, the school system needs to provide multilingual students with opportunities to develop the linguistic assets they bring to the classroom and accurately measure and recognize their full sets of skills and progress. English-only and English-dominant classrooms inevitably place students who are developing English proficiency at a disadvantage. Two-way dual language instruction has the potential for elevating both the multilingual student and their home language (Garcia, 2002). However, as Valdés (2018) cautioned, providing two-way dual language instruction alone is not sufficient; programming and instruction must be thoughtfully designed to meet the needs of all students, including immigrant and racialized groups. The greater task of establishing linguistic equity and social justice begins with education that grows the resources students bring to school. Assessment plays a crucial role in informing programs and policy to meet this goal. Students must be given equitable opportunities to demonstrate their true and complete skills and proficiencies. At the very least, appropriate accommodations should be provided to EL students on monolingual English assessments to reduce bias that results from unnecessary linguistic complexity (Abedi et al., 2020). To the greatest extent possible, dual-language or home-language assessments should be provided in addition to English assessments to measure student achievement and progress, especially when scores are used to make high-stakes decisions such as service eligibility, course placement, staff evaluation, and school accountability. It is important to assess bilingual students in both languages. The home-language assessment removes language barriers to help schools track student achievement more accurately over time; this is critical to ensuring that students are learning and making progress. The English assessment, on the other hand, allows schools to determine if students are making improvements over time as measured in English. These complementary measures of achievement, in addition to assessments of English proficiency, will help schools to identify programs and services that best support the students' needs.

This study also contributes novel evidence on seasonal patterns of learning, identifying summer learning loss as an ongoing challenge facing EL students who need the most support. My results are consistent with findings from two recent studies. Johnson (2020a) showed that Spanish-English dual language students, who grew more in math during the academic years, also lost more learning during the summers. Johnson and Barker (2021) found that students ever in special education services, and specifically dually-identified students, grew more than or as much as students never in special education during some early grades; but students with special needs also lost more learning during the summers. To achieve equitable learning outcomes in the long run, differential summer loss must first be addressed; otherwise, faster growth rates during academic years alone will be at least partly undone when school is out of session.

Prior work on ELs' academic access shed light on systemic barriers to success, including limited course access, low teacher expectations, and insufficient support. For example, current-EL status has been shown to preclude middle school students from taking a full load of academic content courses and from taking upper-level classes in content subjects (Umansky, 2016b, 2018). These studies pointed to programs and services offered during the academic years as potential explanations for opportunity gaps and points of intervention. I show that above and beyond instruction and services delivered in the school year, summers offer additional opportunities to provide support to the most vulnerable students. In order to help students maintain and continue to develop the language and academic skills they gained during the school year, schools and communities might consider offering programs during summers and other out-of-school time. Research has shown targeted summer programs of extensive duration to have positive effects on academic outcomes for high school newcomer ELs (Johnson, 2020b) and elementary school students from socioeconomically-disadvantaged families (Augustine et al, 2016). Experts in research and practice might consider designing programs that would expand learning opportunities beyond the academic year for current-ELs and dually-identified students.

In addition, my results highlight that schools need to provide dually-identified students with better support across the elementary grades, especially for reading during kindergarten and 1st grade. In order to address this important issue, researchers and practitioners must collaborate to create a system for monitoring student growth. Research on multilingual students' learning and growth is developing, and our current understanding of achievement and growth for students who are eligible for SPED services is extremely limited. Only one study of which I am aware estimated rates of academic progress for this group of students (Johnson & Barker, 2021). We do not know much, at any policy level (e.g., local, state, or federal), about where students eligible for various SPED services start academically in kindergarten or how much progress they make during each academic year, much less interactions between need for SPED services and language proficiency. To support dually-identified students' academic development, school leaders and experts in psychology, language acquisition, and educational assessment must work together to collect, analyze, and interpret high-quality data.

Concluding Remarks

Using data on a unique sample of students across the nation, this study provides the first estimates for academic growth for ever-ELs from kindergarten to 4th grade. My findings show that ever-EL students grow academically at rates comparable to or higher than their never-EL peers during the academic year but also suggest that more support is needed for summer learning. In order to close opportunity gaps in the long run, schools must provide support for ELs to grow more in every grade and address loss during out-of-school time. Recognizing and nurturing the linguistic and cultural assets students bring to school is critical to fostering their academic progress. Schools should provide opportunities for multilingual students to develop literacies and academic skills in their home language and ensure that assessment practices align with the goal of accurately measuring and affirming students' skills and progress. Future research should continue to leverage rich achievement and growth data to help schools better support their linguistically-diverse student populations.

A few limitations for this study merit consideration. First, I used English MAP Growth assessment scores as measures of achievement. Scores at a given time point may be underestimates for students who were developing English proficiency. Thus, the underlying disparities in achievement between always-EL and dually-identified students and the other groups may be overestimated. Future research should compare achievement scores for ELs and other students using bilingual or home-language assessments, such as MAP Growth administered in both Spanish and English.

The sample is comprised of students across the nation, but it is not nationally representative. The average student in the sample had higher achievement than the national average, and the schools were more likely to be urban and had higher proportions of minoritized students than all public schools in the US. The districts in the sample also likely differ from others in the nation in unobservable ways. All of these districts which provided rich, complete data on special services eligibility for research. Compared to other districts, they may have been more experienced or more motivated to use data for the purpose of addressing opportunity gaps and supporting historically-underserved student populations. For this reason, we might interpret the achievement gaps in my sample to be underestimates. However, it is also possible that the districts reported EL and SPED data because they were aware of having larger achievement gaps than other schools and wanted to address this issue through research. In this case, my estimated achievement gaps may be overestimates. We also need to bear in mind that EL classification criteria vary between states, thus, the initial English proficiency of ever-ELs also varies across states in the sample.

Additional challenges involve limitations to the data time frame and availability. The data did not allow me to examine achievement and growth beyond 4th grade for the students in my sample, and multilingual students' growth in the middle grades remains a crucial but understudied topic. Though prior research has shown that ELs' academic achievement and paths to English proficiency may differ by socioeconomic status and language program, I was not able to explore heterogeneity along these dimensions. EL and SPED programs and services differ substantially across schools. I observe program information only for a small number of students and was not able to make explore variations by program. I also did not have student-level variables for socioeconomic status, which has been shown by a lot of research to be a strong predictor of student achievement. Although I was able to conduct heterogeneity analysis for achievement scores by ethnicity, these results should be interpreted with the consideration that socioeconomic status effects likely confounds the relation between ethnicity and achievement. Finally, I was not able to estimate growth rates separately for EL groups by race/ethnicity because of the small subgroup sample sizes. Future studies should explore this very important topic, as recent scholarly attention focuses on the academic achievement of Black ELs and speakers of African American varieties of English (Valdés, 2018).

These limitations are not unique to this study and reflect larger issues in EL data reporting. For example, research using NAEP data must rely on reported language proficiency instead of EL classification. District and state administrative data may include indicators that allow researchers to disaggregate by EL and SPED service eligibility, but the results are not generalizable to other contexts. In addition, assessments currently used in many local and state contexts do not have psychometric properties that support the estimation of academic growth within and across years. These issues point to the urgent need to improve data reporting and tracking for EL and SPED-eligible students. A national initiative should be established to collect and curate actionable achievement and growth data, without which generalizability will remain a challenge.

Notes

[1] I interrogate the quality of the district-reported service eligibility data in two ways. First, I match GRD data to the CCD (NCES, 2017) to compare the total number of students eligible for EL and SPED services reported for each district. I retain districts for which the reported number of EL students from the two data sources that were within 10% of each other. Second, I examine the data files, which contain binary indicators for service eligibility as well as text fields for classification results or program participation. In this qualitative check, I verify that the text fields provided descriptions that were relevant to EL and SPED services. For instance, many of the observations included ELs' English proficiency level and language program type; and many of the SPED text fields included disability categories.

[2] Average test duration was slightly longer for ever-ELs compared to never-ELs. Standard error of measurement and percentage of rapidly-guessed items were similar for ever-ELs and never-ELs.

References

Abedi, J. & Levine, H. G. (2013). Fairness in assessment of English learners. *Leadership*, *42*(3), 26–38. https://eric.ed.gov/?id=EJ1004517

Albus, D., Lazarus, S. S., & Thurlow, M. L. (2013). *2012-13 Publicly Reported Assessment Results for Students with Disabilities and ELLs with Disabilities* (Technical Report 70). University of Minnesota, National Center on Educational Outcomes.

Atteberry, A. C. & McEachin, A. J. (2020). Not where you start, but how much you grow: An addendum to the Coleman Report. *Educational Researcher, 49*(9), 678-685.

Augustine, C. H., McCombs, J. S., Pane, J. F., Schwartz, H. L., Schweig, J., McEachin, A., & Siler-Evans, K. (2016). Learning from summer: Effects of two years of voluntary summer learning programs on low-income urban youth. Santa Monica, CA: RAND Corporation.

Bilingual Education Act. (1968). Title VII, Bilingual Education, Language Enhancement, and Language Acquisition Programs of Elementary and Secondary Education Act. P. L. 90-247.

Burr, E. (2019). *Guidance Manuals for Educators of English Learners with Disabilities: Ideas and Lessons from the Field* (NCEO Report 410). University of Minnesota, National Center on Educational Outcomes.

Carnock, J. T., & Silva, E. (2019). *English Learners with Disabilities: Shining a Light on Dual-Identified Students*. New America.

Carnoy, M. & Garcia, E. (2017). Five key trends in U.S. student performance. Economic Policy Institute Report. http://www.epi.org/publication/five-key-trends-in-u-s-student-performance-progress-by-blacks-and-hispanics-the-takeoff-of-asians-the-stall-of-non-english-speakers-the-persistence-of-socioeconomic-gaps-and-the-damaging-effect/.

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

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review, 46*(4), 535-573.

Estrada, P. & Wang, H. (2018). Making English learner reclassification to fluent English proficient attainable or elusive: When meeting criteria is and is not enough. *American Educational Research Journal, 55*(2), 207-242.

Flores, N., Kleyn, T., & Menken, K. (2015). Looking holistically in a climate of partiality: Identities of students labeled long-term English Language Learners. *Journal of Language, Identity & Education*, *14*(2), 113-132. 10.1080/15348458.2015.1019787

Fry, R. (2007). *How far behind in math and reading are English Language Learners?* PEW Hispanic Center.

Gándara, P. & Orfield, G. (2012). Why Arizona matters: The historical contexts of Arizona's instructional linguistic hegemony. *Language Policy, 11*, 7-19.

Garcia, E. E. (2002). Bilingualism and schooling in the United States. *International Journal of the Sociology of Language, 2002*(155–156), 1–92. 10.1515/ijsl.2002.028

Greenberg Motamedi, J., Singh, M., & Thompson, K. D. (2016). English learner student characteristics and time to reclassification: An example from Washington state (REL 2016–128). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. http://ies.ed.gov/ncee/edlabs

Hemphill, F.C., and Vanneman, A. (2011). Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress (NCES 2011-459). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

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

Johnson, A. (2020a). Dual language education and academic growth. (EdWorkingPaper: 20-300). Annenberg Institute at Brown University: https://doi.org/10.26300/7zth-7373

Johnson, A. (2020b). Summer credit recovery impact on newcomer English Learners. *American Educational Research Journal, 57*(4), 1757-1790.

Johnson, A. & Barker, E. (2021). Understanding differential growth during school years and summers for students in special education. (EdWorkingPaper: 21-409). Annenberg Institute at Brown University: https://doi.org/10.26300/0d43-qv21

Kanno, Y. & Kangas, S. E. (2014). "I'm not going to be, like, for the AP": English Language Learners' limited access to advanced college-preparatory courses in high school. *American Educational Research Journal, 51*(5), 848-878.

Kieffer, M. J. & Parker, C. E. (2016). *Patterns of English learner student reclassification in New York City public schools* (REL 2017–200). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http:// ies.ed.gov/ncee/edlabs.

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391–398.

Lau v. Nichols. 414 U.S. 563, 569. (1974).

Linquanti, R., Cook, H. G., Bailey, A. L., & MacDonald, R. (2016). Moving toward a More Common Definition of English Learner: Collected Guidance for States and Multi-State Assessment Consortia. Council of Chief State School Officers.

Macswan, J., & Rolstad, K. (2006). How language proficiency tests mislead us about ability: Implications for English language learner placement in special education. *Teachers College Record*, *108*(11), 2304–2328.

NCES. (2017). Common Core of Data.

Ortiz, A. A. & Polyzoi, E. (Eds.) (1986). *Characteristics of limited English proficient Hispanic students in programs for the learning disabled: Implications for policy, practice, and research.* University of Texas at Austin.

Ovando, C. J. (2003). Bilingual education in the United States: Historical development and current issues, *Bilingual Research Journal*, *27*(1), 1-24. 10.1080/15235882.2003.10162589

Polat, N., Zarecky-Hodge, A., & Schreiber, J. B. (2016) Academic growth trajectories of ELLs in NAEP data: The case of fourth- and eighth-grade ELLs and non-ELLs on mathematics and reading tests, *The Journal of Educational Research*, 109(5), 541-553, DOI: 10.1080/00220671.2014.993461

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.

Reardon, S.F. (2019). Educational opportunity in early and middle childhood: Using full population administrative data to study variation by place and age. *RSF: The Russell Sage*

Foundation Journal of the Social Science, 5(2), 40-68. https://cepa.stanford.edu/content/educational-opportunity-early-and-middle-childhood-using-full-population-administrative-data-study-variation-place-and-age,

Saunders, W. M. & Marcelletti, D. J. (2013). The gap that can't go away: The catch-22 of reclassification in monitoring the progress of English Learners. *Educational Evaluation and policy analysis*, *35*(2), 139-156.

Soland, J., & Sandilos, L. E. (2020). English Language Learners, Self-efficacy, and the Achievement Gap: Understanding the Relationship between Academic and Social-Emotional Growth. *Journal of Education for Students Placed at Risk*, 1-25.

Stewner-Manzanares, G. (1988). The Bilingual Education Act: Twenty years later. *The National Clearinghouse for Bilingual Education Occasional Papers in Bilingual Education*, 6.

Thompson, K. D. (2017). English Learners' Time to Reclassification: An Analysis. *Educational Policy*, *31*(3), 330–363. https://doi.org/10.1177/0895904815598394

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.

Thurlow, M. L., Albus, D. A., & Lazarus, S. S. (2017). *2014-15 publicly reported assessment results for students with disabilities and ELs with disabilities* (NCEO Report 405). University of Minnesota, National Center on Educational Outcomes.

Umansky, I. (2016a). "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.

Umansky, I. (2016b). "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.

Umansky, I. M. (2018). According to plan? Examining the intended and unintended treatment effects of EL classification in early elementary and the transition to middle school. *Journal of Research on Educational Effectiveness*, *11*(4), 588-621. DOI: 10.1080/19345747.2018.1490470

U.S. Department of Education. (2018). *Educational experiences of English Learners: What are their characteristics?* https://www2.ed.gov/datastory/el-characteristics/index.html#one

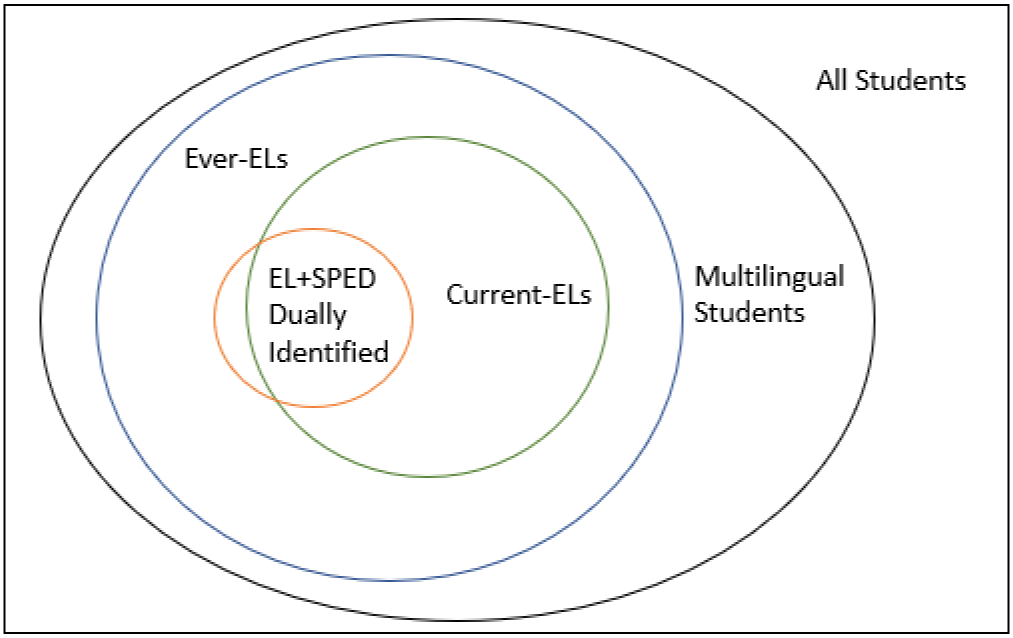
Valdés, G. (2018). Analyzing the curricularization of language in twoway immersion education: Restating two cautionary notes, Bilingual Research Journal, 41(4), 388-412. 10.1080/15235882.2018.1539886

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.

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.

Tables and Figures

Figure 1. English Learner (EL) Subgroups



__**Definitions**__

Multilingual students: students who use a language other than English, including students with fluent English proficiency and students who are developing English proficiency

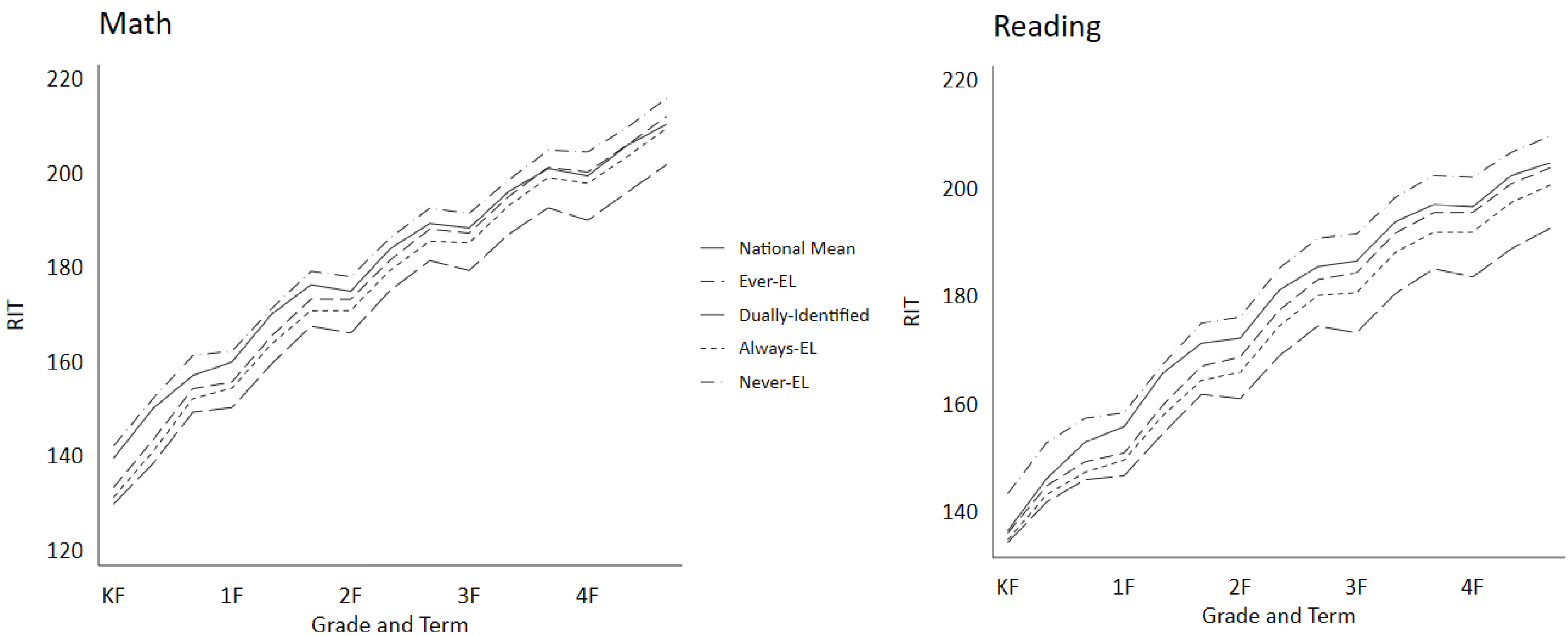
Ever-ELs: students who have ever been classified as English Learners

Current-ELs: students who are currently classified as English Learners, are eligible for language services, and have not attained fluent English proficiency

EL + SPED dually-identified students: students who are eligible for EL services as well as special education services

Never-ELs (not shown in Figure 1): English monolingual students and multilingual students who have never been classified as English learners

Figure 2. Achievement Scores by Student Group



Notes: KF = fall of kindergarten; 1F = fall of 1st grade, etc. EL = English Learner. Dually-identified = students who were ever eligible for both EL and Special Education services.

Figure 3. Estimated Monthly Growth Rates by Student Group

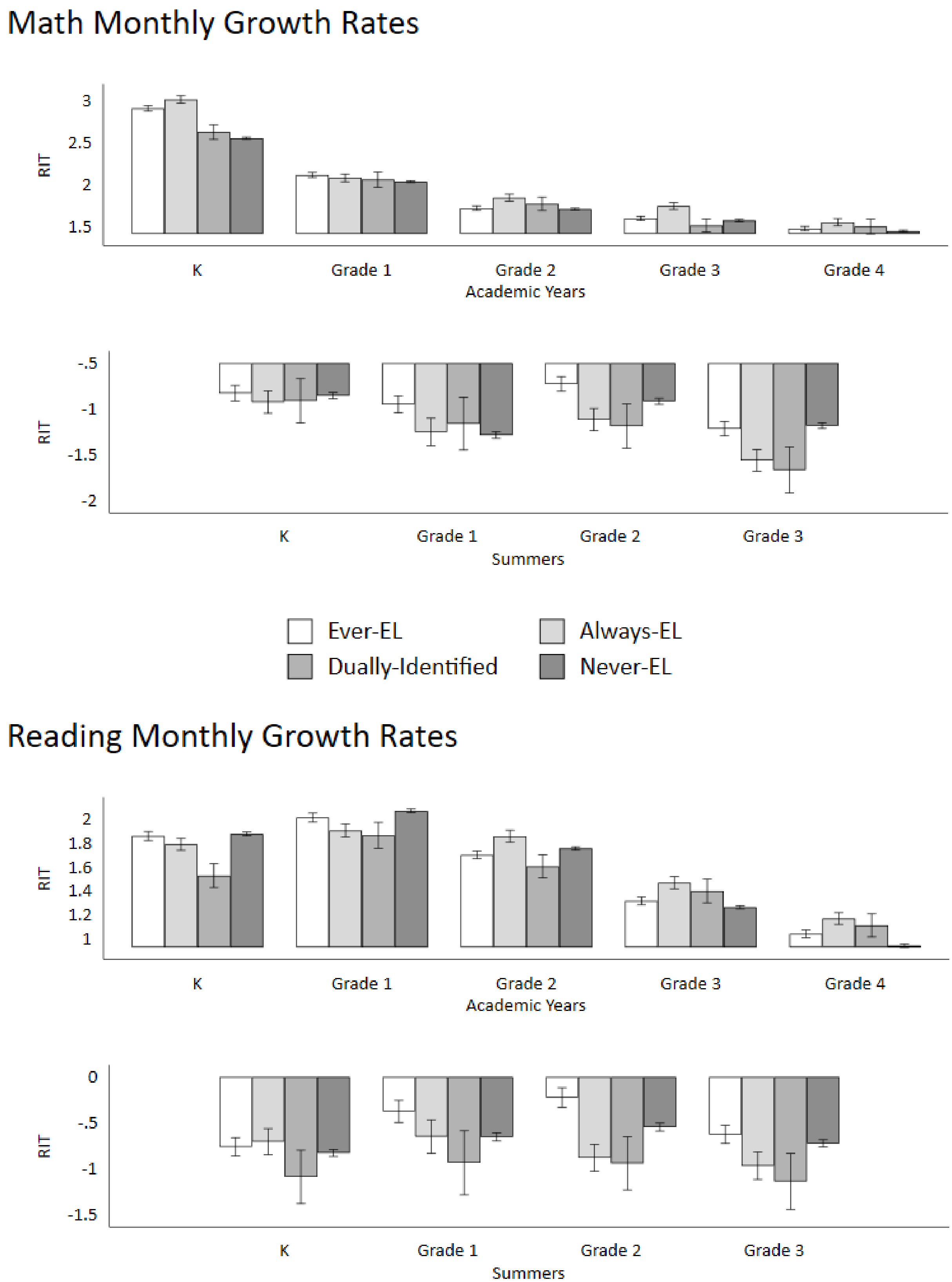


Table 1. Sample Summary Statistics

	All		Ever-EL		Always-EL		Dually-Identified		Never-EL	
Math	N=55946		N=8206		N=4088		N=1084		N=47740	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.49	0.50	0.47	0.50	0.45	0.50	0.32	0.47	0.49	0.50
Asian	0.05	0.23	0.16	0.37	0.15	0.36	0.11	0.31	0.04	0.19
Black	0.23	0.42	0.06	0.24	0.05	0.21	0.06	0.23	0.25	0.44
Hispanic	0.19	0.39	0.61	0.49	0.67	0.47	0.63	0.48	0.12	0.32
White	0.45	0.50	0.10	0.30	0.08	0.28	0.12	0.33	0.51	0.50
Reading	N=54090		N=7779		N=3971		N=1046		N=46311	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.49	0.50	0.47	0.50	0.45	0.50	0.33	0.47	0.49	0.50
Asian	0.05	0.22	0.16	0.37	0.15	0.36	0.11	0.31	0.03	0.18
Black	0.23	0.42	0.06	0.24	0.05	0.21	0.06	0.24	0.26	0.44
Hispanic	0.19	0.39	0.60	0.49	0.66	0.47	0.62	0.49	0.12	0.32
White	0.44	0.50	0.11	0.31	0.08	0.28	0.13	0.33	0.50	0.50

Notes: Means are column means (proportions). SD = standard deviation. EL = English Learner. Dually-Identified = students ever eligible for both EL and Special Education services.

Online Supplemental Materials

Appendix A. Supplemental Tables and Figures

Table A1. Comparison of Sample and All Public School Characteristics

	Sample Schools			All NCES 2015-16 Schools Serving Kindergarten		
	Mean	SD	N	Mean	SD	N
% Free/Reduced Price Lunch Eligible	0.57	0.29	915	0.55	0.30	55119
% Asian	0.05	0.07	915	0.04	0.09	55294
% Black	0.24	0.28	915	0.15	0.24	55294
% Hispanic	0.21	0.23	915	0.25	0.29	55294
% White	0.46	0.31	915	0.50	0.34	55294
City	0.42	0.49	915	0.30	0.46	55824
Town	0.09	0.28	915	0.11	0.31	55824
Rural	0.12	0.32	915	0.26	0.44	55824
Title I Eligible	0.82	0.39	915	0.77	0.42	55377
School-wide Title I	0.69	0.46	912	0.66	0.47	55010

Notes: Means = column means or proportions of students in school. SD = standard deviation. N = number of schools.

Table A2. Number of Students with Test Scores

Panel A: By Test Term

	N students	
Test Term	Math	Reading
K Fall	43514	42282
K Winter	46448	46207
K Spring	50282	47947
G1 Fall	40643	38188
G1 Winter	39253	36857
G1 Spring	39798	37376
G2 Fall	42242	40299
G2 Winter	40084	38480
G2 Spring	41202	39493
G3 Fall	37346	34538
G3 Winter	35572	32458
G3 Spring	37417	34387
G4 Fall	32601	31472
G4 Winter	31459	29804
G4 Spring	33195	31508

Panel B: By Total Number of Test Terms

	Math			Reading		
Number of Terms	Students	%	Cumul. %	Students	%	Cumul. %
1	1,829	3.27	3.27	1,787	3.30	3.30
2	2,078	3.71	6.98	2,053	3.80	7.10
3	3,331	5.95	12.94	3,394	6.27	13.37
4	1,269	2.27	15.21	1,260	2.33	15.70
5	1,772	3.17	18.37	1,886	3.49	19.19
6	3,207	5.73	24.11	2,894	5.35	24.54
7	1,300	2.32	26.43	1,489	2.75	27.29
8	1,601	2.86	29.29	1,930	3.57	30.86
9	2,577	4.61	33.90	3,407	6.30	37.16
10	1,980	3.54	37.44	1,721	3.18	40.34
11	3,456	6.18	43.61	3,181	5.88	46.22
12	7,157	12.79	56.41	6,663	12.32	58.54
13	3,490	6.24	62.64	3,344	6.18	64.72
14	5,756	10.29	72.93	4,983	9.21	73.94
15	15,143	27.07	100.00	14,098	26.06	100.00
Total	55,946	100.00	100.00	54,090	100.00	100.00

Table A3. Achievement (RIT) by Student Group

	Math				Reading			
Test Term	Ever-EL	Always-EL	Dually-ID	Never-EL	Ever-EL	Always-EL	Dually-ID	Never-EL
K Fall	133.36	131.30	129.83	142.20	136.21	134.96	134.45	143.52
K Winter	143.44	140.98	138.44	152.26	144.95	143.31	141.95	152.92
K Spring	154.34	152.19	149.29	161.35	149.45	147.51	146.14	157.53
G1 Fall	155.80	154.48	150.30	162.33	151.03	149.73	146.83	158.49
G1 Winter	165.55	163.84	159.62	171.38	159.90	157.96	154.62	167.49
G1 Spring	173.36	170.93	167.62	179.31	167.14	164.53	161.92	175.12
G2 Fall	173.31	170.96	166.16	178.12	168.88	166.03	161.14	176.24
G2 Winter	181.68	179.49	175.20	186.35	177.45	174.51	169.00	185.24
G2 Spring	188.24	185.68	181.56	192.69	183.19	180.28	174.57	190.82
G3 Fall	187.40	185.28	179.47	191.59	184.45	180.78	173.33	191.66
G3 Winter	195.21	193.21	187.15	198.72	191.81	188.22	180.60	198.41
G3 Spring	201.35	199.19	192.75	204.99	195.61	192.00	185.20	202.52
G4 Fall	200.34	197.93	190.11	204.63	195.64	192.01	183.66	202.17
G4 Winter	206.11	203.66	196.07	209.81	200.96	197.51	188.89	206.79
G4 Spring	212.20	209.66	201.95	216.08	203.92	200.68	192.70	209.78

Table A4. Estimated Monthly Growth Rates (RIT)

	Math				Reading			
	Ever-EL	Always-EL	Dually-ID	Never-EL	Ever-EL	Always-EL	Dually-ID	Never-EL
Intercept	132.607***	128.279***	131.431***	143.948***	136.462***	133.455***	137.932***	145.179***
	(0.434)	(0.651)	(1.157)	(0.101)	(0.448)	(0.647)	(1.135)	(0.098)
K Year	2.908***	3.013***	2.628***	2.554***	1.862***	1.791***	1.531***	1.878***
	(0.017)	(0.022)	(0.045)	(0.007)	(0.019)	(0.025)	(0.050)	(0.008)
K Summer	-0.827***	-0.924***	-0.910***	-0.854***	-0.759***	-0.702***	-1.085***	-0.825***
	(0.043)	(0.062)	(0.123)	(0.017)	(0.051)	(0.072)	(0.148)	(0.020)
G1 Year	2.117***	2.082***	2.063***	2.038***	2.014***	1.907***	1.867***	2.070***
	(0.016)	(0.023)	(0.046)	(0.006)	(0.019)	(0.027)	(0.055)	(0.007)
G1 Summer	-0.948***	-1.248***	-1.159***	-1.283***	-0.374***	-0.648***	-0.932***	-0.651***
	(0.047)	(0.076)	(0.146)	(0.019)	(0.061)	(0.093)	(0.177)	(0.023)
G2 Year	1.723***	1.848***	1.775***	1.713***	1.704***	1.859***	1.608***	1.758***
	(0.014)	(0.021)	(0.041)	(0.006)	(0.017)	(0.026)	(0.049)	(0.007)
G2 Summer	-0.727***	-1.114***	-1.185***	-0.915***	-0.223***	-0.879***	-0.938***	-0.545***
	(0.041)	(0.062)	(0.123)	(0.017)	(0.054)	(0.075)	(0.147)	(0.022)
G3 Year	1.602***	1.748***	1.518***	1.578***	1.321***	1.472***	1.403***	1.268***
	(0.013)	(0.021)	(0.039)	(0.006)	(0.017)	(0.026)	(0.050)	(0.007)
G3 Summer	-1.213***	-1.558***	-1.662***	-1.179***	-0.625***	-0.965***	-1.134***	-0.721***
	(0.040)	(0.059)	(0.128)	(0.016)	(0.050)	(0.076)	(0.154)	(0.020)
G4 Year	1.481***	1.557***	1.505***	1.452***	1.047***	1.174***	1.117***	0.948***
	(0.014)	(0.022)	(0.044)	(0.006)	(0.016)	(0.025)	(0.049)	(0.007)
Female	-0.750***	-0.759**	-2.476***	-0.391***	1.611***	1.408***	-0.441	2.275***
	(0.235)	(0.320)	(0.743)	(0.104)	(0.254)	(0.333)	(0.783)	(0.112)
Asian	5.243***	7.300***	4.173**	5.815***	4.741***	6.432***	4.243***	5.663***
	(0.526)	(0.794)	(1.621)	(0.315)	(0.547)	(0.783)	(1.593)	(0.321)
Black	-3.222***	-1.525	-2.192	-10.597***	-1.899***	0.309	-0.848	-9.821***
	(0.636)	(0.977)	(2.020)	(0.127)	(0.665)	(0.964)	(2.075)	(0.136)
Hispanic	-3.464***	-0.905	-5.674***	-5.368***	-4.187***	-1.463**	-6.602***	-5.515***
	(0.428)	(0.643)	(1.158)	(0.158)	(0.453)	(0.646)	(1.171)	(0.171)
Other Ethnicity	-0.618	-0.697	-2.491	-3.503***	-0.375	-0.354	-3.028*	-3.410***
	(0.598)	(0.925)	(1.641)	(0.202)	(0.633)	(0.967)	(1.699)	(0.210)
Tests	92274	43070	13053	498782	85489	40585	12380	475807
Students	8206	4088	1084	47740	7779	3971	1046	46311
Intercept-Variance	109.80	99.32	134.00	125.00	119.40	102.00	139.90	137.20

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column is a separate regression with sample restricted to the titular subgroup. K=kindergarten. G1= 1st grade. Estimates are from Model 2 described on page 14.

Table A5. Achievement Means (RIT) for Students with All 15 Terms

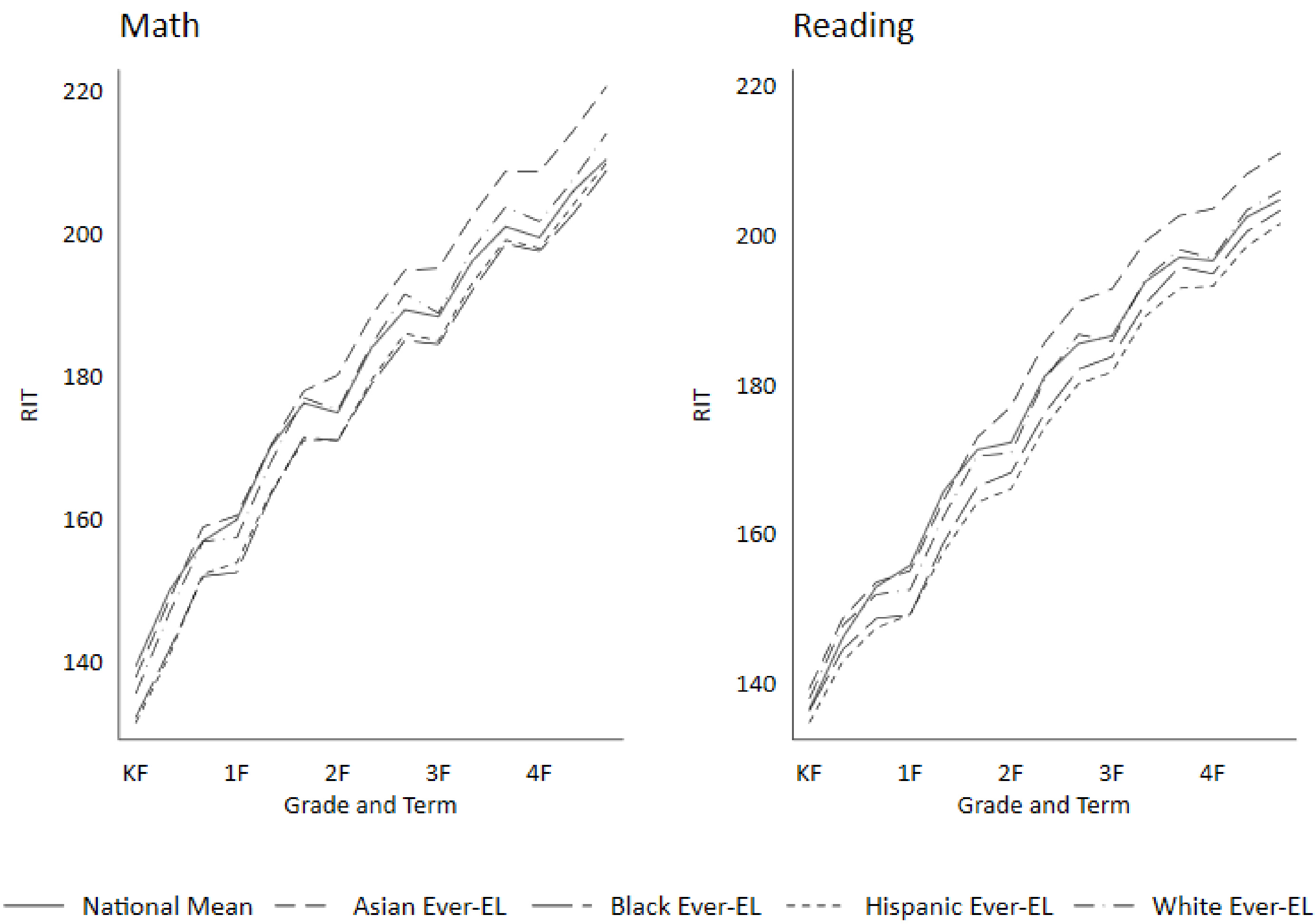
	Math				Reading			
Test Term	Ever-EL	Always-EL	Dually-ID	Never-EL	Ever-EL	Always-EL	Dually-ID	Never-EL
K Fall	134.02	131.97	130.89	143.69	137.09	135.77	134.79	144.99
K Winter	145.12	142.45	139.74	154.83	146.64	144.97	143.72	155.00
K Spring	156.57	153.94	151.28	163.53	152.09	150.13	148.44	159.78
G1 Fall	158.16	156.08	151.94	164.82	153.52	151.58	148.47	161.20
G1 Winter	167.60	165.39	161.09	173.91	162.56	159.76	156.53	170.64
G1 Spring	175.52	172.37	168.80	181.51	170.43	166.82	164.02	177.81
G2 Fall	174.98	172.71	167.68	180.04	171.42	168.80	163.23	178.48
G2 Winter	183.37	181.35	176.35	188.27	180.28	177.57	171.05	187.88
G2 Spring	190.17	187.75	183.19	194.59	186.09	183.00	177.10	193.36
G3 Fall	189.00	186.80	181.16	193.68	186.32	183.69	176.01	192.97
G3 Winter	196.95	194.86	188.85	200.97	193.66	191.07	183.68	199.86
G3 Spring	202.71	200.38	194.03	206.66	197.48	194.54	187.60	203.63
G4 Fall	201.33	199.37	191.73	205.37	197.43	195.05	186.64	202.86
G4 Winter	207.17	205.19	198.07	210.97	202.44	200.09	191.80	207.73
G4 Spring	213.37	211.15	203.74	216.72	205.60	203.24	195.56	210.44

Table A6. Monthly Growth Estimates for Students with All 15 Terms

	Math				Reading			
	Ever-EL	Always-EL	Dually-ID	Never-EL	Ever-EL	Always-EL	Dually-ID	Never-EL
Intercept	134.612***	128.133***	134.371***	144.945***	138.504***	131.975***	139.796***	145.447***
	(0.706)	(0.840)	(1.658)	(0.170)	(0.762)	(1.015)	(1.684)	(0.168)
K Year	3.059***	3.074***	2.747***	2.643***	1.963***	1.883***	1.714***	2.005***
	(0.025)	(0.035)	(0.065)	(0.011)	(0.028)	(0.040)	(0.074)	(0.012)
K Summer	-0.988***	-0.868***	-1.124***	-0.838***	-0.686***	-0.687***	-1.143***	-0.696***
	(0.069)	(0.089)	(0.187)	(0.027)	(0.081)	(0.110)	(0.229)	(0.031)
G1 Year	2.126***	2.081***	2.111***	2.055***	2.079***	1.977***	1.948***	2.070***
	(0.024)	(0.032)	(0.064)	(0.009)	(0.029)	(0.040)	(0.084)	(0.011)
G1 Summer	-1.174***	-1.192***	-1.563***	-1.427***	-0.488***	-0.523***	-1.396***	-0.598***
	(0.078)	(0.107)	(0.228)	(0.030)	(0.099)	(0.139)	(0.285)	(0.039)
G2 Year	1.724***	1.815***	1.821***	1.723***	1.688***	1.794***	1.696***	1.776***
	(0.022)	(0.031)	(0.060)	(0.010)	(0.027)	(0.041)	(0.075)	(0.012)
G2 Summer	-0.861***	-1.086***	-1.350***	-0.964***	-0.353***	-0.631***	-1.074***	-0.732***
	(0.063)	(0.085)	(0.191)	(0.026)	(0.082)	(0.110)	(0.224)	(0.034)
G3 Year	1.587***	1.663***	1.507***	1.591***	1.292***	1.376***	1.414***	1.255***
	(0.019)	(0.028)	(0.053)	(0.009)	(0.024)	(0.036)	(0.072)	(0.011)
G3 Summer	-1.287***	-1.352***	-1.613***	-1.317***	-0.595***	-0.670***	-1.194***	-0.784***
	(0.054)	(0.075)	(0.172)	(0.023)	(0.072)	(0.104)	(0.229)	(0.029)
G4 Year	1.507***	1.538***	1.479***	1.468***	1.025***	1.092***	1.116***	0.959***
	(0.018)	(0.026)	(0.052)	(0.008)	(0.022)	(0.035)	(0.066)	(0.010)
Female	-1.402***	-1.672***	-5.346***	-0.740***	1.498***	1.497**	-1.864	2.006***
	(0.403)	(0.533)	(1.221)	(0.180)	(0.450)	(0.608)	(1.244)	(0.205)
Asian	4.395***	8.202***	3.972*	4.785***	3.737***	9.019***	3.600	5.451***
	(0.862)	(1.057)	(2.303)	(0.541)	(0.928)	(1.225)	(2.386)	(0.569)
Black	-1.904	3.117*	-2.914	-7.863***	-1.214	4.468**	-5.569	-6.654***
	(1.316)	(1.688)	(3.862)	(0.243)	(1.464)	(2.093)	(3.601)	(0.282)
Hispanic	-4.532***	0.615	-6.997***	-4.903***	-4.931***	1.207	-7.168***	-5.189***
	(0.709)	(0.832)	(1.709)	(0.244)	(0.791)	(1.022)	(1.784)	(0.279)
Other Ethnicity	-0.127	2.365	-0.406	-3.196***	-1.256	4.989**	-1.428	-2.935***
	(0.942)	(1.944)	(2.811)	(0.360)	(1.035)	(2.277)	(2.873)	(0.393)
Tests	36975	19635	5985	190170	33045	16320	5310	178425
Students	2465	1309	399	12678	2203	1088	354	11895
Intercept-Variance	98.36	90.55	127.60	100.50	108.80	97.78	122.50	121.00

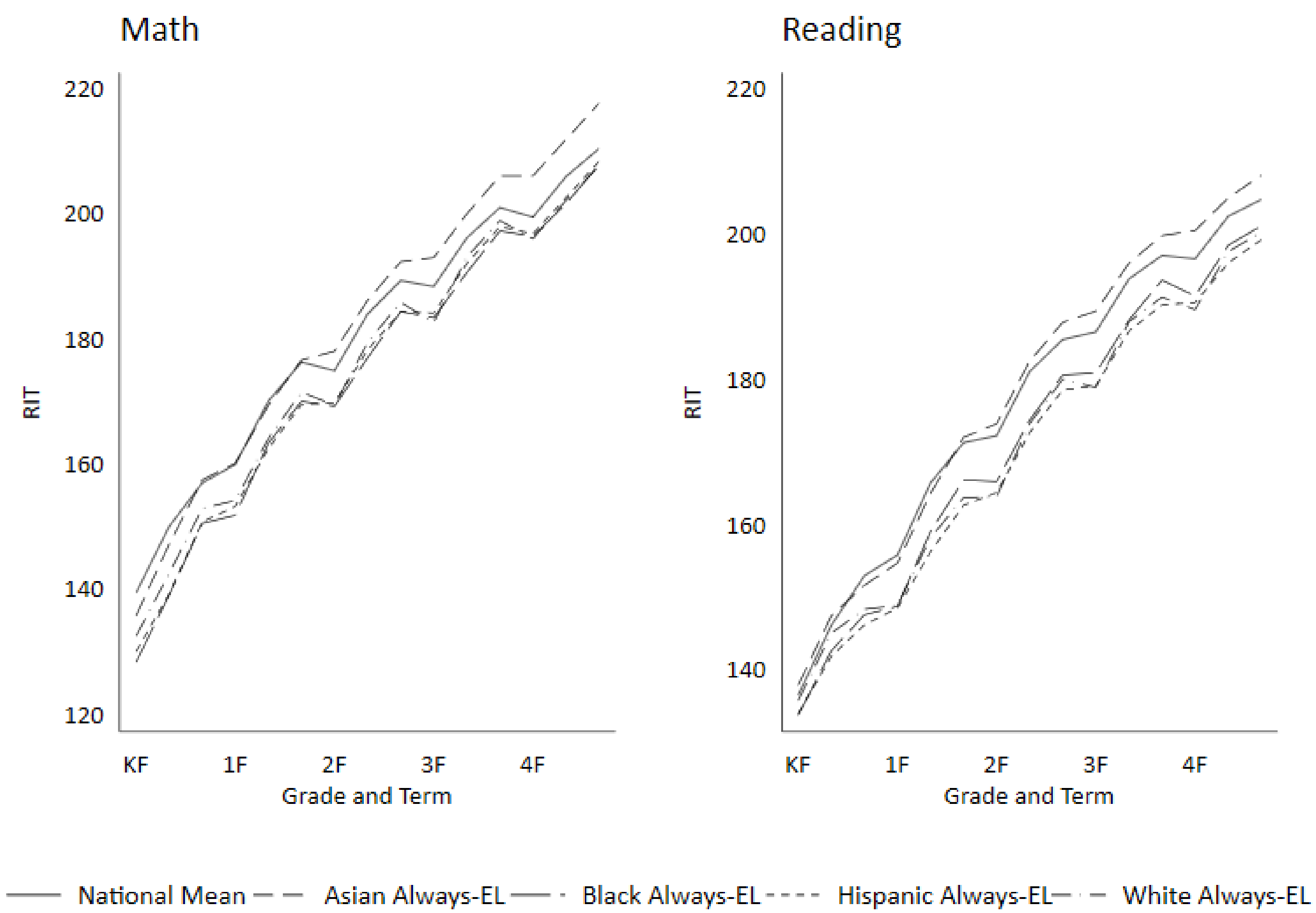
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column is a separate regression with sample restricted to the titular subgroup. K=kindergarten. G1= 1st grade. Estimates are from Model 2 described on page 14.

Figure A1. Math and Reading Achievement for Ever-EL Students by Race/Ethnicity



Notes: National mean are based on NWEA 2020 Norms (Thum & Kuhfeld, 2020). Sample includes students ever identified as English Learners at any time between kindergarten and 4th grade. KF = kindergarten fall. 1F = 1st grade fall.

Figure A2. Math and Reading Achievement for Always-EL Students by Race/Ethnicity



Notes: National mean are based on NWEA 2020 Norms (Thum & Kuhfeld, 2020). Sample includes students consistently identified as English Learners during every term tested. KF = kindergarten fall. 1F = 1st grade fall.

Appendix B. 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 1st grade, they have been exposed to all of kindergarten, a couple months of summer vacation after kindergarten, and one or two months of 1st grade. Since they have not been exposed to another summer vacation or 2nd grade, the values for those predictors are set to zero.

Table B1. Monthly Exposure Rates for a Hypothetical Student Testing in Kindergarten and 1st Grade

Grade/Term	School Start Date	School End Date	Test date	Monthly Exposure Design Matrix					
				Int.	K	SumK	G1	Sum1	G2
Fall K	8/20/2014	6/12/2015	9/1/2014	1.00	0.39	0.00	0.00	0.00	0.00
Winter K	8/20/2014	6/12/2015	12/1/2014	1.00	3.39	0.00	0.00	0.00	0.00
Spring K	8/20/2014	6/12/2015	5/1/2015	1.00	8.23	0.00	0.00	0.00	0.00
Fall 1st	8/19/2015	6/11/2016	9/15/2015	1.00	9.82	2.25	0.89	0.00	0.00
Winter 1st	8/19/2015	6/11/2016	11/20/2015	1.00	9.82	2.25	3.11	0.00	0.00
Spring 1st	8/19/2015	6/11/2016	4/1/2016	1.00	9.82	2.25	7.26	0.00	0.00