ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-337

Within-Year Achievement Gains for English Learners

Angela Johnson
NWEA

This study reports achievement levels and fall-to-spring gains in grades K to 8 for three groups of English Learners (ELs): (a) ever-ELs who were ever eligible for service; (b) current-ELs who continue to require service; and (b) dually-identified students eligible for both EL and Special Education services. I leverage unique data that include 186,139 ever-ELs and their never-EL peers in 1,520 schools. All three EL groups have lower mean math and reading achievement than the average of all students in kindergarten. Compared to the all-student average, ever-ELs and current-ELs make larger gains in the early grades but smaller gains in the middle grades. Dually-identified students make smaller gains than the all-student average across all grades in math and in kindergarten and 1st grade in reading. The expanding achievement gaps in the middle grades suggest better academic support is urgently needed for all multilingual students, especially dually-identified students.

VERSION: December 2020

Suggested citation: Johnson, Angela. (2020). Within-Year Achievement Gains for English Learners. (EdWorkingPaper: 20-337). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/w2bg-s488

Within-Year Achievement Gains for English Learners

English Learners (ELs) are one of the most vulnerable and least understood student populations in US public schools. 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). The vast majority of ELs come from low-income families and are enrolled in schools with high concentrations of ELs and other racial/ethnic minority students. Many are migrant and have experienced interrupted formal education prior to immigrating to the US; a significant fraction additionally require Special Education services.

Studies comparing the achievement of students currently receiving EL service and students in mainstream classes tended to show large gaps and highlight EL underachievement (e.g., Carnoy & Garcia, 2017; Hemphill & Vanneman, 2011; Polat et al., 2016). But recent research points out that every year new immigrant students are entering EL status and students with high English proficiency are exiting EL status; as a result, only contrasting the achievement of current ELs and native or fluent English users can lead to overestimation of achievement gaps and minimize the academic progress of multilingual students as a group (Hopkins et al., 2013; Kieffer & Thompson, 2018; Saunders & Marcelletti, 2013). Instead, we need to disaggregate data within the population of students who have ever been identified as ELs ("ever-EL"s) in order to see the full picture.

As shown in Figure 1 (and detailed in the next section), ever-ELs are multilingual students who were eligible to receive language services at some point, but 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 vastly between ELs who have gained sufficient academic English proficiency and exited EL status ("former-

ELs") and ELs who continue to require language support ("current-ELs"). In addition to language support, some ELs are identified by specialists as also requiring Special Education (SPED) services; these "dually-identified" students, an especially high-need population, is protected by multiple strands of federal education policy.

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. 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 intact cohorts of students over time to track their academic achievement (Johnson, 2020; Soland & Sandilos, 2020; Valentino & Reardon, 2015). These longitudinal studies tended to use data from a single school district, generating findings that were 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 provided evidence on the achievement of both ever-ELs and current-ELs; and (b) no large-scale research has investigated the academic achievement of dually-identified students.

This study estimates the fall-to-spring achievement gains in math and reading for ever-EL students and highlights two high-need subgroups within this population: (a) current-ELs who continue to require language support; and (b) dually-identified students who were eligible for both EL and SPED services at some point. I leverage a unique data set from the Growth Research Database (GRD) at NWEA that contains longitudinal achievement measures for

186,139 ever-ELs across 1,520 schools in 36 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, current-ELs, and dually-identified students) compare to the average student?
2. How do fall-to-spring achievement gains for these EL subgroups compare to those of an average student?

Background

EL Classification and Reclassification

Following the Civil Rights Act of 1964 and two landmark Supreme Court decisions (*Lau v. Nichols*, 1974; *Castañeda v. Pickard*, 1981), school districts are required to identify students who have not gained 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 and lower reading or ELA achievement. For this reason, recent literature on EL achievement highlights the need to examine the population of both 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 by the larger multilingual student population.

In the middle grades, ELs currently receiving service are comprised of (a) students who have risen from the lower grades but have not yet developed sufficient English proficiency and (b) students who recently arrived in the US. Within this current-EL population, educators are particularly concerned about long-term ELs, who have not developed fluent English proficiency after five or more years. Though national and state data are absent, estimates from large districts and reports from administrators reveal that a large fraction of long-term ELs are students additionally eligible for SPED services.

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 English language 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 most difficult to identify. Discerning between developing language proficiency and disability needs is a hard 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.

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 help and not hinder 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. Studies have examined national and local data to compare achievement for EL and non-EL students, yielding contrasting results.

Using National Assessment for Educational Progress (NAEP) data, research consistently reported large education achievement and attainment gaps between current-ELs and their non-ELs peers over the past two decades (e.g., Carnoy & Garcia, 2017; Fry, 2007; Polat et al., 2016). For example, Carnoy and Garcia (2017) reported that after adjusting for family socioeconomic status, the achievement gap between current-ELs and their peers remained larger than all other gaps observed between ethnicities and SES groups. Specifically, on the 8th grade math test, the gap between Hispanic ELs and white students was 0.9 SD larger than Hispanic non-ELs and white students; similarly, the white-Asian gap was 1.2 SD larger for ELs than non-ELs. In 8th grade reading, the white-Asian gap was 0.9 SD larger for ELs than non-ELs, and the white-Hispanic gaps was 0.8 SD larger for ELs than non-ELs.

However, the actual gap between students who do and do not use English at home may be smaller than these estimates suggest. This is because the composition of current-ELs presents challenges to analysis because the population of the group is unstable. That is, new students are constantly being classified as ELs and entering the group while existing ELs with higher proficiency are passing annual exams and exiting the group. Thus, at any point in time, students in the current-EL group only include those with lower English proficiency and, in contexts that require ELA scores for reclassification, lower academic achievement (Fry, 2007). Therefore,

current-ELs' academic performance provides valuable but incomplete information about the achievement of all ever-EL students. In fact, contrasting current-ELs' achievement with the pooled group of native and reclassified fluent English users can cause an overestimation of achievement gaps.

Indeed, Kieffer and Thompson (2018) demonstrate that NAEP score comparisons are substantially different when reclassified former-ELs are grouped with other initially classified ELs, rather than with never-ELs. That is, the gap between monolingual English users and multilingual students (which includes current- and former-ELs) appears to shrink over time for both 4th and 8th grade test takers in math and reading. Saunders and Marcelletti (2013), using state-level ELA achievement data from California, also show that looking only at current-EL data while excluding reclassified former-ELs fails to provide accurate estimates for the achievement level of all initially classified ELs and for the gaps between ELs and non-ELs. The authors of these papers urge future research to present data for all ever-ELs, as well as disaggregate analyses by subgroups so that policy and practice can target their specific needs.

While cross-sectional NAEP data allowed researchers to observe trends across cohorts of 4th and 8th graders, district administrative data have been used to track academic performance longitudinally from elementary to middle school. For instance, Valentino and Reardon (2015) followed students in a 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 (2020) 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 but lost less during the summers. 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.

Current Study

This study builds on Johnson (2020) to provide novel evidence on academic achievement and growth for ever-ELs and key EL subgroups. Using student characteristics and rich assessment data collected in the fall, winter, and spring of each year over five years, I estimate within-year academic gains for the pooled ever-EL sample and the current-EL and dually-identified subgroups. This study makes two main contributions. It is the first to provide within-year academic gain estimates for a large ever-EL sample comprised of students from across the nation. It is also the first to disaggregate findings for achievement and achievement gains and report estimates separately for dually-identified students.

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 can inform programs and policies at the school and district levels. 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 schools have on learning (Hegedus, 2018; Reardon, 2018). Because of structural inequalities faced by EL students and their families, ELs tend to enter school with lower achievement levels 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; examining growth, on the other hand, will reward schools and educators for the progress they are helping students make despite the disadvantages students and their communities face. Improving upon previous research in scale and richness of findings, this study greatly expands our understanding of academic achievement and development for one of the most disadvantaged student populations.

Data

The data for this study come from NWEA's GRD. The data fields include MAP Growth assessment scores and school-reported student demographic and EL and SPED service eligibility indicators. School districts choose to administer MAP Growth assessments for various purposes, including monitoring student achievement and growth, staff evaluation, and school

accountability. Data from the GRD are not nationally representative but cover a sizable portion (more than 20%) of the K-12 student population.

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 has provided complete data in these fields. Having verified the number of students eligible for EL and SPED services in the district against the Common Core of Data (CCD), I restrict my analysis to the districts that provided complete data.[1]

Sample

My sample includes more than 149,000 kindergarten to 8th grade students in 1,520 public schools in 213 districts across the nation in academic years 2014-15 to 2018-19. While the GRD includes private and international schools, I focus only on US 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 presents a comparison of summary statistics of the schools in my sample to all public schools serving grades between K and 8 in the 2016-17 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 White students.

Table 1 shows summary statistics for students in my sample. I include all students attending the 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. I describe the math sample below. Pooling students in all grades

between 2014-15 and 2018-19, the full sample is 49% female, 5% Asian, 27% Black, 19% Hispanic, and 41% White. There is little variation in demographic composition by grade or year.

[Table 1 here]

Students eligible for EL service in at least one year between 2014-15 and 2018-19 are categorized as "ever-ELs" in the data regardless of the timing of EL service.[2] Ever-ELs comprise 13% of the full sample. About 47% of the ever-ELs in my sample are girls, 13% are Asian, 8% Black, 61% Hispanic, and 12% White. Ever-ELs additionally eligible for SPED services at any time between 2014-15 and 2018-19 (dually-identified students) comprise 1.5% of the full sample. Only 35% of these dually-identified students were girls; Asian and Black students formed a smaller fraction of this subgroup compared to all ever-ELs, and White students formed a larger fraction. Students eligible for EL service during each grade are flagged as "current-ELs" for that grade. Current-ELs form a subgroup of ever-ELs who have not attained fluent English proficiency; their gender and racial/ethnic composition were similar to ever-ELs'. 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.

Measures of Achievement

Students were tested using the math and reading MAP Growth assessments up to three times (fall, winter, and spring) during the school year. The 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.[3]

Exposure to Instruction

MAP Growth tests are administered on dates determined by districts and schools. Test administration dates varies considerably between and even within school districts. To account for differences in testing schedules for academic gain analyses, I calculate the number of months each student has been in school on test date using the school start date for each year (months in school = (test date – start date) ÷30).

Analysis

Comparison of Achievement Levels

Pooling data for all five academic years,[4] I plot the mean achievement levels in the fall, winter, and spring of each grade for all students, ever-ELs, dually-identified students, and current-ELs. The plot also shows the national average from NWEA norms for comparison (Thum & Kuhfeld, 2020).

Spring Achievement Gaps

To understand achievement gaps between EL students and the average of all students, I calculate standardized mean difference in spring test scores between all students and the EL subgroups (ever-ELs, dually-identified, and current-ELs) using Cohen's *d*. The standardized mean difference between all students' and each subgroup' scores is

$$\frac{\overline{\mathrm{RIT}}_{Ag}-\overline{\mathrm{RIT}}_{Sg}}{\sqrt{\frac{(\mathrm{N_{Ag}}-1)\mathrm{SD}_{Ag}^2+(\mathrm{N}_{Sg}-1)\mathrm{SD}_{Sg}^2}{\mathrm{N}_{Ag}+\mathrm{N}_{\mathrm{S}g}-2}}}, \quad (1)$$

where, for every grade g, $\overline{\mathrm{RIT}}_{Ag}$ is the overall mean student test score in the spring, $\overline{\mathrm{RIT}}_{Sg}$ is the mean subgroup (ever-EL, dually-identified, or current-EL) test score in the spring, SD_{Ag} and SD_{Sg} are the standard deviations (SDs) for all students and the subgroup, and N_{Ag} and N_{Sg} are the corresponding number of students in each group. The mean, SD, and group sample size estimates used in these calculations are reported in Appendix Table A2.

Fall-to-Spring Achievement Gains

Then, I use fall and spring achievement scores to calculate standardized gain scores for each grade. I estimate the standardized mean difference effect sizes between fall and spring, first without accounting for repeated testing within student (Equation 2). The standardized gain score between fall and spring test scores is

$$\frac{\overline{\mathrm{RIT}}_{Sg}-\overline{\mathrm{RIT}}_{Fg}}{\sqrt{\frac{(\mathrm{N_{Sg}}-1)\mathrm{SD}_{Sg}^2+(\mathrm{N}_{Fg}-1)\mathrm{SD}_{Fg}^2}{\mathrm{N}_{Sg}+\mathrm{N_{F}}g-2}}}, \tag{2}$$

where, for every grade g, $\overline{\mathrm{RIT}}_{Sg}$ is the mean spring test score, $\overline{\mathrm{RIT}}_{Fg}$ is the mean fall test score, SD_{Sg} and SD_{Fg} are the SDs in the spring and fall, and N_{Sg} and N_{Fg} are the number of students. I do this for all students and then repeat for each EL subgroup separately. As a sensitivity check, I adjust the calculation to account for repeated testing.[5]

The standardized gain score approach described above does not account for variations in time elapsed between fall and spring testing due to differences in school testing schedules. I conduct an additional analysis that accounts for differences in time passed, calculating average monthly growth between the fall and spring as

$$\frac{\sum_{i=1}^{N_g}\frac{\mathrm{RIT}_{Si}-\mathrm{RIT}_{Fi}}{\mathrm{Mo}_{Si}-\mathrm{Mo}_{Fi}}}{N_g}, \tag{3}$$

where RIT_{Si} is student *i*'s spring test score, RIT_{Fi} is student *i*'s fall test score, Mo_{Si} is the number of months of instruction the student experienced before the spring test, Mo_{Fi} is the number of months of instruction before the fall test, and N_g is the number of unique students in grade g with both fall and spring test scores. Once again, I calculate this for all students, then separately for each EL subgroup.

Findings

Comparison of Achievement Levels

Figure 2 compares achievement means for all students in the sample to ever-ELs, dually-identified students, and current-ELs. The corresponding mean values are reported in Appendix Table A2. Compared to the national average in each spring, the average student in the sample had the same or very slightly higher achievement across grades and subjects.

For both math and reading, the between-group comparisons are similar. Ever-EL students start kindergarten with lower initial achievement than the average student. During each of the early grades, ever-ELs grow in parallel to the average student; but in the middle grades, ever-ELs show larger summer drops and flatter academic-year gains than the average student. As a result, the gap between ever-ELs and the average student widens from kindergarten to 8th grade.

Current-ELs' trajectories largely overlap with those of ever-ELs in the early grades. This is to be expected since the two groups are indistinguishable until students start to exit EL status in significant numbers, usually around 4th grade. Unsurprisingly, current-EL achievement starts to deviate from the ever-EL achievement in the middle grades, largely due to starting each fall with lower achievement. This could be because higher-achieving and high-English-proficiency students were exiting the current-EL group each spring, leaving lower-achieving, low-English-proficiency students to start in the current-EL group the following fall.

EL and SPED dually-identified students start kindergarten with achievement lower than the average student and slightly lower than ever-ELs. In math, dually-identified students follow flatter trajectories during each grade as well as larger summer drops, resulting in a wide achievement gap between dually-identified students and the other groups by the end of 8th grade. In reading, slower growth and larger summer drops led to widening gaps in kindergarten to 5th grade between dually-identified students and the ever-EL group as a whole; however, dually-

identified students experienced smaller summer losses compared to ever-ELs between 5th and 8th grade, reducing the gap between the two. Between dually-identified students and the average student in the same schools, the reading achievement gap consistently widened in each grade between K and 8.

There is a striking drop in the summer between 5th and 6th grade, for all students but especially for EL students. The drop is visible for math and reading, though more dramatic for math.

[Figure 2 here]

Spring Achievement Gaps

Table 2 presents the spring achievement gaps between all students and ever-ELs, dually-identified students, and current-ELs. All gaps are statistically significant. In kindergarten, the math achievement gap was .43 SD between all students and ever-ELs, .42 SD between all students and current-ELs, and .80 SD between all students and dually-identified students. For both ever-ELs and current-ELs, the achievement gap shrank between kindergarten and 3rd grade and widened between 3rd and 8th grade, such that by the end of 8th grade, the gaps were .64 SD and .82 SD, respectively. For dually-identified students, the math achievement gap stayed .80 SD until 2nd grade and widened in each subsequent year to 1.10 SD in 8th grade.

For reading, gaps started at .49 SD in kindergarten for both ever-ELs and current-ELs, fluctuated or shrank slightly in the first few years, then grew wider starting in 4th grade. For dually-identified students, the reading gap started at .74 SD in kindergarten and grew consistently in every grade. By the end of 8th grade, the reading gap was .78 SD between all students and ever-ELs, 1.02 SD for current-ELs, and 1.20 SD for dually-identified students. These estimates are consistent with the achievement trajectories presented in Figure 2.

[Table 2 here]

Fall-to-Spring Achievement Gains

Table 3 shows the fall-to-spring academic gains for all students and each EL subgroup, which are consistent with the results in Table 2. Ever-ELs and current-ELs experienced significantly larger academic gains than all students from kindergarten to 3rd grade in math and reading. For example, in kindergarten, all students gained 1.54 SD in math, while ever-ELs gained 1.66 SD, and current-ELs gained 1.67 SD. But ever-ELs and current-ELs had significantly smaller gains than all students in math starting in 5th grade and statistically similar gains as all students in reading in the middle grades. In math, dually-identified students experienced significantly smaller gains compared to all students in across grades. In reading, dually-identified students gained significantly less than all students in kindergarten and 1st grade and had statistically-similar (and some numerically larger) gains in 2nd to 8th grade. Effect sizes estimated after adjusting for repeated testing and estimated monthly growth rates (Appendix Tables A3 and A4) show qualitatively similar patterns. In math, ever-ELs and current-ELs grow more during each of the earlier grades and less during grades 6 and 7 than all students; dually-identified students grow less than all students in every grade. In reading, ever-ELs and current-ELs grow at similar or faster rates than all students in most grades, and dually-identified students grew less in the lower grades but grew more in the middle grades than all students.

[Table 3 here]

Discussion

This study estimates within-year academic gains for a unique sample of EL students across the nation. I report three main findings. First, in each grade, large fall achievement status gaps exist between ELs and the average grade-level peer. Second, compared to the average

student, ever-ELs and current-ELs experience larger within-year gains in grades K to 4 in both math and reading but less gains in math and similar gains in reading in the middle grades. Third, dually-identified students experience smaller gains than the average student in every grade in math and in the early grades in reading. These patterns explain the larger achievement gaps in both subjects between ELs and the average student at the end of 8th grade.

My results add to the body of recent evidence (e.g., Kieffer & Thompson, 2018; Valentino & Reardon, 2015) affirming ELs' academic 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 within-year estimates for the early grades complement their findings. Challenging the deficit narrative portraying ELs as underachieving, my results show ELs are capable of *growing faster* during the school year than the average student.

That achievement gaps widen between 5th and 8th grade for ever-ELs as a group highlights the need for close monitoring in these middle grades and better specialized support. The drop in the summer between 5th and 6th grade (i.e., from elementary to middle school for many students) is especially alarming. This sizable drop and the subsequent expansion of the achievement gap through 8th grade, confirm earlier research evidence suggesting that historically disadvantaged student subgroups are more vulnerable to achievement decline and require more support than other students during the transition to middle school (e.g., Eccles & Roeser, 2004; Schwerdt & West, 2013). In this case, the growing gap in the middle grades applies to the whole ever-EL group, which includes current-ELs as well as students who have already gained fluent English proficiency in earlier grades. Thus, the 5th grade summer drop and the widening achievement gap may not simply be a result of higher-achieving students' leaving the subgroup,

but a larger issue facing all ELs. In contrast to Kieffer and Thompson's (2018) finding that gaps close between multilingual and monolingual students, my data show that recently reclassified students, as well as current-ELs, would benefit from better support during the critical transition to the middle grades.

Prior work on ELs' middle school academic access shed light on potential factors that contribute to the growing achievement gap, including limited course access, low teacher expectations, and insufficient support. Current-EL status has been shown to preclude students from taking a full load of academic content courses in middle school and from taking upper level classes in content subjects (Umansky, 2016b, 2018). Former ELs may continue to face low academic expectations, even after reclassification, from teachers and counselors who want to protect the students from difficult academic materials. Alternatively, some former ELs may struggle to grow academically if they were reclassified too soon. Reclassification criteria set at the state or district level may be higher or lower than optimal; when the criteria are set too low, students can be reclassified before they are ready to succeed without linguistic support (Robinson, 2011). It is possible that premature removal of language support is contributing to some of the academic drop experienced by ever-ELs. My data did not allow me to identify how reclassification contributed to ELs' achievement and growth patterns, but recommendations from existing research are worth consideration. To ensure that ever-ELs make progress toward mastery of grade-level content and academic English proficiency, schools need to provide them with rigorous curriculum, high-quality instruction, and unrestricted access to peers who are native or fluent English users, both before and after reclassification (Johnson & Goldenberg, 2020). Providing these critical support elements in the early grades will help ELs develop English proficiency and prevent students from entering long-term (5 years +) EL status, which is

negatively associated with self-efficacy and expectations, as well as academic outcomes (Flores et al., 2015; Freeman et al., 2003; Johnson, 2019; Kim & Garcia, 2014).

This study also reveals that dually-identified students need better support across the elementary and middle grades. It is not surprising that students whose home language is not English and has an identified or unknown disability would arrive in kindergarten with lower levels of achievement than their peers. But their within-year growth rates, consistently lower than non-ELs and ELs across all grades for math and early grades for reading, suggest that the services students receive are not meeting their needs. A first step to addressing this issue is investigating the EL and SPED identification processes, which are crucial to providing students with appropriate support. In the early grades, it is often difficult to distinguish between low language proficiency and a disability. As a result, districts may choose to prioritize classifying a student for one type of service over the other. For example, some districts default to classifying multilingual students for EL service when they enter the district and postpone assessment for SPED service needs until 2nd or 3rd grade. This may be a practical option, given the challenges in distinguishing the two types of needs and limited district resources. However, delayed identification can result in delayed provision of much-needed services and cost the student crucial opportunities for development and learning.

In order to address this critical issue facing students with special needs, researchers and practitioners must collaborate to create a system for monitoring student growth. Currently, our understanding of achievement and growth for students who are eligible for SPED services is extremely limited. We do not know 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 large sample of ever-ELs from across the nation, this study provides the first estimates for within-year academic gains. My findings suggest that ELs, especially dually-identified students and students in the middle grades, need improved support programs and practices. Linguistic as well as academic services are necessary to ensure that ELs making progress toward academic English proficiency and content knowledge. Monitoring achievement growth is a crucial aspect of fostering academic success. However, simply maintaining pace with non-EL peers, without acquiring grade-level content skills, is not sufficient. In order to close the achievement gaps in the long run, schools must provide support for ELs to make larger gains in every grade and address any loss during out-of-school time. Most importantly, educators have to set high expectations for students who enter school with additional social and linguistic assets. This study demonstrates that ELs are capable of making greater academic growth than their peers and of making progress toward closing achievement gaps. 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. 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 closing achievement gaps and supporting disadvantaged students. 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 program information. EL and SPED service eligibility flags were only available for 2014-15 to 2018-19. There may have been students who received and then exited services before entering the data set and as a result were not classified as ever-EL in the analysis. To the extent that early-reclassification students are high-achieving, this may have resulted in overestimated achievement gaps. Though prior research has shown that ELs' academic achievement and paths to English proficiency may differ by demographic characteristics and language program, I was not able to explore heterogeneity along these dimensions. I observe language program only for a small number of ELs and did not have student-level variables for SES. Although student ethnicity was available, I refrained from conducting heterogeneity analysis by ethnicity because the relation between ethnicity and achievement likely confounds SES effects.

These limitations are not unique to this study and reflect larger issues in EL data reporting. 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. Future research should aim to establish a national initiative to collect and curate actionable achievement 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 Common Core of Data 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] Students who gained fluent English proficiency before 2014-15 or before their district first administered the MAP Growth assessments between 2014-15 and 2018-19 are not flagged as ever-ELs. As such, this categorization may be slightly underclassifying ever-ELs.

[3] Average test duration was slightly longer for ELs compared to all students. Standard error of measurement and percentage of rapidly-guessed items were similar for ELs and all students.

[4] Test scores varied very little from year to year between 2014-15 and 2018-19. Score summary statistics by year are available upon request.

[5] To account for repeated testing, I apply the following formula to students with both spring and fall test scores and calculate the average difference score divided by the standard deviation of the difference scores (Borenstein et al., 2009):

$$\frac{\overline{\text{RIT}}_{Sg} - \overline{\text{RIT}}_{Fg}}{\frac{SD_{RITs-RITf}}{\sqrt{2(1 - r_{RITs,RITf})}}} \quad (4)$$

where, for every grade g, $\overline{\text{RIT}}_{Sg}$ is still the mean spring test score, $\overline{\text{RIT}}_{Fg}$ is still the mean fall test score, $SD_{RITs-RITf}$ is the SD of the difference in spring and fall test scores for individual students, and $r_{RITs,RITf}$ is the correlation between spring and fall test scores.

References

Borenstein, M., Hedges, L.V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley.

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).

Eccles, J.S. & Roeser, R. (2004). Schools, academic motivation, and stage-environment fit. In: Lerner, RM.; Steinberg, L. (eds) *Handbook of Adolescent Psychology*, (pp. 125-153). John Wiley and Sons.

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, and Education, 14,* 113-132.

Freeman, Y. S., Freeman, D. E., & Mercuri, S. (2002). *Closing the achievement gap: How to reach limited-formal-schooling and long-term English learners.* Portsmouth, NH: Heinemann.

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

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

Hakuta, K., Butler, Y., & Witt, D. (2000). How long does it take English learners to attain proficiency? *Policy Report 2000-1*. The University of California Linguistic Minority Research Institute.

Hegedus, A. (2018). *Evaluating the relationships between poverty and school performance.* https://www.nwea.org/research/publication/evaluating-the-relationships-between-poverty-and-school-performance-whitepaper/

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

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.

Kim, W. G. & Garcia, S. (2014). Long-Term English Language Learners' perceptions of their language and academic learning experiences. *Remedial and Special Education*, *35*(5), 300-312.

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.

National Center for Education Statistics (NCES) (2019). https://nces.ed.gov/programs/coe/indicator_cgf.asp

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

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,

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

Saunders, W. M. & Marcelletti, D. J. (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.

Schwerdt, G., & West, M. R. (2013). The impact of alternative grade configurations on student outcomes through middle and high school. *Journal of Public Economics, 97*(3), 308–326.

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.

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.

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

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.

Tables and Figures

Figure 1. English Learner (EL) Subgroups

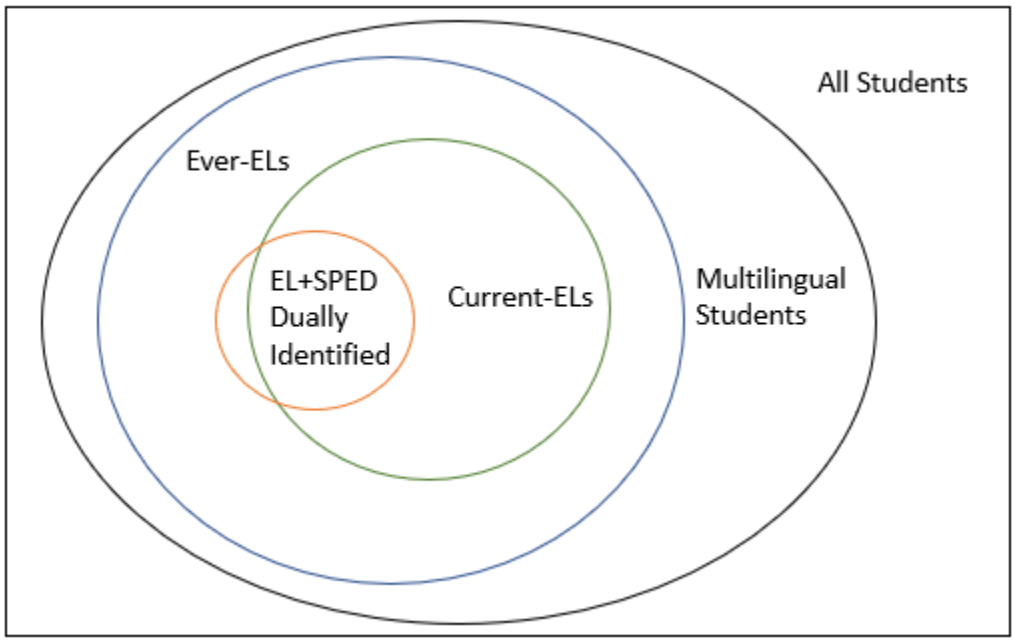

Figure 2. Achievement by EL Subgroup Status

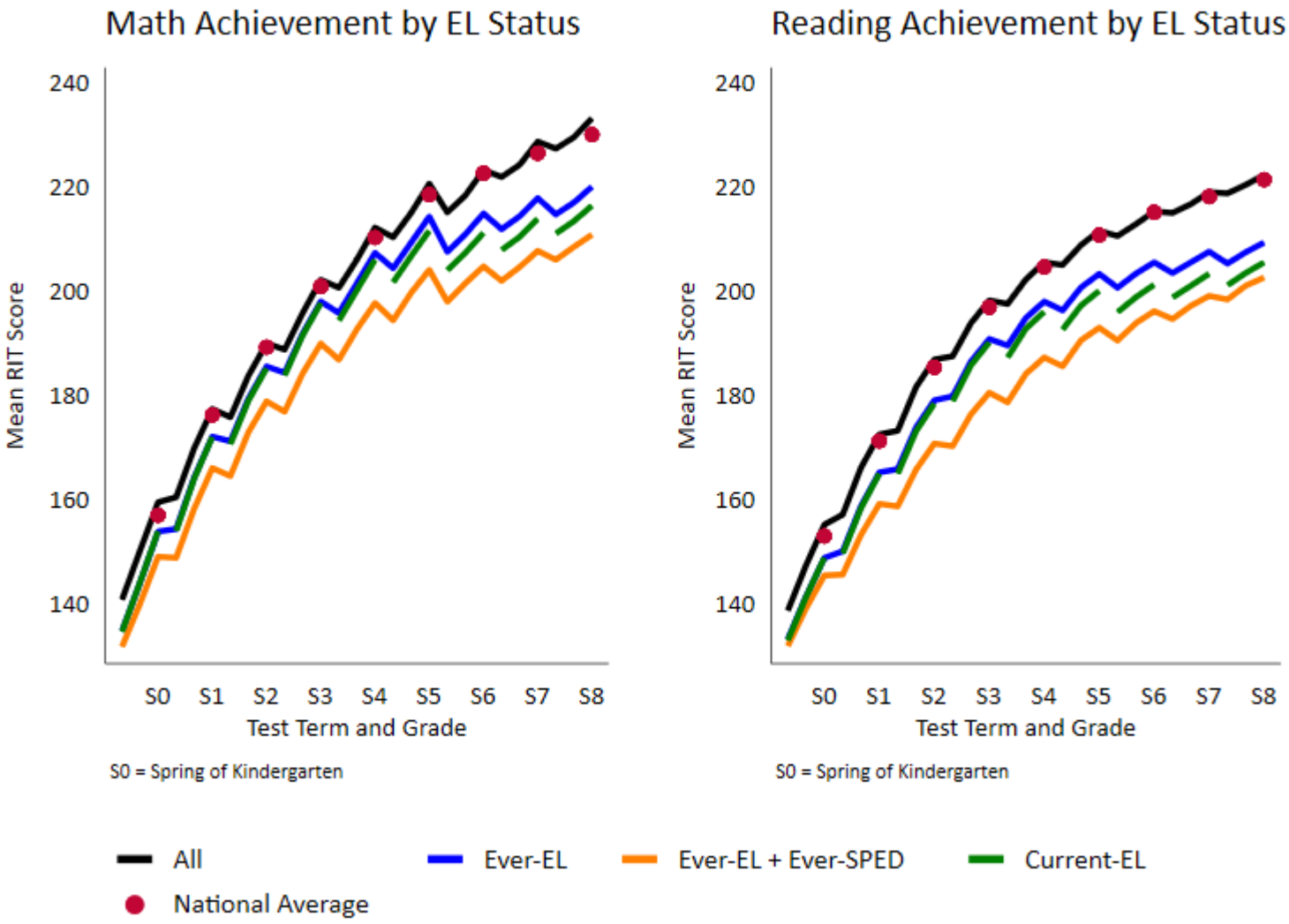

Table 1. Sample Summary Statistics

	all (N=1,353,500)		ever-EL (N=185,608)		ever-EL+SPED (N=21,256)		current-EL (N=148,576)	
Math	mean	sd	mean	sd	mean	sd	mean	sd
All Grades								
Female	0.49	0.50	0.47	0.50	0.35	0.48	0.47	0.50
Asian	0.05	0.22	0.13	0.34	0.10	0.30	0.13	0.34
Black	0.26	0.44	0.07	0.26	0.06	0.23	0.08	0.27
Hispanic	0.20	0.40	0.62	0.49	0.63	0.48	0.62	0.48
White	0.41	0.49	0.12	0.32	0.13	0.34	0.11	0.31
% ELs by								
Grade	total students		% ever-EL		% ever-EL + SPED		% current-EL	
K	272740		14.8%		1.5%		13.1%	
1	338498		16.2%		1.9%		13.7%	
2	368838		17.9%		2.1%		14.7%	
3	377629		17.6%		2.1%		13.9%	
4	377961		16.7%		2.1%		11.9%	
5	377865		15.4%		2.1%		9.9%	
6	359668		13.5%		1.9%		8.0%	
7	349123		11.5%		1.7%		7.0%	
8	338425		9.9%		1.5%		6.6%	
Total	1353500		13.7%		1.6%		11.0%	

	all (N=1,344,071)		ever-EL (N=181,834)		ever-EL+SPED (N=21,015)		current-EL (N=144,910)	
Reading	mean	sd	mean	sd	mean	sd	mean	sd
All Grades								
Female	0.49	0.50	0.47	0.50	0.35	0.48	0.47	0.50
Asian	0.05	0.22	0.13	0.34	0.10	0.30	0.13	0.34
Black	0.26	0.44	0.07	0.26	0.06	0.24	0.08	0.26
Hispanic	0.20	0.40	0.61	0.49	0.63	0.48	0.61	0.49
White	0.41	0.49	0.12	0.32	0.14	0.34	0.11	0.32
% ELs by								
Grade	total students		% ever-EL		% ever-EL + SPED		% current-EL	
K	261484		13.8%		1.4%		12.1%	
1	326703		15.4%		1.8%		12.9%	
2	359700		17.3%		2.1%		14.1%	
3	370660		17.1%		2.1%		13.5%	
4	371139		16.3%		2.1%		11.5%	
5	371514		15.1%		2.0%		9.6%	
6	352349		13.4%		1.9%		8.0%	
7	340495		11.4%		1.7%		6.9%	
8	331261		9.9%		1.5%		6.5%	
Total	1344071		13.5%		1.6%		10.8%	

* Note: Table presents fall test student counts. Total is not sum of K-8. Some students are represented in multiple grades. In the full sample, 30% of students had 1 year of data; 22% had 2 years; 18% had 3 years; 16% had 4 years, and 13% had 5 years of data. Among ever-ELs, 23% had 1 year of data; 22% had 2 years; 20% had 3 years; 19% had 4 years; and 16% had 5 years of data.

Table 2. Standardized Spring Achievement Gaps

	All - Ever-EL		All - Ever-EL+SPED		All - Current-EL	
	Gap (SD)	SE	Gap (SD)	SE	Gap (SD)	SE
Math						
K	0.43	0.00	0.80	0.01	0.42	0.01
1	0.37	0.00	0.80	0.01	0.38	0.00
2	0.32	0.00	0.80	0.01	0.34	0.00
3	0.28	0.00	0.83	0.01	0.30	0.00
4	0.30	0.00	0.89	0.01	0.38	0.00
5	0.35	0.00	0.92	0.01	0.50	0.01
6	0.46	0.00	1.03	0.01	0.67	0.01
7	0.56	0.01	1.09	0.01	0.77	0.01
8	0.64	0.01	1.10	0.01	0.82	0.01
Reading						
K	0.49	0.01	0.74	0.01	0.49	0.01
1	0.49	0.00	0.89	0.01	0.51	0.01
2	0.47	0.00	0.96	0.01	0.50	0.00
3	0.44	0.00	1.06	0.01	0.48	0.00
4	0.46	0.00	1.11	0.01	0.58	0.01
5	0.51	0.00	1.15	0.01	0.71	0.01
6	0.60	0.00	1.18	0.01	0.86	0.01
7	0.69	0.01	1.21	0.01	0.95	0.01
8	0.78	0.01	1.20	0.01	1.02	0.01

Notes: Achievement gaps calculated using Formula (1). SD = standard deviation. SE = standard error. “All” refers to all students in the sample. Ever-EL refers to students ever flagged as “EL” between 2014-15 and 2018-19. Ever-EL+SPED refers to dually-identified students who were eligible to receive EL and ever eligible to receive SPED services between 2014-15 and 2018-19. Current-ELs are students eligible to receive EL service during each grade. All estimated gaps are statistically significant.

Table 3. Fall-to-Spring Achievement Gain Estimates

	All		Ever-EL		Ever-EL+SPED		Current-EL	
	Effect Size	SE	Effect Size	SE	Effect Size	SE	Effect Size	SE
Math								
K	1.54	0.00	**1.66**	0.01	**1.42**	0.02	**1.67**	0.01
1	1.23	0.00	**1.31**	0.01	**1.18**	0.02	**1.34**	0.01
2	1.01	0.00	**1.07**	0.01	**0.98**	0.02	**1.10**	0.01
3	0.94	0.00	**0.98**	0.01	**0.85**	0.02	**1.01**	0.01
4	0.75	0.00	0.76	0.01	**0.65**	0.02	**0.77**	0.01
5	0.60	0.00	**0.58**	0.01	**0.54**	0.02	**0.58**	0.01
6	0.48	0.00	**0.43**	0.01	**0.39**	0.02	**0.43**	0.01
7	0.37	0.00	**0.32**	0.01	**0.31**	0.02	**0.33**	0.01
8	0.30	0.00	**0.27**	0.01	0.26	0.02	0.28	0.01
Reading								
K	1.40	0.00	**1.44**	0.01	**1.22**	0.02	**1.47**	0.01
1	1.07	0.00	**1.12**	0.01	**1.00**	0.02	**1.14**	0.01
2	0.81	0.00	**0.85**	0.01	0.78	0.02	**0.88**	0.01
3	0.63	0.00	**0.67**	0.01	0.60	0.02	**0.70**	0.01
4	0.48	0.00	**0.50**	0.01	0.48	0.02	**0.53**	0.01
5	0.40	0.00	0.41	0.01	0.41	0.02	**0.44**	0.01
6	0.29	0.00	0.28	0.01	0.31	0.02	0.30	0.01
7	0.24	0.00	**0.23**	0.01	0.25	0.02	0.25	0.01
8	0.22	0.00	0.22	0.01	0.24	0.02	**0.24**	0.01

Notes: Achievement gain estimates calculated using Formula (2). Effect sizes expressed in standard deviations. SE = standard error. "All" refers to all students in the sample. Ever-EL refers to students ever flagged as "EL" between 2014-15 and 2018-19. Ever-EL+SPED refers to dually-identified students who were eligible to receive EL and ever eligible to receive SPED services between 2014-15 and 2018-19. Current-ELs are students eligible to receive EL service during each grade. All estimated gaps are statistically significant. Estimates that are significantly different from the corresponding "All" students estimate are boldfaced.

Appendix Tables (Supplemental Materials)

Table A1. Comparison of Sample and All Public School Characteristics

	Sample Schools			All NCES Schools		
	Mean	SD	Schools	Mean	SD	Schools
Grade 1						
% FRPL	0.59	0.29	1114	0.55	0.30	54426
% Asian	0.04	0.07	1114	0.04	0.09	54590
% Black	0.27	0.31	1114	0.15	0.24	54590
% Hispanic	0.22	0.24	1114	0.25	0.29	54590
% White	0.42	0.31	1114	0.50	0.34	54590
City	0.45	0.50	1114	0.30	0.46	55038
Town	0.08	0.27	1114	0.11	0.31	55038
Rural	0.12	0.33	1114	0.26	0.44	55038
Title I Eligible	0.83	0.38	1114	0.79	0.41	54614
School-wide Title I	0.71	0.45	1111	0.67	0.47	54313
Grade 5						
% FRPL	0.59	0.30	1143	0.56	0.30	53140
% Asian	0.04	0.07	1143	0.04	0.09	53315
% Black	0.27	0.31	1143	0.15	0.24	53315
% Hispanic	0.22	0.24	1143	0.25	0.29	53315
% White	0.42	0.32	1143	0.49	0.34	53315
City	0.45	0.50	1143	0.30	0.46	53867
Town	0.09	0.28	1143	0.10	0.30	53867
Rural	0.12	0.33	1143	0.26	0.44	53867
Title I Eligible	0.82	0.38	1143	0.77	0.42	53410
School-wide Title I	0.71	0.45	1140	0.66	0.47	53085
Grade 8						
% FRPL	0.58	0.28	554	0.54	0.29	32408
% Asian	0.04	0.07	554	0.03	0.07	32598
% Black	0.30	0.34	554	0.16	0.26	32598
% Hispanic	0.19	0.24	554	0.22	0.27	32598
% White	0.43	0.33	554	0.52	0.35	32598
City	0.44	0.50	554	0.27	0.45	33412
Town	0.12	0.32	554	0.13	0.33	33412
Rural	0.16	0.37	554	0.33	0.47	33412
Title I Eligible	0.76	0.43	554	0.70	0.46	32946
School-wide Title I	0.64	0.48	553	0.58	0.49	32564

Notes: Grades 1, 5, and 8 selected to represent the early and middle grades. Characteristics for other grades available upon request. SD = standard deviation. FRPL = free or reduced-price lunch.

Table A2. Achievement by Grade and Term

	all				ever-EL				ever-EL+SPED				current-EL			
RIT by Grade/Term	mean	sd	median pctile	N	mean	sd	median pctile	N	mean	sd	median pctile	N	mean	sd	median pctile	N
Math																
KG Fall	140.8	11.0	52	272740	134.8	9.8	34	40269	131.8	10.2	25	4099	134.7	9.8	34	35690
KG Spring	159.6	13.2	63	318327	154.0	12.9	42	47940	149.1	13.7	26	5068	154.0	12.9	43	41922
G1 Fall	160.6	13.4	56	338498	154.5	13.0	36	54782	148.9	13.7	19	6417	154.1	12.9	35	46328
G1 Spring	177.5	14.2	55	348063	172.2	14.1	41	57246	166.2	15.5	25	6670	172.2	14.1	41	48659
G2 Fall	176.0	14.0	56	368838	171.3	13.2	42	65853	164.6	14.2	23	7914	170.7	13.0	40	54037
G2 Spring	190.1	14.0	56	374296	185.7	13.7	42	67325	179.0	15.2	23	7971	185.4	13.6	41	55283
G3 Fall	189.0	13.9	55	377629	184.5	13.4	42	66291	177.0	14.6	21	8095	184.0	13.3	40	52588
G3 Spring	202.4	14.7	58	380780	198.2	14.4	46	68360	190.1	16.1	23	8217	197.9	14.2	46	54073
G4 Fall	200.9	14.7	58	377961	196.0	14.6	44	63050	187.0	15.9	21	7980	194.5	14.5	40	45002
G4 Spring	212.4	16.2	59	381604	207.6	16.0	46	64461	197.9	17.4	22	7996	206.3	15.9	43	46435
G5 Fall	210.6	16.3	58	377865	204.6	16.3	42	58081	194.6	17.1	18	7769	201.9	16.3	35	37421
G5 Spring	220.9	18.0	58	377294	214.5	17.8	43	58101	204.3	18.6	19	7699	211.8	17.7	37	37900
G6 Fall	215.4	16.2	55	359668	207.8	16.1	37	48386	198.2	16.6	16	6948	204.2	15.9	29	28888
G6 Spring	223.5	18.0	55	354685	215.1	17.8	35	48668	204.9	18.2	15	6832	211.4	17.6	28	29560
G7 Fall	222.2	18.0	58	349123	212.1	18.0	35	40026	202.2	18.2	15	5842	208.1	17.5	26	24525
G7 Spring	229.0	19.3	58	344419	218.1	19.2	34	40925	207.9	19.1	15	5796	214.1	18.8	27	25589
G8 Fall	227.6	19.1	58	338425	215.0	19.0	32	33593	206.2	18.3	16	4973	211.3	18.5	25	22339
G8 Spring	233.4	20.4	58	330524	220.3	20.1	32	34046	211.1	19.4	17	4821	216.7	19.7	27	23181
Reading																
KG Fall	138.7	10.1	56	261484	133.2	9.3	40	36125	131.9	9.7	35	3732	133.0	9.3	39	31590
KG Spring	155.3	13.2	58	304955	148.9	12.1	36	43498	145.5	12.2	25	4693	149.0	12.1	37	37805
G1 Fall	157.2	13.8	55	326703	150.2	12.6	33	50240	145.7	12.4	21	5951	149.7	12.6	32	42060
G1 Spring	172.6	14.9	55	336431	165.3	14.4	34	52707	159.3	14.7	18	6241	165.1	14.4	34	44412
G2 Fall	173.4	17.0	52	359700	166.0	15.1	32	62189	158.8	14.2	15	7531	165.1	14.7	30	50609
G2 Spring	187.0	16.9	58	363843	179.2	16.2	36	63500	170.9	16.5	15	7587	178.6	16.0	35	51777
G3 Fall	187.7	17.3	57	370660	180.0	16.3	38	63414	170.4	16.3	14	7879	179.0	16.0	36	50169
G3 Spring	198.4	16.7	59	372662	191.0	16.7	40	65199	180.7	17.8	16	7970	190.4	16.5	39	51322
G4 Fall	197.7	17.0	58	371139	189.7	16.8	39	60678	178.8	17.5	14	7765	187.4	16.6	34	42829
G4 Spring	205.8	16.5	58	375046	198.2	16.8	40	61996	187.4	18.2	16	7765	196.2	16.7	36	44153
G5 Fall	205.2	16.7	57	371514	196.5	17.2	37	56238	185.7	18.0	14	7592	192.8	17.1	29	35780
G5 Spring	211.8	16.2	57	372445	203.5	17.0	38	56479	193.2	18.1	15	7568	200.3	17.1	31	36347
G6 Fall	210.8	16.5	57	352349	200.9	17.6	35	47362	190.7	18.1	13	6833	196.2	17.5	25	28210
G6 Spring	215.5	16.3	56	348080	205.7	17.5	33	47457	196.3	18.0	13	6671	201.4	17.5	25	28700
G7 Fall	215.3	16.6	58	340495	203.7	18.1	32	38861	194.8	18.1	14	5695	199.0	18.1	23	23632
G7 Spring	219.2	16.4	57	335952	207.8	18.2	32	39584	199.3	17.7	14	5620	203.5	18.4	24	24540

G8 Fall	219.0	16.5	57	331261	205.5	18.8	29	32642	198.6	17.6	15	4851	201.3	18.7	22	21553
G8 Spring	222.6	16.4	57	324152	209.5	18.6	30	33106	202.8	17.3	15	4721	205.7	18.8	23	22500

Table A3. Fall-to-Spring Achievement Gain Estimates, Adjusted for Repeated Testing

	All		Ever-EL		Ever-EL+SPED		Current-EL	
	Effect Size	SE	Effect Size	SE	Effect Size	SE	Effect Size	SE
Math								
K	1.55	0.00	**1.70**	0.01	**1.41**	0.02	**1.72**	0.01
1	1.24	0.00	**1.36**	0.00	**1.17**	0.01	**1.38**	0.00
2	1.02	0.00	**1.12**	0.00	**1.00**	0.01	**1.15**	0.00
3	0.94	0.00	**1.02**	0.00	**0.86**	0.01	**1.06**	0.00
4	0.74	0.00	**0.79**	0.00	**0.66**	0.01	**0.81**	0.00
5	0.59	0.00	**0.62**	0.00	**0.54**	0.01	**0.63**	0.00
6	0.46	0.00	**0.45**	0.00	**0.38**	0.01	**0.45**	0.00
7	0.35	0.00	0.35	0.00	**0.31**	0.01	0.35	0.00
8	0.28	0.00	**0.29**	0.00	**0.25**	0.01	**0.31**	0.00
Reading								
K	1.40	0.00	**1.51**	0.01	**1.26**	0.02	**1.54**	0.01
1	1.08	0.00	**1.16**	0.00	**0.99**	0.01	**1.18**	0.00
2	0.81	0.00	**0.89**	0.00	0.80	0.01	**0.92**	0.00
3	0.63	0.00	**0.71**	0.00	**0.61**	0.01	**0.74**	0.00
4	0.48	0.00	**0.54**	0.00	**0.49**	0.01	**0.57**	0.00
5	0.40	0.00	**0.45**	0.00	**0.42**	0.01	**0.49**	0.00
6	0.28	0.00	**0.31**	0.00	**0.30**	0.01	**0.33**	0.00
7	0.24	0.00	**0.27**	0.00	0.25	0.01	**0.29**	0.00
8	0.21	0.00	**0.26**	0.00	**0.24**	0.01	**0.28**	0.00

Notes: Achievement gain estimates calculated using Formula (4). Effect sizes expressed in standard deviations. SE = standard error. "All" refers to all students in the sample. Ever-EL refers to students ever flagged as "EL" between 2014-15 and 2018-19. Ever-EL+SPED refers to dually-identified students who were eligible to receive EL and ever eligible to receive SPED services between 2014-15 and 2018-19. Current-ELs are students eligible to receive EL service during each grade. All estimated gaps are statistically significant. Estimates that are significantly different from the corresponding "All" students estimate are boldfaced.

Table A4. Monthly Growth Rates (RITs) Between Fall and Spring, Adjusting for Days Elapsed

	All		Ever-EL		Ever-EL+SPED		Current-EL	
	G.Rate	SD	G.Rate	SD	G.Rate	SD	G.Rate	SD
Math								
K	2.64	1.33	2.70	1.32	2.37	1.35	2.71	1.29
1	2.28	1.17	2.42	1.22	2.27	1.31	2.43	1.21
2	1.90	1.13	1.97	1.14	1.91	1.23	1.99	1.12
3	1.78	1.05	1.87	1.08	1.74	1.22	1.90	1.07
4	1.53	1.05	1.58	1.07	1.45	1.23	1.58	1.08
5	1.37	1.08	1.37	1.11	1.27	1.22	1.37	1.13
6	1.05	1.03	1.00	1.09	0.87	1.21	0.97	1.15
7	0.87	1.04	0.84	1.15	0.75	1.23	0.84	1.22
8	0.73	1.06	0.75	1.18	0.63	1.23	0.77	1.23
Reading								
K	2.32	1.44	2.20	1.34	1.85	1.33	2.22	1.31
1	2.08	1.26	2.08	1.28	1.80	1.32	2.09	1.26
2	1.83	1.36	1.83	1.31	1.62	1.32	1.85	1.28
3	1.43	1.28	1.53	1.28	1.36	1.34	1.56	1.28
4	1.05	1.20	1.18	1.21	1.15	1.30	1.22	1.22
5	0.86	1.17	1.00	1.22	0.98	1.34	1.06	1.25
6	0.60	1.16	0.68	1.27	0.70	1.36	0.74	1.35
7	0.51	1.16	0.63	1.30	0.59	1.36	0.69	1.38
8	0.46	1.15	0.63	1.33	0.56	1.37	0.68	1.39

Notes: Estimates calculated using Formula (3). Effect sizes expressed in RIT points. SD = standard deviation. “All” refers to all students in the sample. Ever-EL refers to students ever flagged as “EL” between 2014-15 and 2018-19. Ever-EL+SPED refers to dually-identified students who were eligible to receive EL and ever eligible to receive SPED services between 2014-15 and 2018-19. Current-ELs are students eligible to receive EL service during each grade. All estimated gaps are statistically significant. G.Rate = growth rate.