

# Grade 8 to 12 Academic Growth Patterns for English Language Learners and Students with Disabilities

Dina Bassiri  
Jeff Allen

April 2012

For additional copies, write:  
ACT Research Report Series  
P.O. Box 168  
Iowa City, IA 52243-0168

© 2012 by ACT, Inc. All rights reserved.

# **Grade 8 to 12 Academic Growth Patterns for English Language Learners and Students with Disabilities**

Dina Bassiri  
Jeff Allen



## Table of Contents

<b>Table of Contents</b> .....	<b>ii</b>
<b>Abstract</b> .....	<b>iv</b>
<b>Introduction</b> .....	<b>1</b>
<i>NCLB, AYP, and Student Subgroups</i> .....	<i>1</i>
<i>Population of English Language Learner and Students with Disabilities</i> .....	<i>2</i>
<i>Achievement and AYP Challenges for English Language Learner and Students with Disabilities</i> .....	<i>4</i>
<i>The Need for Measuring Student Growth</i> .....	<i>5</i>
<b>Methodology</b> .....	<b>7</b>
<i>Subgroup Definitions</i> .....	<i>7</i>
<i>Predicted Test Scores</i> .....	<i>9</i>
<b>Results</b> .....	<b>11</b>
<i>Achievement Differences</i> .....	<i>13</i>
<i>Hierarchical Modeling Results</i> .....	<i>14</i>
<i>Growth Differences</i> .....	<i>15</i>
<b>Discussion</b> .....	<b>24</b>
<i>Limitations</i> .....	<i>26</i>
<b>References</b> .....	<b>29</b>
<b>Appendix</b> .....	<b>33</b>



## **Abstract**

Under the current version of the Elementary and Secondary Education Act (ESEA), public schools must show consistent improvement in achievement for all students including English language learners (ELL) and students with disabilities. The assessment of ELL students and students with disabilities can present unique challenges to schools that must make adequate yearly progress (AYP) toward high proficiency rates. Alternatives to accountability systems based on AYP often seek to employ measures of academic *growth*. This study used the longitudinal data of 2011 high school graduates who took EXPLORE, PLAN, and the ACT in grades 8, 10, and 11/12, respectively, to examine growth differences for ELL students and students with disabilities. ELL students experienced less-than-expected growth in English and reading, but above-average growth in mathematics. Students with disabilities experienced less-than-expected growth in all four subject areas. Most of the growth differences, while statistically significant, were small in magnitude.





# **Grade 8 to 12 Academic Growth Patterns for English Language Learners and Students with Disabilities**

## **Introduction**

English language learner (ELL<sup>1</sup>) students are the fastest growing U.S. educational subgroup in size and diversity. Thus, accelerating their learning and ensuring their fair and valid assessment of academic progress is imperative (Francis, Rivera, Lesaux, Kieffer & Rivera, 2006). Similarly, it is critical that assessments for students with disabilities are valid and reliable, and permit stakeholders to measure progress. In this study, we examine growth patterns, based on college readiness assessments, from grade 8 to grade 12 for these two important subgroups.

## **NCLB, AYP, and Student Subgroups**

Under the current accountability framework of No Child Left Behind (NCLB), schools are rewarded or sanctioned based on the percentage of students proficient in math and reading. Schools are also evaluated based on the percentage of students in specific subgroups achieving proficiency. Specifically, NCLB mandates reporting adequate yearly progress (AYP)<sup>2</sup> for all students, including ELL students, students with disabilities, and other subgroups.<sup>3</sup> If any one of the subgroups fails to make AYP under NCLB for two consecutive years, the entire school is designated "in need of improvement" or "failing." Ultimately, all students must reach 100 percent proficiency by the 2013-2014 school year.

---

<sup>1</sup> The term English language learner (ELL), indicates a person who is in the process of acquiring English and has a first language other than English. Other terms interchangeably used in the literature include language minority students, limited English proficient (LEP), English as a second language (ESL), and culturally and linguistically diverse (CLD) (The Free Library, 2008).

<sup>2</sup> No Child Left Behind requires states to measure AYP for schools receiving Title I funds with the goal of all students reaching the proficient level on reading/language arts and mathematics tests by the 2013-2014 school year. States must define minimum levels of improvement as measured by standardized tests chosen by the state. AYP targets must be set for overall achievement and for subgroups of students.

<sup>3</sup> Defined by race/ethnicity and economic disadvantage.

Many ELL students are also racial-ethnic minorities, low-income, or have disabilities, meaning they may count in two or more groups' AYP results. ELL students and students with disabilities typically obtain the lowest test scores, contributing to multiple and difficult hurdles to meeting AYP. Research has shown that more diverse schools are less likely to make AYP under NCLB (Neill, Guisbond, & Shaeffer, 2004; Kane & Staiger, 2003).

### **Population of English Language Learner and Students with Disabilities**

According to the U.S. Government Accountability Office (GAO), approximately 5 million ELL students were enrolled in public schools during the 2003-2004 school year, representing about 10 percent of the total school population (U.S. Government Accountability Office, 2006). Over a 10-year period between the 1994–95 and 2004–05 school years, the enrollment of ELL students grew over 60 percent, while the total K–12 growth was just over 2 percent (Office of English Language Acquisition [OELA], n.d.). The increase in ELL students is much more pronounced in some states.<sup>4</sup> While ELL students reside throughout the United States, they are heavily concentrated in Arizona, California, Texas, New York, Florida, and Illinois. These six states contain 61 percent of the nation's ELL population.<sup>5</sup> Approximately 79 percent of ELL students nationally are from Spanish language backgrounds. California serves about one-third of all the nation's ELL students – 1.6 million students, 85 percent of whom are Spanish speaking (Payán & Nettles, n.d.). Not only is the ELL population growing, but the group is also becoming more diverse. Over 400 different languages are reported among these students, and quality of prior schooling varies considerably – some may have been well educated in their

---

<sup>4</sup> For instance, North Carolina and Nevada have reported their ELL population growth rate as 500% and 200%, respectively during this period (Batalova, Fix, & Murray, 2005, as cited in Short & Fitzsimmons, 2007).

<sup>5</sup> Despite the high concentration in six states, other states have experienced a 300 percent or higher growth of ELL students in a ten-year period from 1995 to 2005. These states include Alabama, Indiana, Kentucky, Nebraska, North Carolina, South Carolina, and Tennessee. In addition, the U.S. territory of Puerto Rico accounts for another 1 percent of the U.S. ELL population (Payán & Nettles, n.d.).

country of origin while others have enormous educational gaps when they begin their schooling in the U.S.

Before the formation of the Education for All Handicapped Children Act (EHA) of 1975, U.S. public schools educated only 1 out of 5 children with disabilities. Until that time, many states had laws that explicitly excluded children with certain types of disabilities from attending public school, including children who were blind, deaf, or labeled "emotionally disturbed" or "mentally retarded." At the time the EHA was enacted, more than 1 million children in the U.S. had no access to the public school system. From 1988-89 to 1998-99, special education enrollments grew about twice as quickly as those of all students (33% versus 15% growth). Special education costs per student also rose during this period. The population of students served under Individuals with Disabilities Education Act (IDEA)<sup>6</sup> has grown at nearly twice the rate of the general education population. During the twenty-five year period between 1980 and 2005, the population of students served under IDEA increased by 37 percent, while the general education population grew by only 20 percent. Students served under IDEA in 2006 account for about 14 percent of the total education population, up from about 10 percent in the 1980s<sup>7</sup>

---

<sup>6</sup> IDEA is a United States federal law enacted in 1990 and reauthorized in 1997. It governs how states and public agencies provide early intervention, special education, and related services to children with disabilities between the ages of 3 to 21 who meet the eligibility criteria in one of thirteen specified categories of disability. It is designed to protect the rights of students with disabilities by ensuring that everyone receives a Free Appropriate Public Education (FAPE) regardless of ability that prepares them for further education, employment, and independent living. The categories of disabilities are: autism, deaf/blind, deafness, visual impairment including blindness, hearing impaired, mental retardation, multiple disabilities, orthopedic impairment, serious emotional disturbance, specific learning disabilities, speech or language impairment, traumatic brain injury, and other health impairment. To be eligible, a student must have a disability that adversely affects her or his educational performance and must need special education in order to receive an appropriate education (National Resource Center on ADHD, n.d).

<sup>7</sup> This increase has not been uniform across all disability types. The biggest increases have been in the "specific learning disabilities" category, which grew over 200 percent during this time; the "other health-impairments" category, which began to increase in 1991 when the federal definition was expanded to include ADD and attention deficit hyperactivity disorder (ADHD); and in autism, which has increased partly due to an expanded definition that includes children with milder forms of the disorder. The increase in students identified as having disabilities is a reflection of a variety of factors, including an increased acceptance of the "disabled" label among parents and students, a better understanding of disabilities, and changes in the definition of what qualifies as a disability (Dillon, 2007).

(Federal Education Budget Project (FEBP), 2011a). Three out of every four students with disabilities spend part or all of their school day in a general education classroom. In turn, nearly every general education classroom across the country includes students with disabilities (National Education Association, IDEA / Special Education, n.d.). Over six million children in the United States currently receive services under IDEA (Federal Education Budget Project (FEBP), 2011b).

### **Achievement and AYP Challenges for English Language Learner and Students with Disabilities**

The academic achievement gap between ELL and non-ELL students is considerable. ELL students' math proficiency level averaged 20 percent lower than that of the overall population in 2003–2004 across 48 states (U.S. Government Accountability Office, 2006).

ELL students are *defined* by their level of English language proficiency. Students enter the subgroup often speaking little English while others leave after being reclassified as fully proficient in English, with a common exit criterion being the 36<sup>th</sup> percentile in an English reading/language arts assessment. When they have demonstrated proficiency in English, they exit the subgroup and their scores are no longer counted in the computation of AYP (Crawford, 2004). Therefore, the ELL student group is a highly fluid population.

Many ELL students attend schools with large concentrations of at-risk minority and low-income students, so these schools struggle to meet AYP proficiency goals for multiple subgroups (Neill, Guisbond & Schaeffer, 2004). These schools face additional issues for ELL students, as they are not proficient in English, and the tests that are used to measure academic achievement are often administered in English (Abedi, 2002, 2004, 2007).

Making AYP has been particularly challenging for the special education subgroup. In a study on fourth-grade readers in 39 states, there was an achievement gap of 30% between students with disabilities and their general education peers in 30 of the 39 states (Pew Charitable Trust, c.f. Eckes & Swando, 2009, p. 2481). Data from California, Florida, and Texas suggest that the AYP requirements often are not met for the students with disabilities subgroup (Eckes & Swando, 2009.). Furthermore, Eckes & Swando make the following observation:

The failure of the special education subgroup to make AYP occurs mainly because the students with disabilities subgroup is expected to maintain the exact same proficiency levels as their general education peers—a standard that has proved to be problematic because special education students often start out with lower average test scores than general education students... [In addition], the “students with disabilities” subgroup is the only subgroup in which actual limitations on ability to learn might come into play. The existence of these limitations calls into question the wisdom of trying to close the general education–special education “achievement gap” at the same pace as the race- or class-based achievement gaps.

The importance placed on meeting AYP targets, and the difficulty in doing so, has raised numerous concerns, particularly for ELL students and students with disabilities. An alternative to setting targets for the percentage of students meeting proficiency would be to set targets for the percentage of students meeting growth standards. In this study, we measure growth between grades 8 and 12 for a large sample of students and examine whether growth is different for ELL students and students with disabilities. The study provides insight into whether these student subgroups would be more or less likely than their peers to meet growth standards.

### **The Need for Measuring Student Growth**

Using a single test score, it is impossible to tell whether a student made significant progress, made no progress, or even regressed. Students with disabilities may have made more than a year’s growth in math or reading and still not meet the state's proficiency target. This is supported by a study that examined the effect of initial achievement and cognitive ability on

academic progress over time (Rescorla & Rosenthal, 2004). Students not meeting proficiency must make more progress than their peers in order to “catch up.” Within the current framework of NCLB, these students would not be given credit for growing more than their peers, unless that growth resulted in passing the state’s proficiency mark.<sup>8</sup>

Growth measurement could be an important contribution to accountability systems insofar as it represents a method for measuring learning that has occurred. Examples of growth-centric alternatives are referenced in calls for ESEA reauthorization to change the current provisions required by NCLB (U.S. Department of Education, 2010; Council of Chief State School Officers, 2010). Growth models measure how much progress a student has made from one point in time to another. A “simple” growth model measures the growth of individual students over time, without comparing their progress to an annual growth target chosen to keep them on track for achieving proficiency in the future. Models which include a growth target are often referred to as “growth-to-proficiency” models. Growth-to-proficiency models establish an expected rate of growth—based on previous year’s achievement, the grade level of the student, and the time frame required to demonstrate proficiency—and determine whether or not the expected growth rate was achieved.

Predictive models relate students’ test scores to background factors, prior test scores, and, in some cases, teacher or school characteristics. In this study, we apply predictive models to longitudinal data from the high school graduating cohort of 2011 who took EXPLORE, PLAN, and the ACT<sup>9</sup>. Growth is defined as the difference between an observed test score and a

---

<sup>8</sup> As of January 2012, several states are seeking waivers to NCLB that permit growth-centric accountability measures.

<sup>9</sup> ACT’s College and Career Readiness System includes EXPLORE (for 8<sup>th</sup> graders), PLAN (for 10<sup>th</sup> graders), and the ACT (for 11<sup>th</sup> and 12<sup>th</sup> graders). Each assessment contains four subject tests: English, Mathematics, Reading, and Science (the ACT also contains an optional Writing Test). The maximum possible score on each subject test is 25 for EXPLORE, 32 for PLAN, and 36 for the ACT. All three assessments measure academic achievement, respective to the curriculum of the grade level for which it is intended, and use the same score scale.

predicted test score based on prior test scores and other variables. The goal of the study is to measure the extent that growth from grade 8 to grade 12 varies for ELL and students with disabilities.

A *value-added* model uses predictive modeling to estimate how much a particular school, teacher, or intervention program has “added value” to students’ test scores (Rubin, Stuart, & Zanutto, 2004). Value-added models decompose the variance of the test scores into segments that are explained by student inputs (i.e., adjusting for entering achievement level and socio-economic status), and into others that are attributed to school, teacher, or program effectiveness. Schools and teachers are held accountable only for the portions of variance over which they have control. Value-added models control for extraneous variables that, if not accounted for, could lead to biased conclusions regarding a school’s or teacher’s effectiveness. This study does not consider inferences about the effectiveness of schools or teachers.

### **Methodology**

To determine if growth varies for ELL students and students with disabilities, this study used longitudinal data from a large sample of students who took EXPLORE in 8<sup>th</sup> grade, PLAN in 10<sup>th</sup> grade, and the ACT in 11<sup>th</sup> or 12<sup>th</sup> grade. Data collected as part of the assessment programs were used to classify students into subgroups and to measure growth between grade 8 and grade 11/12.

### **Subgroup Definitions**

As part of the EXPLORE and PLAN programs, some schools provide ACT with extra demographic and subgroup data, including indicators for which students are classified as ELL and which are in special education (SPED) programs. For the purposes of this study, students with ELL and/or SPED indicators in the EXPLORE or PLAN data were classified into the ELL

or SPED subgroups, respectively. As discussed earlier, the ELL subgroup is a highly fluid population, with students entering and departing the subgroup based on their English proficiency. In this study, however, we treat the ELL subgroup as fixed between grade 8 (EXPLORE) and grade 11/12 (the ACT). Because not all schools that use EXPLORE and PLAN provide ACT with ELL and SPED indicators, we cannot always distinguish non-ELL (non-SPED) students from students missing this information. In other words, some students may be ELL or SPED, but are not classified as such because their school did not provide the data. Thus, we excluded students from schools with no reported ELL or SPED students.

Included in the Student Profile Section of the ACT registration form is the question *Do you have a disability that requires special provisions from the educational institution?* (The institution is referring to a college the student plans to attend.) Students who responded “visual impairment (not correctable),” “hearing impairment,” or “motor impairment” to this question were classified as having a *physical disability*. Those who responded “learning or cognitive disability” were classified as having a *cognitive/learning disability*, and those who responded “multiple disabilities” or “other disability” were classified as having *other disability*.

Because the school-reported SPED indicator is not demarcated by type of disability, we also included the student-reported disability subgroup in order to examine growth differences for students with different types of disabilities. We believe that the SPED subgroup is a more “official” classification because it is based on school-provided data; the use of student-reported disability type is a limitation of this study.

We identified seven groups of students for analysis. The first group was used as a reference group and included those who were neither considered as ELL, SPED, nor a student with a self-reported disability. The second and third groups were ELL and SPED students,



respectively. The fourth group was students with any self-reported disability. Groups 5-7 are subsets of the fourth group (any disability), demarcated by type of disability: physical disability, cognitive/learning disability, or other type of disability, respectively<sup>10</sup>. It is important to note that the ELL, SPED, and any disability groups are not mutually exclusive – students could be members of more than one of these groups. Moreover, we assume that students in the SPED group have a disability and thus treat the SPED and student-report disability groups as representative of the *students with disabilities* population.

### Predicted Test Scores

We modeled PLAN (ACT) test scores for each student, using as predictors the student's prior EXPLORE (PLAN) scores in all four subject areas, the number of months that passed between the two tests, and the prior subject-specific mean EXPLORE (PLAN) score for the high school. To account for unobserved high school effects, a random intercept model was used. The model for projected scores is given explicitly in Equation 1.

$$\hat{Y}_{ij} = \hat{\beta}_0 + \sum_{p=1}^4 \hat{\beta}_p X_{ijp} + \hat{\beta}_5 S_{ij} + \hat{\beta}_6 \bar{X}_j + \hat{\tau}_j \quad (1)$$

In Equation 1,  $\hat{Y}_{ij}$  is the projected later test score for the  $i^{\text{th}}$  student from the  $j^{\text{th}}$  high school,  $\hat{\beta}_0$  is the estimate for the overall intercept term,  $X_{ijp}$  ( $p=1,2,3,4$ ) are the prior test scores in the four subject areas for the  $i^{\text{th}}$  student from the  $j^{\text{th}}$  high school,  $\hat{\beta}_p$  ( $p=1,2,3,4$ ) are the estimated regression coefficients for the four prior test scores,  $S_{ij}$  is the number of months

---

<sup>10</sup> Disabilities categorized as *physical* may include cerebral palsy, deafness, epilepsy, physical disability, speech/language disability, Tourette syndrome, non-correctable visual, hearing impairment, and motor impairment. Disabilities categorized as *cognitive/learning* may include attention deficit disorder, anxiety disorder, emotionally disabled, limited intellectual capacity, psychological disability, developmental arithmetic disorder, developmental writing disorder, learning disability, reading disability, and learning or cognitive disability. Disabilities categorized as *other* may include autism, homebound, multiple disabilities, and other disability.

between the prior and later tests (EXPLORE to PLAN, PLAN to the ACT, or EXPLORE to the ACT) for the  $i^{\text{th}}$  student from the  $j^{\text{th}}$  high school,  $\hat{\beta}_5$  is the estimated regression coefficient for the testing span,  $\bar{X}_j$  is the mean prior test score in the subject area of interest for the  $j^{\text{th}}$  high school,  $\hat{\beta}_6$  is the estimated regression coefficient for the mean prior test score, and  $\hat{\tau}_j$  is the estimated school-specific intercept (random intercept term assumed to be normally distributed with mean 0 and variance  $\tau_{00}$ ). The model is a special case of a hierarchical linear model (Raudenbush & Bryk, 2002) and can be fit using statistical software packages such as HLM<sup>®</sup> or SAS<sup>®</sup>.

By using the amount of time between tests as a covariate (Allen, Bassiri, & Noble, 2009), the projected scores accommodate varying time spans between tests, as more growth is expected as more time passes. Additionally, because of possible peer effects (Angrist & Lang, 2002), the projected scores also take into account the subject-specific mean prior achievement observed in each high school. The random intercept model adjusts the projected scores for unobserved school effects on student growth, so that subgroup differences in residual scores (described next) are not confounded by the school effects.

Residual scores are defined as the difference between students' actual and projected scores. We aggregated residual scores by subgroups to form measures of aggregate growth for each subgroup. This was done separately for each subject area (English, mathematics, reading, and science), and for each grade level span (8 to 10, 10 to 11/12, and 8 to 11/12).

By definition, the mean residual scores across all students in the study sample are zero. Because the reference group is by far the largest group in the study, it is expected that the mean residual score for the reference group will be close to zero. Mean residual scores that are positive

(negative) and deviate significantly from zero are evidence of greater-than-expected (less-than-expected) growth between tests. For each mean residual score, we tested whether the mean was significantly different than zero. Statistical tests with p-values less than 0.01 were considered significant.

## **Results**

The sample sizes, gender, and racial/ethnic breakdowns for the seven groups of students, the total sample, and the 2011 ACT-tested high school graduating class are presented in Table 1. As was previously mentioned, groups 5-7 are subsets of the fourth group (any disability). The total number of students in the study sample is 103,725. ELL students made up 2% of the sample, SPED students made up 7%, and students with disabilities made up 6% of the sample. Of the students with disabilities, 60% had a cognitive/learning disability, 17% had a physical disability, and the other 23% were classified as having some other type of disability. Note that the gender and racial/ethnic breakdowns of the reference group and the total sample is similar to that for the 2011 ACT-tested high school graduating class.

Table 1

*Demographics of Study Sample*

Group	N	Gender		Race/Ethnicity				
		Female	Male	White	African American	Hispanic	Asian	Other
Reference	89,344	54%	45%	62%	12%	12%	4%	7%
English language learner	2,548	50%	50%	17%	6%	45%	18%	6%
Special education	7,411	40%	60%	53%	15%	13%	2%	7%
Any disability	6,538	49%	51%	71%	13%	6%	2%	8%
Physical disability	1,089	54%	45%	62%	20%	7%	3%	7%
Cognitive/learning Disability	3,949	47%	53%	75%	9%	6%	1%	8%
Other disability	1,500	51%	49%	65%	17%	6%	2%	9%
Sample total	103,725	53%	47%	61%	12%	13%	4%	7%
ACT-tested graduating class of 2011	1,623,112	54%	46%	60%	14%	12%	4%	9%

**Note:** Percentages do not sum to 100 due to no response to gender or race/ethnicity questions.

Most of the reference group's students were White (62%), with African American (12%), Hispanic (12%), Asian (4%), and students of other racial/ethnic groups (7%) also represented. Most of the ELL students were Hispanic (45%) or Asian (18%). SPED students and the students with disabilities were predominantly White (53% and 71%, respectively). Note that large portions of the ELL and SPED subgroups (8% and 10%, respectively) have unknown or missing race/ethnicity. Students with physical disability and the students with other disability had relatively more African American students (20% and 17%, respectively).

The gender breakdown of ELL students was evenly distributed (50% female vs. 50% male). Compared to the reference group, ELL females were slightly underrepresented (50% vs. 54%) and ELL males were slightly overrepresented (50% vs. 45%). On the other hand, SPED students and students with disabilities were more likely to be male (60% and 51%, respectively). Students with physical disability and the students with other disability had relatively more

female students (54% and 51%, respectively). However, the cognitive/learning disability group had more male students (53%).

### **Achievement Differences**

Summary statistics of EXPLORE, PLAN, and ACT scores were computed for each of the seven groups of students (see Table 2). Across subject areas, the average EXPLORE, PLAN, and ACT scores were largest for the reference group and smallest for ELL and SPED students. Among students with disabilities, the average scores across subject areas were lower for students with cognitive/learning disabilities than those for students with physical disabilities. Following the reference group, the largest average scores were achieved by students with physical disabilities.

TABLE 2

*Summary Statistics of EXPLORE, PLAN, and ACT Scores*

Group	N	Test <sup>11</sup>	Subject area			
			English	Math	Reading	Science
			Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Reference	89,344	EXPLORE	15.7 (4.1)	16.5 (3.5)	15.5 (3.8)	17.4 (3.0)
		PLAN	18.2 (4.3)	19.0 (4.7)	18.0 (4.5)	19.2 (3.9)
		ACT	21.6 (6.2)	21.6 (5.4)	21.8 (6.0)	21.7 (5.1)
English language learner	2,548	EXPLORE	11.4 (3.5)	13.2 (3.8)	12.2 (2.8)	14.9 (2.7)
		PLAN	14.4 (4.0)	16.1 (4.2)	14.5 (4.2)	16.9 (3.2)
		ACT	15.9 (6.1)	18.5 (4.8)	17.0 (5.4)	18.0 (4.9)
Special education	7,411	EXPLORE	11.6 (3.6)	12.6 (4.0)	12.3 (3.0)	14.7 (2.9)
		PLAN	14.0 (4.2)	15.3 (4.3)	14.6 (4.2)	16.5 (3.3)
		ACT	15.5 (6.0)	17.4 (4.5)	17.2 (5.8)	17.5 (5.2)
Any disability	6,538	EXPLORE	13.8 (4.0)	14.6 (3.7)	13.8 (3.5)	16.0 (3.0)
		PLAN	16.0 (4.3)	16.8 (4.1)	16.0 (4.3)	17.5 (3.4)
		ACT	18.7 (6.1)	19.1 (4.6)	19.6 (5.9)	19.6 (5.0)
Physical disability	1,089	EXPLORE	14.1 (4.3)	15.0 (3.7)	14.2 (3.6)	16.3 (3.0)
		PLAN	16.6 (4.4)	17.3 (4.1)	16.6 (4.5)	17.9 (3.5)
		ACT	19.0 (6.2)	19.4 (4.7)	19.7 (6.1)	19.8 (5.0)
Cognitive/learning Disability	3,949	EXPLORE	13.6 (4.0)	14.4 (3.7)	13.6 (3.4)	15.8 (3.0)
		PLAN	15.8 (4.2)	16.6 (4.1)	15.8 (4.2)	17.4 (3.3)
		ACT	18.5 (6.1)	19.1 (4.6)	19.5 (5.9)	19.6 (5.0)
Other disability	1,500	EXPLORE	14.1 (4.0)	14.8 (3.6)	13.9 (3.4)	16.2 (2.9)
		PLAN	16.2 (4.3)	17.0 (4.0)	16.2 (4.4)	17.6 (3.4)
		ACT	18.8 (6.0)	19.0 (4.6)	19.6 (5.9)	19.4 (4.9)

### Hierarchical Modeling Results

We now examine results from the two-level linear regression model (Equation 1) used to generate the projected and residual scores. Twelve models were fit – one for each combination of subject area (English, mathematics, reading, and science) and grade level span (8 to 10, 10 to 11/12, and 8 to 11/12). Tables in the Appendix contain the parameter estimates for the projection models from EXPLORE to PLAN, PLAN to the ACT, and EXPLORE to the ACT, respectively.

<sup>11</sup> EXPLORE was administered in grade 8, PLAN in grade 10, and the ACT in grades 11/12.

As presented in Tables A1 and A3 in the Appendix, the level of homogeneity of projected scores within schools, as measured by the intraclass correlation coefficients (ICC), are higher for the EXPLORE to PLAN and EXPLORE to ACT models, ranging from 0.14 to 0.24. The ICCs for the PLAN to ACT models ranged from 0.02 to 0.06. Across models, the majority of overall variance in later test scores is due to student characteristics. All regression coefficients for prior test scores and time span are positive and statistically significant ( $p$ -value  $< 0.01$ ). These results suggest that prior test scores and time span are predictive of later test scores, and that score projections increase with increasing prior test scores and more time in between tests. With the exception of the PLAN to ACT models (Table A2 in the Appendix), school subject-specific mean prior test scores were not predictive of later test scores. As expected, the same-subject regression coefficients of prior test scores were the largest across all models. It is also apparent that the second largest regression coefficients tend to be the off-subject member of the English/reading and math/science test pairs. The correlation of projected and observed later test scores (Multiple  $R$ ) ranged from 0.71 to 0.87. In this study, Multiple  $R$  reflects the strength of relationship between the projected and actual later test scores. In sum, these results signify the importance of prior achievement and number of months between tests in projecting later test scores.

### **Growth Differences**

Averaging residual scores by subgroup provides a measure of each subgroup's average relative growth (for each subject area and for each grade span). Table 3 contains the mean residual scores for English and reading and each grade span (grade 8 to 10, grade 10 to 11/12, and grade 8 to 11/12), as well as their standard errors. Similarly, Table 4 contains the mean residual results for mathematics and science.

TABLE 3

*Summary Statistics of Mean Residual Scores in English and Reading*

Group	N	Grade Span	Subject Area	
			English	Reading
			Mean (SE)	Mean (SE)
Reference	89,344	8-10	<b>0.07</b> (0.01)	<b>0.04</b> (0.01)
		10-11/12	<b>0.08</b> (0.01)	<b>0.01</b> (0.01)
		8-11/12	<b>0.09</b> (0.01)	0.03 (0.01)
English language learner	2,548	8-10	<b>-0.16</b> (0.06)	<b>-0.19</b> (0.07)
		10-11/12	<b>-0.81</b> (0.06)	<b>-0.61</b> (0.07)
		8-11/12	<b>-0.40</b> (0.08)	<b>-0.27</b> (0.08)
Special education	7,411	8-10	<b>-0.67</b> (0.04)	<b>-0.32</b> (0.04)
		10-11/12	<b>-0.78</b> (0.04)	<b>-0.06</b> (0.04)
		8-11/12	<b>-0.92</b> (0.05)	<b>-0.20</b> (0.05)
Any disability	6,538	8-10	<b>-0.36</b> (0.03)	<b>-0.28</b> (0.04)
		10-11/12	<b>-0.07</b> (0.04)	<b>0.19</b> (0.05)
		8-11/12	<b>-0.37</b> (0.05)	<b>-0.17</b> (0.05)
Physical disability	1,089	8-10	-0.05 (0.08)	0.03 (0.09)
		10-11/12	<b>-0.29</b> (0.10)	<b>-0.23</b> (0.12)
		8-11/12	-0.22 (0.10)	-0.24 (0.12)
Cognitive/ learning disability	3,949	8-10	<b>-0.47</b> (0.04)	<b>-0.39</b> (0.05)
		10-11/12	<b>0.03</b> (0.06)	<b>0.38</b> (0.07)
		8-11/12	<b>-0.40</b> (0.06)	-0.11 (0.07)
Other disability	1,500	8-10	<b>-0.31</b> (0.07)	-0.20 (0.08)
		10-11/12	<b>-0.15</b> (0.09)	<b>0.00</b> (0.10)
		8-11/12	<b>-0.38</b> (0.09)	<b>-0.27</b> (0.10)

**Note:** Means in bold are statistically significant at  $p < .01$ .



TABLE 4

*Summary Statistics of Mean Residual Scores in Math and Science*

Group	N	Grade Span	Subject Area	
			Math	Science
			Mean (SE)	Mean (SE)
Reference	89,344	8-10	0.02 (0.01)	0.02 (0.01)
		10-11/12	<b>0.02 (0.01)</b>	<b>0.04 (0.01)</b>
		8-11/12	0.01 (0.01)	<b>0.04 (0.01)</b>
English language learner	2,548	8-10	<b>0.43 (0.06)</b>	<b>0.28 (0.05)</b>
		10-11/12	<b>0.15 (0.05)</b>	<b>-0.44 (0.06)</b>
		8-11/12	<b>0.68 (0.07)</b>	0.00 (0.07)
Special education	7,411	8-10	<b>-0.14 (0.04)</b>	<b>-0.11 (0.03)</b>
		10-11/12	<b>-0.24 (0.03)</b>	<b>-0.42 (0.04)</b>
		8-11/12	<b>-0.12 (0.04)</b>	<b>-0.36 (0.05)</b>
Any disability	6,538	8-10	<b>-0.31 (0.03)</b>	<b>-0.23 (0.03)</b>
		10-11/12	<b>-0.08 (0.03)</b>	<b>0.06 (0.04)</b>
		8-11/12	<b>-0.33 (0.04)</b>	<b>-0.22 (0.04)</b>
Physical disability	1,089	8-10	-0.11 (0.08)	-0.06 (0.07)
		10-11/12	<b>-0.10 (0.08)</b>	<b>-0.14 (0.09)</b>
		8-11/12	-0.16 (0.09)	-0.19 (0.10)
Cognitive/ learning disability	3,949	8-10	<b>-0.36 (0.05)</b>	<b>-0.28 (0.04)</b>
		10-11/12	<b>0.04 (0.04)</b>	<b>0.21 (0.06)</b>
		8-11/12	<b>-0.28 (0.05)</b>	<b>-0.16 (0.06)</b>
Other disability	1,500	8-10	<b>-0.33 (0.07)</b>	<b>-0.24 (0.06)</b>
		10-11/12	<b>-0.37 (0.07)</b>	<b>-0.19 (0.09)</b>
		8-11/12	<b>-0.58 (0.08)</b>	<b>-0.41 (0.09)</b>

**Note:** Means in bold are statistically significant at  $p < .01$ .

**Grade 8 to 10.** We begin with the growth from grade 8 to grade 10 (EXPLORE to PLAN). From Tables 3 and 4, we see that the mean residual scores for larger groups tend to have less sampling error (i.e., smaller standard errors). Thus, the larger the group size, the more precise is the measurement of average relative growth. ELL students were the only group that had statistically significant ( $p$ -value  $< 0.01$ ) and positive value-added scores in mathematics and science between EXPLORE to PLAN testing (grade 8 to 10). For example, ELL students had a value-added score of 0.43 (SE=0.06) for the grade 8 to grade 10 model (Table 4). Students in this subgroup demonstrated 0.43 score points of growth above what would have been expected, given

their 8<sup>th</sup> grade test scores, their school's mean 8<sup>th</sup> grade scores in mathematics, the time between the two tests, and other variation in growth attributed to school differences.

SPED students and students with any disabilities, including those with cognitive/learning disabilities and other disabilities had statistically significant negative value-added scores in mathematics and science ( $p\text{-value} < 0.01$ ), meaning that these groups of students grew less than projected in mathematics and science between grade 8 and 10. On the other hand, for students with physical disabilities, value-added scores in mathematics and science were not significantly different than the average.

Less-than-expected growth in English and reading was observed in four of the subgroups. Compared to their non-ELL non-disabled peers (the reference group), ELL students, SPED students, those with any disabilities, and those with cognitive/learning disabilities scored below projections in both English and reading. Students with other disabilities grew less than expected in English, but their growth in reading was not significantly different than the average. The smallest growth was achieved by SPED students in English ( $M=-0.67$ ) and by students with cognitive/learning disabilities in reading ( $M=-0.39$ ). Again, for students with physical disabilities, value-added scores in English and reading were not significantly different than the average.

Figure 1 contains the value-added scores from EXPLORE to PLAN for ELL students, SPED students, and students with any disabilities, including those with physical, cognitive/learning, and other disabilities. Figure 1 includes 99% confidence intervals around each value-added score. As mentioned before, it is important to take into account the error associated with the value-added scores. For example, the sample size of the SPED subgroup is larger than the sample size of the other five subgroups; therefore, the standard errors and

corresponding confidence intervals associated with the value-added scores for the SPED subgroup are smaller (i.e., narrower error bands). In other words, the larger the group size, the more precise is the measurement of average relative growth. If the confidence interval bands contain 0, then we conclude that the observed growth was not significantly different from expected. However, if the confidence interval bands do not contain 0, then we are more confident (in this case, 99% confident) that the average relative growth was significantly above or below expected.

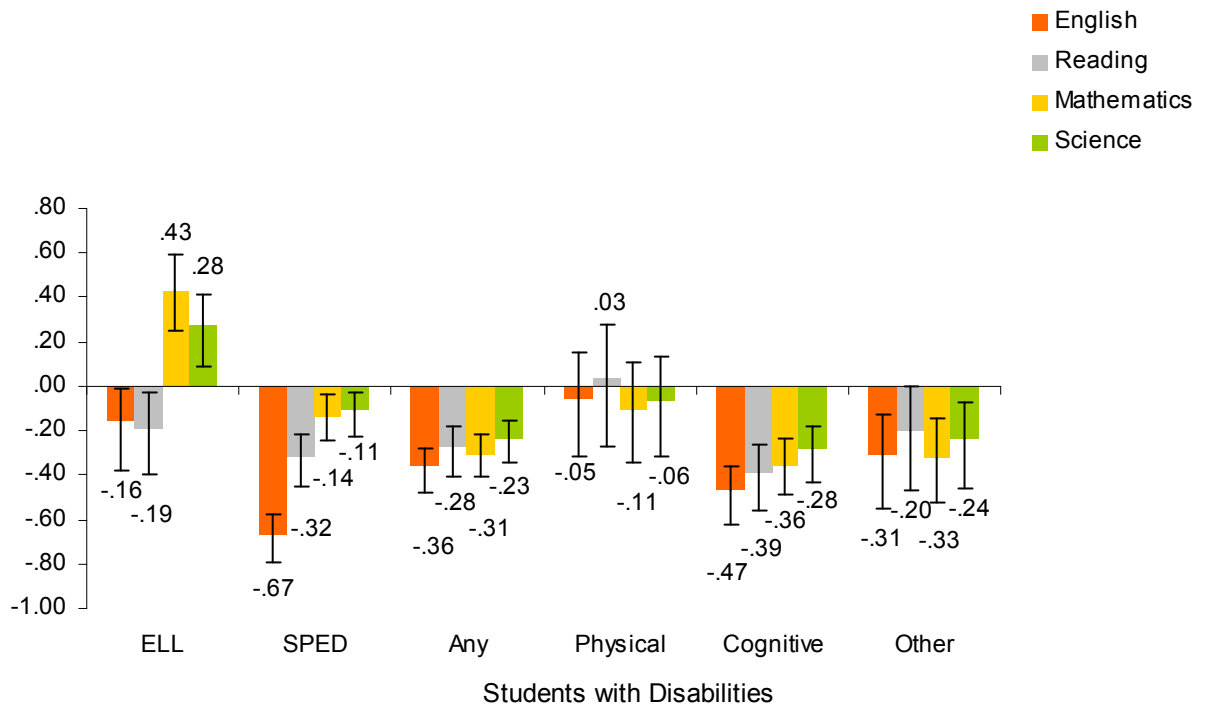


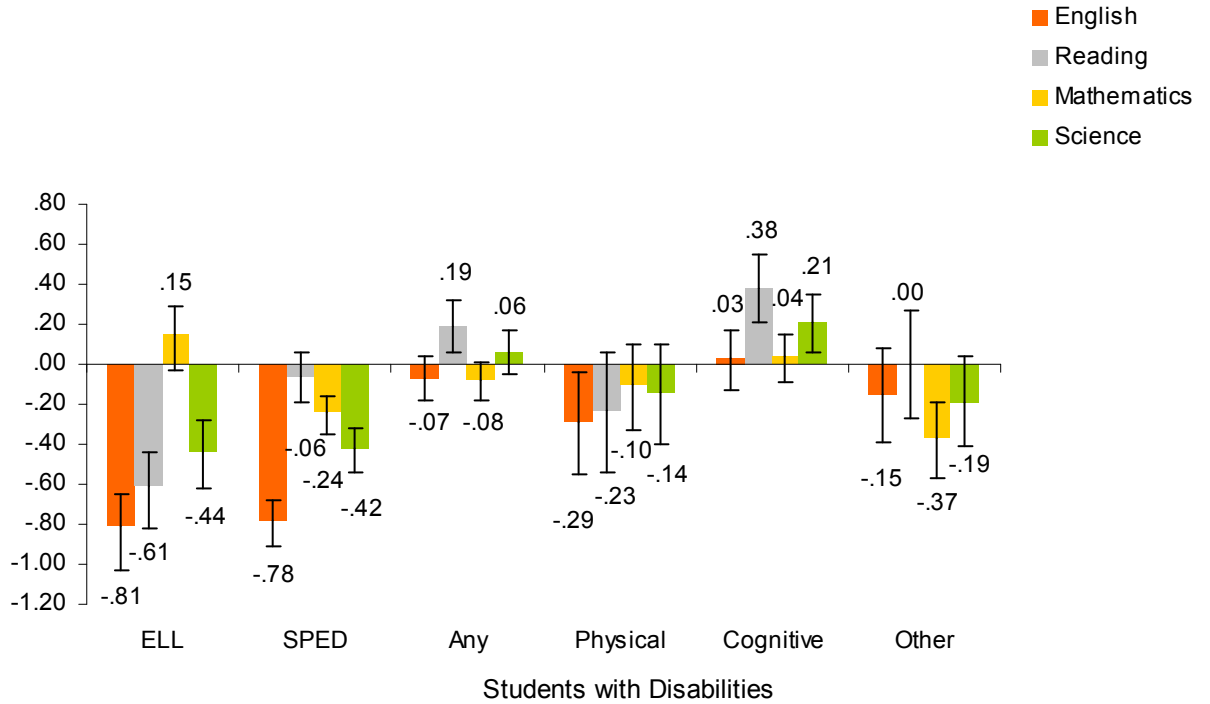
Figure 1. Mean Residual Scores from EXPLORE to PLAN, by Subgroup

**Grade 10 to 11/12.** Next, we examine growth from grade 10 to grade 11/12 (PLAN to the ACT). During this period, ELL students continued to have above-average growth in mathematics ( $M=0.15$ ), but below-average growth in science ( $M=-0.44$ ). SPED students continued to have less-than-expected growth in mathematics ( $M=-0.24$ ) and in science ( $M=-$

0.42). Interestingly, above-average growth in science ( $M=0.21$ ) was achieved by students with cognitive/learning disabilities. Students with other disabilities continued to have below-average growth in mathematics ( $M=-0.37$ ), but their growth in science was not significantly different than the average. For students with any disabilities and students with physical disabilities, value-added scores in mathematics and science were not significantly different than 0.

In English, less-than-expected growth was observed for ELL students ( $M=-0.81$ ), SPED students ( $M=-0.78$ ), and students with physical disabilities ( $M=-0.29$ ), suggest that ELL and SPED students experienced the least growth. In reading, contrary to the less-than-expected growth observed between EXPLORE to PLAN testing (grade 8 to 10), students with any disabilities ( $M=0.19$ ) and those with cognitive/learning disabilities ( $M=0.38$ ) observed above-expected growth between PLAN and the ACT testing (grade 10 to 11/12), but ELL students ( $M=-0.61$ ) continued to have less-than-expected growth, again showing the least growth.

Figure 2 graphically presents the value-added scores from PLAN to the ACT by subgroup. Figure 2 includes 99% confidence intervals around value-added scores for ELL students, SPED students, and students with any disabilities, including those with physical, cognitive/learning, and other disabilities.



*Figure 2.* Mean Residual Scores from PLAN to the ACT, by Subgroup

**Grade 8 to 11/12.** Finally, we consider growth between grade 8 and grade 11/12 (EXPLORE to the ACT). Figure 3, contains the value-added scores from PLAN to the ACT for ELL students, SPED students, and students with any disabilities, including those with physical, cognitive/learning, and other disabilities. Figure 3 displays 99% confidence intervals (error bands) around each value-added score.

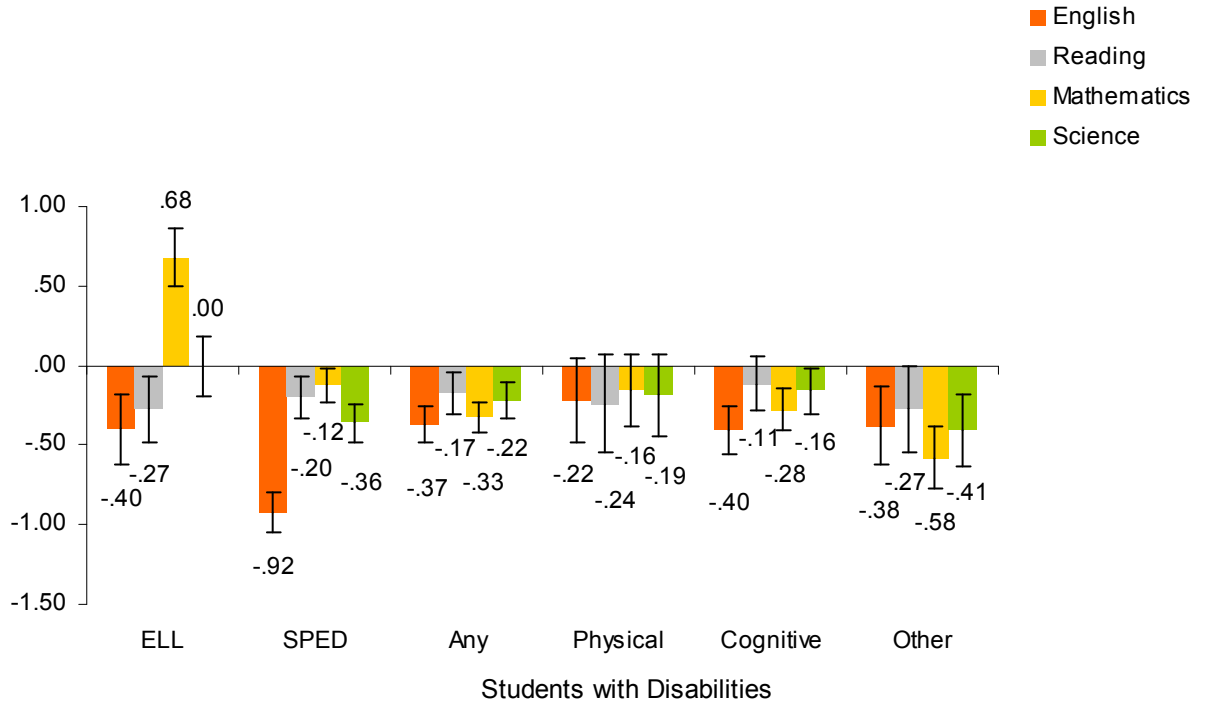


Figure 3: Mean Residual Scores from EXPLORE to the ACT, by Subgroup

Generally, the value-added scores for this grade span are statistically significant if one or both of the value-added scores are significant for the shorter spans (grade 8 to 10 and/or grade 10 to 11/12). There are some exceptions. For students with other disabilities, less-than-expected growth was observed in reading ( $M=-0.27$ ), but the value-added scores were not significantly different than 0 for the other two grade level spans (8 to 10 and 10 to 11/12). Moreover, in English, students with physical disabilities grew significantly less-than-expected between grades 10 and 11/12, but their growth between grades 8 and 11/12 was not significantly below average ( $M=-0.22$ ,  $SE=0.10$ ,  $p\text{-value} > .01$ ). In some cases, growth for the grade 8 to grade 11/12 span was not significantly different than average, despite one of the shorter spans showing a difference, but in opposite directions. For example, in science, ELL students grew significantly more than average between grades 8 and 10, but less than average between grades 10 and 11/12.

Similarly, in reading, students with cognitive/learning disabilities grew significantly less-than-expected between grades 8 and 10, but more-than-expected between grades 10 and 11/12.

Once again, ELL students continued to show above-average growth in mathematics ( $M=0.68$ ). SPED students and students with disabilities (any, cognitive/learning, and other) grew significantly less than average in mathematics and science. Compared to their non-ELL non-disabled peers (the reference group), below-average growth in English was significant for all subgroups except for students with physical disabilities. During this period, the least growth in English was observed for SPED students ( $M=-0.92$ ). In reading, ELL students grew significantly less-than-average ( $M=-0.27$ ), as did SPED students ( $M=-0.20$ ), students with any disabilities ( $M=-0.17$ ), and those with other disabilities ( $M=-0.27$ ).

**Recap of value-added results.** Our analysis of growth differences generally revealed non-significant or below-expected average growth across subgroups and subject areas. The only exceptions were:

- ELL students, who continually demonstrated more relative growth (though lower overall achievement) than their general education peers (reference group) in mathematics between grade 8 to 10, grade 10 to 11/12, and grade 8 to 11/12. This group also demonstrated above-average growth in science from grade 8 to 10.
- Students with any disabilities, who outgrew their reference group peers in reading between grade 10 and grade 11/12.
- Students with cognitive/learning disabilities who showed more relative growth than their general education peers (reference group) in reading and science between grade 10 to 11/12.

- In contrast to the achievement differences in Table 2, the growth differences in Tables 3 and 4 tend to be smaller in magnitude.

### **Discussion**

As the data in this study and others suggest, the achievement gaps between student subgroups is striking. As shown in Table 2, ELL students and students with disabilities had lower EXPLORE, PLAN, and ACT scores in all subject areas than their reference group peers. The growth measures presented in Tables 3 and 4 reveal a different picture. Compared to the reference group, ELL students had consistently higher growth in mathematics and also had higher growth in science between grade 8 and grade 10. Students with any disabilities also experienced above-average growth in reading between grade 10 and grade 11/12, as did students with cognitive/learning disabilities in reading and science. In several cases, the ELL and disability groups demonstrated significantly less growth than the reference group. In other cases, subgroup mean growth was not significantly different than the reference group's mean growth.

In cases where a subgroup's mean growth was significantly above or below average, it is important to understand the practical magnitude of the difference. For example, students with a cognitive/learning disability grew 0.28 points less-than-average in mathematics between grade 8 (EXPLORE) and grade 11/12 (the ACT). From Table 2, we see that the average gain in mathematics from grade 8 to grade 11/12 for the reference group was 5.1 score points (16.5 for EXPLORE, 21.6 for the ACT). This helps place the 0.28 score point growth difference in context, letting us see that the growth difference – while statistically significant – is only about 5% ( $0.28 / 5.10$ ) of the overall average growth, and so is not very large in magnitude.

While many of the growth differences were not statistically significant or statistically significant but small in magnitude, some differences were more striking. In English, ELL



students grew 0.81 points less between grade 10 and grade 11/12, on average, than their general education peers. Similarly, SPED students grew nearly a full point less in English between grade 8 and grade 11/12. In mathematics, ELL students grew nearly  $\frac{3}{4}$  of a point more than their general education peers, whereas students in the other disability subgroup grew over half a point less between grades 8 and 11/12.

Clearly, it is impossible to tell whether a student made significant progress, made no progress, or even regressed using a single test score. Cognizant of this limitation, we measured the growth of individual students over time (using multiple test scores), but without comparing their progress to an annual growth target. Note that AYP targets are required under NCLB to keep students on track for achieving proficiency in the future as envisioned in "growth-to-proficiency" models. Point-in-time achievement differences fail to fully account for the amount of progress a student has made between grade levels. When reviewing progress, measures of growth can be an important addition to point-in-time achievement measures (Yen, 2009). But they need not function as a substitute for achievement, and as Betebenner and Linn (2009, p. 17) have noted, school accountability systems are better served by combining achievement and growth in their assessment of school quality and accountability:

For any approach that takes into account both status and growth, it is important to set ambitious goals, but ones that are reasonable for schools to reach given sufficient effort. One way of determining whether goals are reasonable is to base them on normative information obtained from past experience with high performing schools. It is also possible to use a combination of norm referenced and criterion referenced performance and growth goals.

Aside from uses within accountability systems, both achievement and growth measures have larger roles within school improvement and educators' efforts to use assessment data to target instruction, intervention, and measure program or course effectiveness.

With the design and implementation of new accountability systems on the horizon (either through reauthorization of ESEA, state waivers from NCLB, or other reforms), analyses of the achievement of student subgroups and analyses of their growth in achievement across time are highly relevant. Policymakers, researchers, educators, and parents alike want to know how to best assess the academic progress of all students, including ELL students and those with disabilities. This is a difficult question to answer, but the results of the current study demonstrate the potential of assessing the achievement growth of subgroups such as ELL students and students with disabilities using growth models and value-added analyses.

### **Limitations**

A limitation of the study is that we reported on grade 8 to 12 growth patterns for ELL students and students with disabilities without controlling for covariates like race/ethnicity, low-income status, or type of high school coursework. Thus, the growth differences we observed might have been explained by these other factors. By adjusting/accounting for the effects of race and income, we'd expect to see some changes in the value-added scores for ELL students, but perhaps less so for students with disabilities. As was pointed out earlier, this is because ELL students are a highly diverse population (Crawford, 2004) and many are also racial-ethnic minority and low-income. Another limitation is that our definitions of the disability subgroups were based on student responses to a question of disability type; inaccuracies in the response could affect the study findings. Another limiting factor of the study is that in examining growth differences between grades 8 and 12 (EXPLORE to PLAN, PLAN to the ACT, or EXPLORE to the ACT), the value-added analysis did not take into account (or control for) potential differences among subgroups in the fixed effect for prior achievement. In the EXPLORE to PLAN models, for example, the coefficients for EXPLORE tests might be different for the ELL students and the

SPED students. Such differences could affect the resulting value-added scores and assessment of relative growth differences.

Our study examined growth differences between grades 8 and 12, but the EXPLORE achievement data suggest that growth differences begin well before 8th grade. For example, the mean EXPLORE mathematics score was 16.5 for general education students without a disability (the reference group), but was 13.2 for ELL students, 12.6 for SPED students, and 14.6 for students who reported a disability. Prior studies have suggested that raising the achievement level of students before they enter high school is likely to be more effective in improving college readiness than other interventions, and that lower-achieving 8th graders tend to grow less during high school (Sawyer, 2008). Thus, it is perhaps most important to understand why pre-high school achievement gaps exist and to adopt policies that close those gaps.



## References

- Abedi, J. (2002). Standardized Achievement Tests and English Language Learners: Psychometrics Issues. *Educational Assessment*, 8(3), 231-257.
- Abedi, J. (2004). The No Child Left Behind Act and English Language Learners: Assessment and accountability issues. *Educational Researcher* 33(1), 4-14.
- Abedi, J. (2007). English Language Proficiency Assessment & Accountability under NCLB Title III: An Overview. In J. Abedi (Ed.), *English Language Proficiency Year in the Nation: Current Status and Future Practice*, 3-10. Davis: University of California.
- Allen, J., Bassiri, D., & Noble, J. (2009). *Statistical Properties of Accountability Measures Based on ACT's Educational Planning and Assessment*. (ACT Research Report 2009-1). Iowa City, IA: ACT.
- Angrist, J., & Lang, K. (2002). *How Important Are Classroom Peer Effects? Evidence from Boston's Metco Program*, NBER Working Paper 9263. Retrieved from <http://www.nber.org/papers/w9263>
- Betebenner, D. W., & Linn, R. L. (2009). *Growth in Student Achievement: Issues of Measurement, Longitudinal Data Analysis, and Accountability*. Paper presented at the Exploratory Seminar: Measurement Challenges within the Race to the Top Agenda, organized by Center for K-12 Assessment & Performance Management. Retrieved March 16, 2010 from <http://www.k12center.org/publications.html>.
- Council of Chief State School Officers. (2010). ESEA Reauthorization Principles and Recommendations. Retrieved November 1, 2011 from [http://www.ccsso.org/Documents/2009/ESEA\\_Task\\_Force\\_Policy\\_Statement\\_2010.pdf](http://www.ccsso.org/Documents/2009/ESEA_Task_Force_Policy_Statement_2010.pdf)
- Crawford, J. (2004). *No Child Left Behind: Misguided Approach to School Accountability for English Language Learners*. Retrieved from [http://www.nabe.org/documents/policy\\_legislation/NABE\\_on\\_NCLB.pdf](http://www.nabe.org/documents/policy_legislation/NABE_on_NCLB.pdf)
- Dillon, E. (2007). Labeled: The Students Behind NCLB's 'Disabilites' Designation. *Education Sector*. Retrieved from <http://www.educationsector.org/publications/labeled-students-behind-nclbs-disabilities-designation>
- Eckes, S. E., & Swando, J. (2009). *Special Education Subgroups under NCLB: Issues to Consider*. Teachers College Record 111(11), 2479-2504. New York: Columbia University.
- Federal Education Budget Project (FEBP). (2011a). *Background & Analysis: Individuals with Disabilities Education Act - Cost Impact on Local School Districts*. (June 7). Retrieved from <http://febp.newamerica.net/background-analysis/individuals-disabilities-education-act-cost-impact-local-school-districts>

- Federal Education Budget Project (FEBP), New America Foundation. (2011b). Background & Analysis: Individuals with Disabilities Education Act Overview. (June 7). Retrieved from <http://febp.newamerica.net/background-analysis/individuals-disabilities-education-act-overview>
- Francis, D. J., Rivera, M. O., Lesaux, N., Kieffer, M., & Rivera H. (2006). *Practical Guidelines for the Education of English Language Learners: Research-based Recommendations for Instruction and Academic Interventions*. Portsmouth: RMC Corporation.
- Kane, T.J., & Staiger, D.O. (2003). "Unintended Consequences of Racial Subgroup Rules. In P. E. Peterson & M. R. West (Eds.), *No Child Left Behind? The Politics and Practice of Accountability*, 152-176. Washington, DC: Brookings Institution Press, 2003,
- National Education Association. IDEA / Special Education. (n.d.). Access to a Free, Quality Education for All. Washington, DC. Retrieved from <http://www.nea.org/specialed>
- National Resource Center on ADHD. (n.d). IDEA (The Individuals with Disabilities Education Act). Retrieved from <http://www.help4adhd.org/education/rights/idea>
- Neill, M., Guisbond, L., & Schaeffer, B. (2004). *Failing our Children: How No Child Left Behind Undermines Quality and Equity in Education*. Retrieved from [http://www.fairtest.org/Failing\\_Our\\_Children\\_Report.html](http://www.fairtest.org/Failing_Our_Children_Report.html)
- Office of English Language Acquisition, Language Enhancement, and Academic Achievement for Limited English Proficient Students. (n.d.). *The Growing Numbers of Limited English Proficient Students: 1994/95–2004/05* [poster]. Washington, DC: U.S. Department of Education. Retrieved from [http://www.ncele.gwu.edu/policy/states/reports/statedata/2004LEP/GrowingLEP\\_0405\\_Nov06.pdf](http://www.ncele.gwu.edu/policy/states/reports/statedata/2004LEP/GrowingLEP_0405_Nov06.pdf)
- Payán, R. M., & Nettles, M. T. (n.d.) Current State of English Language Learners in the U.S. K-12 Student Population. Retrieved from [http://www.ets.org/Media/Conferences\\_and\\_Events/pdf/ELLSymposium/ELL\\_factsheet.pdf](http://www.ets.org/Media/Conferences_and_Events/pdf/ELLSymposium/ELL_factsheet.pdf)
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage Publications.
- Rescorla, L., & Rosenthal, A. S. (2004). Growth in Standardized Ability and Achievement Test Scores from 3rd to 10th Grade. *Journal of Educational Psychology*, 96, 85-96.
- Rubin, D.B., Stuart, E.A., & Zanutto, E.I. (2004). A Potential Outcomes View of Value-Added Assessment in Education. *Journal of Educational and Behavioral Statistics*, 29, 103-116.

- Sawyer, R. (2008). Benefits of Additional High School Course Work and Improved Course Performance in Preparing Students for College, *ACT Research Report Series*, 2008-1.
- Short, D. J., & Fitzsimmons, S. (2007). *Double the Work: Challenges and Solutions to Acquiring Language and Academic Literacy for Adolescent English Language Learners—A Report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education. Retrieved from <http://www.all4ed.org/publications/DoubleWork/DoubleWork.pdf>
- The Free Library: Educational Decisions about the Culturally and Linguistically Diverse (CLD) Students under the No Child Left Behind (NCLB) Act. (2008). Retrieved from <http://www.thefreelibrary.com/Educational+decisions+about+the+Culturally+and+Linguistically+Diverse...-a0197721365>
- U.S. Government Accountability Office. (2006). *No Child Left Behind Act: Assistance from Education Could Help States Better Measure Progress of Students with Limited English Proficiency* (GAO-06-815). Washington, DC: Author. Retrieved from <http://www.gao.gov>
- U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2010). *ESEA Blueprint for Reform*, Washington, D.C.
- Yen, W., M. (2009). Measuring Student Growth with Large-Scale Assessments in an Education Accountability System. Retrieved from <http://www.k12center.org/publications.html>.





## Appendix

Table A1

Parameter Estimates for Predicting PLAN scores from EXPLORE

Effect	Subject Area			
	English	Math	Reading	Science
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
<i>Intercept</i> ( $\hat{\beta}_0$ )	1.26 (0.36)	0.26 (0.43)	1.71 (0.42)	5.22 (0.42)
<i>EXPLORE test</i>				
English	0.46 (<0.01)	0.13 (<0.01)	0.27 (<0.01)	0.12 (<0.01)
Mathematics	0.14 (<0.01)	0.50 (<0.01)	0.04 (<0.01)	0.21 (<0.01)
Reading	0.19 (<0.01)	0.09 (<0.01)	0.38 (<0.01)	0.15 (<0.01)
Science	0.14 (<0.01)	0.29 (0.01)	0.20 (0.01)	0.30 (<0.01)
<i>Subject-specific mean EXPLORE</i>	-0.00 (0.02) <sup>ns</sup>	-0.03 (0.02) <sup>ns</sup>	0.04 (0.02) <sup>ns</sup>	-0.03 (0.02) <sup>ns</sup>
<i>Time span</i>	0.08 (0.01)	0.11 (0.01)	0.06 (0.01)	0.06 (0.01)
<i>Variance of intercept, (<math>\hat{\tau}_{00}</math>)</i>	1.63	1.97	1.61	1.30
<i>Intraclass correlation coefficient (ICC)</i>	0.17	0.16	0.14	0.15
<i>Multiple R</i>	0.78	0.75	0.72	0.71

Note: Statistically insignificant coefficients ( $p > 0.05$ ) are marked as 'ns'

Table A2

Parameter Estimates for Predicting ACT scores from PLAN

Effect	Subject Area			
	English	Math	Reading	Science
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
<i>Intercept</i> ( $\hat{\beta}_0$ )	-8.19 (0.21)	-2.67 (0.18)	-3.90 (0.19)	-2.29 (0.22)
<i>PLAN test</i>				
English	0.70 (<0.01)	0.17 (<0.01)	0.40 (<0.01)	0.20 (<0.01)
Mathematics	0.24 (<0.01)	0.63 (<0.01)	0.17 (<0.01)	0.36 (<0.01)
Reading	0.27 (<0.01)	0.01 (<0.01)	0.48 (<0.01)	0.14 (<0.01)
Science	0.12 (<0.01)	0.25 (<0.01)	0.17 (<0.01)	0.32 (<0.01)
<i>Subject-specific mean PLAN</i>	0.18 (0.01)	0.18 (0.01)	0.14 (0.01)	0.20 (0.01)
<i>Time span</i>	0.09 (<0.01)	0.04 (<0.01)	0.04 (<0.01)	0.03 (<0.01)
<i>Variance of intercept, (<math>\hat{\tau}_{00}</math>)</i>	0.56	0.44	0.29	0.31
<i>Intraclass correlation coefficient (ICC)</i>	0.05	0.06	0.02	0.03
<i>Multiple R</i>	0.87	0.87	0.82	0.81

Note: Statistically insignificant coefficients ( $p > 0.05$ ) are marked as 'ns'

Table A3

Parameter Estimates for Predicting ACT scores from EXPLORE

Effect	Subject Area			
	English	Math	Reading	Science
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
<i>Intercept</i> ( $\hat{\beta}_0$ )	-4.46 (0.37)	1.96 (0.37)	0.47 (0.41)	3.16 (0.46)
<i>EXPLORE test</i>				
English	0.57 (<0.01)	0.16 (<0.01)	0.36 (0.01)	0.18 (<0.01)
Mathematics	0.25 (0.01)	0.55 (<0.01)	0.12 (0.01)	0.36 (<0.01)
Reading	0.30 (0.01)	0.07 (<0.01)	0.50 (0.01)	0.18 (<0.01)
Science	0.26 (0.01)	0.33 (0.01)	0.29 (0.01)	0.35 (0.01)
<i>Subject-specific mean EXPLORE</i>	0.05 (0.02)	-0.01 (0.02) <sup>ns</sup>	0.03 (0.03) <sup>ns</sup>	-0.00 (0.03) <sup>ns</sup>
<i>Time span</i>	0.06 (<0.01)	0.02 (<0.01)	0.01 (<0.01)	0.02 (<0.01)
<i>Variance of intercept, (<math>\hat{\tau}_{00}</math>)</i>	4.64	3.46	3.42	2.87
<i>Intraclass correlation coefficient (ICC)</i>	0.24	0.23	0.17	0.19
<i>Multiple R</i>	0.81	0.78	0.76	0.74

Note: Statistically insignificant coefficients ( $p > 0.05$ ) are marked as 'ns'





\* 0 5 0 2 0 1 1 2 0 \*

Rev 1