MIGRATION POLICY INSTITUTE
NATIONAL CENTER ON IMMIGRANT INTEGRATION POLICY

Issue Brief

May 2018 Issue No. 1

A Guide to Finding and Understanding English Learner Data

By Julie Sugarman

EL INSIGHTS

More Data, More Questions

Data on schools, students, and their outcomes have been much in the news. All states are required by law—first by the *No Child Left Behind Act of 2001* (NCLB) and then by its successor, the *Every Student Succeeds Act of 2015* (ESSA)—to publish data on how K-12 students are performing in areas such as reading, math, and graduation rates. These data are critical to decisions about how English Learners (ELs) are educated in the United States. With access to more detailed information, for example, advocates have called attention to gaps between the standardized test scores of White, middle-class students and traditionally underserved groups. Community stakeholders are also now encouraged—even required—to use data when engaging with policymakers on critical questions such as resource allocation, school accountability, and program effectiveness, including recently when the federal government required states to include feedback from stakeholders in their plans for implementing ESSA.

For all of these reasons, it is more important than ever that members of the public understand what types of data are available, where to find them, and how to use them. But with so much data available online, deciding where to start can be daunting. And without an understanding of how data are collected, analyzed, and reported, users run the risk of coming to incorrect conclusions about what they mean. This guide breaks down these questions to help parents, practitioners, policymakers, advocates, and other interested members of the public understand how to find and use student outcome data and information on school and community contexts. It aims to help readers use these data to explain results and appropriately compare schools and districts.

What Data Do School Systems Collect?

The first step to becoming an educated data consumer is knowing exactly what information school systems collect about students and their families.[1] The first four types of data covered in this section describe student characteristics, and the last two focus on schools and districts overall.

School systems collect and use these data for a variety of purposes. These include keeping track of a student's individual needs (such as eligibility for support programs) and flagging individual students as part of certain subgroups to understand differences between group outcomes or to decide how to allot state and federal funds.[2] Student data are also used to develop and improve school policies and programs, and to determine whether schools and districts have met their accountability benchmarks.

Student Background Data: What Do We Know about Students?

This includes data on…	Data can be used to…
▪ Ethnicity/race ▪ Gender ▪ Home language ▪ Place of birth ▪ Prior school enrollment ▪ Eligibility for free or reduced-price lunch	Understand how different groups fare on assessments Understand how the characteristics of the student body change over time

The first information districts collect about students is usually personal and family background data provided on enrollment forms.[3] The school district website is a good place to learn what information is collected when students enroll and how this is done because most districts prominently post their registration forms for new families to download.

Under federal guidelines, districts must ask all families[4] about the language(s) the student and other family members use at home as a first step toward identifying students who require EL services. This data can be collected either using a home-language questionnaire or in a section of the standard enrollment form. When a family responds that a language other than English is used in the household, staff ask the student to complete an English language proficiency test to determine whether he or she needs additional support in learning English.[5]

Importantly, student background data may have limited usefulness if information is not collected in a way that is sensitive to cultural and linguistic differences. For example, students may end up with their name incorrectly listed if it does not conform to the U.S. standard first, middle, and last name fields.[6] Or, teachers may have inaccurate information about the home languages of their students if forms do not distinguish between the languages family members speak and those they read and write. And the categories many districts use to collect information about students' race/ethnicity—usually following U.S. Census questions[7]—may not provide sufficient detail to differentiate among key EL subgroups (such as Black students from Africa and those from the Caribbean). Some school districts work with community partners to ensure that they collect data in linguistically and culturally appropriate ways to make sure they are both accurate and useful.

Can Schools Ask about Immigration Status?

Schools may not ask families questions about their immigration status for school enrollment.[8] The concern is that unauthorized immigrant parents might be wary of enrolling their children if they must reveal their immigration status to people they may see as government employees. And because public schools must take all children regardless of immigration status,[9] the question is not relevant to the enrollment process. Because different districts interpret these requirements in different ways, some do collect information about whether students belong to specific categories, such as refugees and unaccompanied minors, that may need additional support, but it is not a widespread practice. Districts are allowed to ask families for documents that prove they live within the boundaries of the district or school neighborhood and that students are of school age, but they cannot require specific documents (such as state-issued identification or a driver's license) that might discourage unauthorized families from enrolling their children.

That said, districts that receive federal funding from Title III of ESSA (for EL education) must report to the state not only the number of identified ELs in their schools but also the number of

recent immigrant students they serve.[10] To gather this information, the federal government suggests that districts ask for the student's date of birth, place of birth (including country, if not the United States), and where they previously attended school.[11] States then use this information to set aside a portion of their Title III funding for districts that have experienced a notable increase in the number or percentage of recent immigrants.

Data on Services: Who Is Eligible, and Who Uses Them?

This includes data on…	Data can be used to…
▪ English Learner services ▪ Gifted-and-talented programs ▪ Migrant education ▪ Special education ▪ Title I of ESSA (funds for underperforming students in schools with large shares of low-income students)	Identify students eligible for services Count students for federal and state funding allocations Understand how different groups fare on assessments

Data systems generally include fields to track whether students are eligible for different services, what services they receive, and dates when their eligibility status changed. All districts must track whether students are currently ELs, or whether they were previously ELs but have since achieved English proficiency (under ESSA, districts must monitor former ELs for two to four years after achieving proficiency). Districts may or may not track the EL program model or specific types of EL support services that each student receives, although they generally identify students who receive no services because their parents have opted them out. While the type of services received may be obvious for students enrolled in certain schools (for example, one with a whole-school dual language program, meaning all students receive bilingual services), for others it is not clear and not recorded.[12]

Although some districts already keep a record of how many years a student is classified as an EL, more will have to do so going forward. Starting in 2018, all districts that receive ESSA Title III funding will be required to record these data to identify long-term ELs.[13] This designation, as well as markers for other EL subgroups (such as students with interrupted formal education or ELs with disabilities) may become part of district data systems when these funding rules come into effect.

Why Look at Data on Services Not Specific to Learning English?

Districts also keep track of other services that students receive that are not related to their EL status. These include special education services (including individualized education programs, known as IEPs, and 504 plans),[14] gifted-and-talented programs, and other federal programs such as Title I and migrant education.[15] Districts also track whether students are eligible for free or reduced-price lunch based on their parents' income level. Although it is an imperfect indicator, eligibility for free or reduced-price meals is frequently used by educators and researchers to represent students' socioeconomic status or poverty level.[16] By examining information about these other services, data users can get a more well-rounded picture of what support ELs are (or are not) getting.

The participation of ELs in other programs—especially special education and gifted-and-talented services—at similar rates to non-ELs is frequently tracked as evidence of equity. Specifically, ELs should not be over-referred or under-referred to these services, which can happen if screening procedures are not linguistically or culturally sensitive. On the other hand, it is not unusual for

higher shares of ELs to qualify for Title I support and migrant education programming than non-ELs given the aims of these programs and the number of ELs whose families live in poverty and/or engage in seasonal, migratory work.

Assessment Data: How Are Test Scores Reported?

This includes data on…	Data can be used to…
▪ English language proficiency ▪ State standardized tests (English language arts, math, science, social studies) ▪ High school assessments (SAT, AP tests) ▪ Other district and school standardized tests	Hold schools accountable to state standards Identify students eligible for support services Group students for instruction (for example, into English as a second language classes based on level of English proficiency)

Districts collect a vast amount of data on how students perform on standardized tests, such as those given to ELs to monitor English language proficiency. Districts usually receive this information from the companies that grades the tests in the form of electronic score reports. These reports might indicate how a student performed on each question[17] but more typically show results for sets of academic skills (known as subscales). Reports that show student-by-student data are only available to school and district administrators, and are used to create individual report cards for parents and average scores (without student names) for release to the public.

The ACCESS for ELLs[18]—an English language proficiency assessment given annually to ELs in 39 states and U.S. territories—is a good example of how this works. Its score reports include information on how each student performed on four subscales: speaking, listening, reading, and writing. Reports also include composite scores for oral language (listening and speaking), literacy (reading and writing), and comprehension (listening and reading), as well as an overall composite score that combines all four language domains. The reports do not show the number of correct answers out of the total number of questions (a raw score); instead, they show how students performed on a scale from 100 to 600 (a scale score). Doing this allows data users to more easily see students' progress from one year to the next. The report also assigns students a proficiency level on a scale of 1.0 to 6.0 (with corresponding descriptors from "entering" to "reaching" English proficiency) for each language domain and composite to help data users connect the test scores to what a student can do with language.[19] Schools and districts generally use these scores to identify students as ELs and to group students into English as a second language (ESL) classes by proficiency level.

State standardized tests of English language arts (ELA), math, and other subjects are usually reported in a similar way. Score reports provide a composite score and subscale scores for each section of the test, using raw and/or scale scores with corresponding descriptions (often "below basic," "basic," "proficient," and "advanced"). Scores can also be reported as a percentile, which describes what percent of other test-takers scored below the student in question. District databases may also track how a student's performance has changed from year to year (a growth score).

General Student Data: What Other Measures Do Schools Track?

This includes data on…	Data can be used to…
▪ Attendance ▪ Discipline ▪ Grade retention (for students who repeat a grade) ▪ Graduation status or reason for withdrawing from the district	Hold schools accountable to state standards (i.e., graduation rate, school quality indicators) Ensure equitable treatment of minority students

Districts keep a record of some student outcomes that go beyond test scores. These data help them follow the progress of individual students as well as of subgroups (such as ELs and low-income students) that may face challenges to succeeding academically.

When tracking student attendance, schools generally distinguish between excused absences (those verified by a parent for a district-approved reason) and unexcused absences. Districts also keep track of whether (and for how many days) students receive in-school or out-of-school suspensions, as well as whether they are expelled (and for what reason). If a student is held back and must repeat a grade, this is also noted in their record (known as "grade retention"). These data have recently been recognized as important indicators of school quality and student success, and states use them in accountability systems alongside data on academic achievement.[20]

Districts also track when students graduate or, if they leave the district without graduating, whether they dropped out or transferred to another district.[21] This classification is important for calculating graduation rates, which are an accountability measure for high schools and districts.[22]

School and Administrative Data: Context Matters

Type	This includes data on…	Data can be used to…
School Climate	▪ Family and community involvement ▪ Parent and student satisfaction ▪ Safety	Ensure equitable treatment of students and families Make program improvements
Administrative	▪ Budget and expenditures ▪ Teacher certifications and licensure	Ensure schools and districts follow spending rules Help communities understand how human and material resources are allocated

System-level data provide a useful look at the broader context in which ELs are educated. While this is important information, it is less likely that users of this guide will need to access and analyze these data themselves. At a very basic level, these data provide information about the school environment and how schools use resources. Data are collected through surveys and administrative record keeping, and summary reports of the findings are typically available on district websites.

Who Has Access to Data, and What Can They Do with It?

Ten to 15 years ago, vast amounts of student data were stored in paper files. Now, most districts store some or all of their data electronically. Many districts also used to keep the various types of data in separate spreadsheets or data repositories, making it cumbersome to access and use them. Today, districts tend to use software packages that allow them to gather student data in one place, with password-protected access for administrators and teachers. Students and their parents may also have access to their own data (or selected portions). Authorized personnel can also generate datasets from these repositories for researchers to use.

What Can Teachers and School Administrators Do?

Most educational data software packages allow users to create data reports, letting them run queries that display and compare the categories of data that matter most to them. Teachers generally have access to every piece of data about the students in their classes, including information from previous school years. Administrators usually have access to that sort of data for all students in the school or district. Together, teachers and administrators use this information to monitor student progress and make decisions about how individuals and groups of students are taught.

Educators are able to run queries to answer questions such as: how many EL students are enrolled in each homeroom? How do student outcomes compare within a grade across teachers? And which student subgroups are making adequate progress in certain subject areas (such as math or science)? Some types of software can also answer more complicated queries, such as: how many Hispanic, low-income students passed the state standardized math test despite scoring below proficient on the literacy subscore of the English language proficiency assessment? This question is more complex because the software must calculate the number of students who have two or more characteristics (known as cross tabulation). And because it looks at multiple characteristics of individual students, it requires access to individual student data.

What Can Members of the Public Do?

Depending on district or state policy, certain individual data users (such as education researchers) may be given a detailed but de-identified dataset—one with all personal information such as names, addresses, and telephone numbers removed—so they, too, can do complex statistical analyses.[23] But not everyone who is interested in student data can be given this level of access. For most other stakeholders—including state administrators, policymakers, and the general public—states have created public-use databases. These online data dashboards allow users to access a wide variety of data points, and to create charts and graphs using data visualization tools. Although state data dashboards generally do not enable users to crosstabulate data (to look at, for example, ELs who are low-income), such capabilities will be included in redesigned dashboards based on ESSA requirements that some data points be publicly available. This increases the range of questions the public can ask and answer without access to internal district databases.[24]

Some limitations on what the public can access are put in place to protect student privacy. Even data that include no names, addresses, or other personal information may be traced back to an individual child if other characteristics are described in enough detail and the groups they are part of are small enough (such as, an EL in 3rd grade at X school who receives reduced-price lunches and scores

highly in math). For this reason, schools, districts, and states do not make available average scores or counts of students when there are fewer than a certain number of students in the group. Data users may also find their access limited if there are too few students in a group to draw reasonable conclusions about what the average score for that group means. When users encounter these kinds of limitations, they will often see an asterisk (*) or dash (–) in their data tables.[25]

State-Level Data Sources

For members of the public, the most comprehensive and up-to-date source for student data is usually the website of the state department of education (often in the section on accountability). Through the state report cards and data dashboards found there—which are made up of data submitted by school districts—users can produce reports that show data at the state, district, or school level. A state may have one webpage visitors can use to view a variety of data types, or it may have several different pages for different types of data.[26] Because data dashboards include information about all public-school students, not just ELs, users can also make comparisons with other groups of students.

EL Data All States Are Required to Share (or Will Be Soon)

States include the following types of data in their report cards or data dashboard. The items with an asterisk (*) may or may not already be included, but will become required under ESSA starting with data for the 2017-18 school year.

- Number of identified ELs
- Number and percent of ELs achieving English proficiency*
- Number and percent of ELs taking state standardized tests (ELA, math, science, etc.)
- Number and percent of recently arrived ELs who are exempt from the ELA test*
- EL subgroup achievement levels on:
  - ELA, math, and science assessments
  - four-year graduation rate
  - optional five- or six-year graduation rate*
  - non-academic school quality or student success indicator*
  - postsecondary enrollment rates*
- Progress toward long-term goals in English language proficiency and on all the subgroup achievement indicators above*
- Some data categories collected for the Civil Rights Data Collection (for example, suspensions, expulsions, and school safety data)*

Under ESSA, states are required to publish annual report cards with information on state, district, and school performance and progress. By requiring additional categories of data at each of these three levels, states will have to make more data available than was required under NCLB.[27] States have until December 31, 2018 (with a possible one-year extension) to redesign their report cards to accommodate the new details.[28]

Almost all of the indicators in the box above must be reported for each state, school district, and individual school, including charter schools (but not private schools). At all three levels, data users can compare ELs to other subgroups (for example, students grouped by race/ethnicity, gender, and special education status).[29]

There are also some types of data that states must report under Title III of ESSA, although it is not clear whether those must be included in the state report card in addition to the Consolidated State Performance Report (CSPR) states submit to the U.S. Department of Education each year. These data include the number and percent of ELs with disabilities, the number and percent of long-term ELs, the number and percent of reclassified former ELs who meet content standards, and disaggregated data on the share of ELs with disabilities making progress toward English proficiency and meeting content standards.

EL Data that Some States Share

Among the other types of information that some states make available are:

- Percent of ELs at each English language proficiency level, often disaggregated by grade
- Percent of EL students with interrupted formal education
- Language(s) other than English spoken at home for all students or for ELs specifically
- Disaggregated counts of ELs by grade, race, country of origin, district (and school), free or reduced-price lunch status, or other variables
- Count of students enrolled in each EL program type (e.g., dual language, ESL)
- Student achievement by EL subgroup (students with interrupted formal education, long-term ELs, recent immigrants) and EL program type

As state data systems continue to evolve, more data may become available to the public. In the future, states might publish data on the number of students who receive the state Seal of Biliteracy,[30] data collected from formative assessments,[31] and data on EL and former-EL outcomes in postsecondary institutions. Many states have or are developing statewide longitudinal data systems that match up data on students' interactions with early childhood, K-12, postsecondary, and workforce development systems to give a more well-rounded, long-term picture of the supports students get and their outcomes over time.[32] The public can use these systems to examine the long-term impacts of programs and initiatives as students move in and out of various state educational systems.

For members of the public who want to dig deeper into EL data, additional information can also be found in state reports that are published on an annual or one-off basis. For example, Illinois and Washington State both produce detailed annual reports on their bilingual education programs that users can access on their state websites.[33] Some districts also publish their own analyses of EL data or the results of EL program evaluations conducted by outside organizations.

National Data Sources

For users interested in national EL data or in comparing EL data across states, there are numerous useful sources. Those described in this section all provide free access to the public via their websites, with data dashboards that allow users to run queries and create reports. Some also publish reports that summarize the data. Like state data dashboards, they provide the number or percent of students with certain characteristics, some name districts or schools but never individual teachers or students, and most allow users to compare the characteristics and progress of ELs to other subgroups and to students as a whole.

Listed below are four data sources produced by the federal government. The first three present information reported by state departments of education (data that they, in turn, have collected from schools and districts). The fourth is the U.S. Census, which collects data from U.S. households using yearly and ten-year surveys.

<table>
<tr><th colspan="2">Consolidated State Performance Report (CSPR)</th></tr>
<tr><th>What Is It?</th><th>What Data Are Included?</th></tr>
<tr><td rowspan="2">Every year, states report certain information required by ESSA to the federal government. It usually takes two years before the report for a school year is available on the U.S. Department of Education website.

Although ESSA was passed in December 2015, its new data requirements didn't take effect immediately. Reports for school year 2017–18 are likely to be the first to follow ESSA requirements, and they are likely to be available online in 2019–20. As this happens, the types of data available in Ed Data Express are likely to grow.</td><td>ED Data Express
https://eddataexpress.ed.gov
Site can be used to create graphs, tables, and maps for state data, including:
the number of identified ELs,
number of students receiving Title III services,
number of monitored former ELs,
number of students speaking the top five languages in each state,
number and percent of students progressing toward and attaining English proficiency,
EL subgroup (and former EL subgroup) passing rates on state standardized tests, and
graduation rate data.</td></tr>
<tr><td>CSPR home page
www2.ed.gov/admins/lead/account/consolidated/index.html
Site can be used to find published state reports with the information included in ED Data Express, plus information on native language assessments, number of recent immigrants, and EL-related professional development.</td></tr>
</table>

Common Core of Data (CCD)	
What Is It?	**What Data Are Included?**
Not to be confused with the Common Core State Standards, these data are collected annually by the National Center for Education Statistics. Like CSPR, data are collected from states, but some can be disaggregated down to the district and school levels. These data are adjusted to enable users to make comparisons across locations and time, but this means that CCD counts may differ from other data sources.	***Elementary / Secondary Information System (ElSi)*** https://nces.ed.gov/ccd/elsi/ This system provides detailed student enrollment counts and staff counts for states, counties, districts, and schools. Counts of identified ELs are only available for states and districts.
	National Center for Education Statistics (NCES) reports https://nces.ed.gov/surveys/annualreports/ Users have access to different publications, such as the *Condition of Education,* that summarize key data such as the number of public school students participating in EL programs by state, as well as national EL enrollment by grade, home language, race/ethnicity, and disability status.

Civil Rights Data Collection (CRDC)	
What Is It?	**What Data Are Included?**
Every two years, the U.S. Department of Education collects school and district reports on topics related to education access and equity.	***CRDC website*** http://ocrdata.ed.gov Data can be broken down by EL status, race, and disability status, and they include information on: ▪ enrollment in early childhood programs, ▪ participation in advanced coursework (such as AP), ▪ discipline, and ▪ staffing.

U.S. Census	
What Is It?	**What Data Are Included?**
The U.S. Census Bureau conducts two major surveys of the U.S. population: the decennial census and the yearly American Community Survey. These collect information about the countries of origin, home languages, limited English proficiency (LEP), and school attendance/ completion of all residents of the United States. It is important to note that for LEP, survey respondents are asked to describe their own level of English proficiency (and parents describe that of their children), so these data may not match a state's count of EL students.	***U.S. Census Bureau's American FactFinder*** http://factfinder.census.gov The American FactFinder website brings together data tables from all of the U.S. Census surveys with an easy-to-use guided search feature.
	Migration Policy Institute's Migration Data Hub www.migrationpolicy.org/programs/migration-data-hub The Migration Data Hub offers a variety of data tools that allow users to explore key census data related to immigrants by state and nationwide.

Other national sources draw on the data collected by CSPR, CCD, CRDC, and the U.S. Census Bureau to report on state and national trends. These include resources from the U.S. Department of Education's Office of English Language Acquisition. Our Nation's English Learners (www2.ed.gov/datastory/el-characteristics) is an interactive website that includes charts on the number of ELs by race/ethnicity; eligibility for other federal programs; enrollment in districts with high, medium, or low concentrations of ELs; enrollment by school district (with maps showing growth); and languages spoken. For data highlights, Fast Facts (https://ncela.ed.gov/fast-facts) offers two-page summaries of key EL trends.

Some national and international tests, described in the table below, also produce information that can be analyzed to better understand student characteristics and outcomes. But because participation in these tests is voluntary, there may be significant limitations in how representative their results are of ELs. Data users should also be cautious about how they interpret data from these tests because differences in which students take the test in each country can affect the results; in addition, they are not aligned to the curriculum of any U.S. state.[34]

Source Name	What Is It?	What Can It Say about ELs?
National Assessment of Educational Progress (NAEP) https://nces.ed.gov/nationsreportcard/	▪ Given every year to a sample of U.S. students in 4th, 8th, and 12th grades ▪ Produces outcomes data for reading, math, and other subjects ▪ Collects data on student background, school characteristics, curriculum and instruction, and students' educational experiences	Data are available for current and former ELs, and in some cases, data users can explore student results based on school characteristics (e.g., percent of students in bilingual or ESL programs, availability of certain kinds of staff to work with ELs).
Programme for International Student Assessment (PISA) www.oecd.org/pisa	▪ Given every three years to a sample of fifteen-year-old students throughout the world ▪ Covers a variety of academic subjects	Data users can compare outcomes across countries for immigrant students, native-born students with foreign-born parents, and native-born students with native-born parents.
Trends in International Mathematics and Science Study (TIMSS) https://nces.ed.gov/timss	▪ Given every four years to a sample of students in 4th and 8th grades and the last year of high school throughout the world	Data users can compare outcomes across countries based on whether students speak the language of the test at home.
Progress in International Reading Literacy Study (PIRLS) https://nces.ed.gov/surveys/pirls/	▪ Given every five years to a sample of students in 4th grade throughout the world	Data users can compare outcomes across countries based on whether students speak the language of the test at home.

Comparing Data: A Cautionary Tale

The spread of education data has enormous benefits for families, community members, and student advocates who want to understand how ELs fare and provide feedback to school decisionmakers. However, without an experienced eye it can be easy to inadvertently misinterpret data that are this complex.[35] This section focuses on three challenges users may face when comparing groups of ELs to other subgroups and across states, data sources, and time.

Can I Compare Group A with Group B?

Unlike most other student characteristics that are self-reported (such as gender and race/ethnicity), EL status is more complex. States have different methods for identifying students as ELs, using different English language proficiency assessments, cutoff scores, or other criteria.[36] In fact, until ESSA was passed in 2015, there was no requirement that states have a standard procedure for determining EL status across districts and schools. As states put these procedures in place, data users will be able to have more confidence when comparing data from districts and schools within the same state. But across states, there will still be some inconsistencies; depending on the rules each state sets, a student might be considered an EL in one state but not in another. National sources that compare EL counts across states (such as NCES) use state definitions and do not adjust data to conform to any one definition of who is an EL.

On top of this, EL status is not a permanent label. Students no longer count as ELs once they score high enough on an English proficiency test to meet state criteria to exit EL status. This means that, by definition, the EL subgroup does not include students who score at grade level on standardized academic assessments in English.[37] The fluid nature of the EL subgroup makes tracking a cohort of students over time complicated. Unlike most other subgroups, the number of ELs generally decreases from elementary through middle and high school, as more students exit EL status than new immigrant children arrive. In other words, the ELA and math achievement of a school's Black student subgroup is likely to be spread across a range of low, medium, and high performance,[38] and this spread is likely to remain even as students move in and out of the school; the students currently identified ELs, on the other hand, are always likely to score in the lower end of the achievement range on academic tests because students who have developed the levels of English proficiency needed to score highly on such tests no longer carry the EL label.

While the ESSA requirement that states standardize EL criteria may make comparisons within a state more reliable, the same law may have inadvertently made comparisons across states more difficult. This is largely a question of how states count former ELs and newcomers. Under NCLB, states were allowed to include data on former ELs in the EL category for up to two years after they were reclassified as fluent English proficient. But under ESSA, there is more variation; some states do not include these reclassified students in their EL subgroup at all, while others include them for two or four years. Similarly, states are now allowed to choose whether to test and how to report the ELA and math scores of newcomer students in their first two years in U.S. schools. Because former ELs are likely to score higher than current ELs on state academic content assessments, and most recently arrived ELs are unlikely to score proficient on such tests, the inclusion or exclusion of these students could have a significant effect on overall counts of ELs meeting state goals.

Why Don't these Numbers Match?

In addition to differences in how states define and count the outcomes of ELs, there are important differences between datasets collected for different reasons. In other words, even if every school system defined ELs using the same criteria, datasets put together for different reporting purposes may use different methods and thus have different totals for the same group. This can make comparisons across data sources challenging.

For example, it might seem like it should be straightforward to find out how many ELs there are in each state. Yet different sources will often have different totals. As the table below shows, data users may come across at least seven different figures when looking for the number of ELs in California in school year 2014–15.

Type of Count	What Makes It Complex?	Example: California
Enrollment Counts	When counting students, states can use a single count date (e.g., all those enrolled on October 1), the average enrollment for a few count dates or the whole year, or the average daily attendance for the whole year. When states report EL enrollment to the federal government under Title III of ESSA, they use a single count on October 1. But enrollment counts for other purposes may use a different day or method, resulting in a different total for the same year.	The California Department of Education reports on its websites that in school year 2014–15, there were 1,392,263 ELs (based on its spring language census). But on its CSPR for the same year, it reports 1,397,841 ELs (based on its October 1 count).
Test Participation	State report cards show the number of students in each subgroup (including ELs) that participate in statewide assessments. Data users should take care not to use these figures as total population counts because they only include students in the tested grades (grades 3 through 8, plus one grade in high school).	On its 2014–15 state accountability report card, California reported having 975,726 ELs enrolled and 957,922 tested in a table on ELA test participation. These figures are far lower than the total count of nearly 1.4 million ELs in all grades discussed above.
Students Identified vs. Served	On its annual CSPR, each state reports the number of identified ELs as well as the number of ELs served by Title III. The latter number excludes ELs in districts that choose not to receive Title III funds or whose parents opted them out of services.	California's 2014–15 CSPR reports 1,397,841 total ELs, but 1,354,691 who received services in a Title III program.
NCES Definitions	The annual NCES Common Core of Data survey asks states to report the number of ELs enrolled in EL instruction, whether funded by Title III or not. These numbers are included in the ElSi data repository and the NCES *Digest of Education Statistics*. But the *Digest* table on participation in EL programs includes a note that it excludes students in certain types of school environments (presumably to ensure comparability across states).	According to ElSi, the number of ELs in California in 2014-15 was 1,392,295, but the *Digest of Education Statistics* shows 1,390,316 ELs.

Note: To navigate the ElSi tableGenerator to find the data on California ELs mentioned in the last row of this table, select "State" for Table Row, then select the years "2014–15," and under "Table Columns" select "Enrollments—Students in Special Programs—Limited English Proficient (LEP) / English Language Learners (ELL) [District]." Finally, select "Filter by individual state(s)" and choose California.

Although data users should always try to choose the dataset most appropriate for the question at hand, this is especially important when making comparisons. When comparing the number of ELs enrolled in different states (or districts in different states), it is better to use one of the federal data sources described in the previous section, such as the *Digest of Education Statistics*, because these systems take steps to ensure that state data are comparable. For most sources—whether federal, state, or local—users can also find information about how the data were collected and which types of students are included, often in the title or caption of a table or in a footnote.

Can I Compare One Year to Another?

There are three main issues to be aware of when comparing data from one point in time to data from another. The first is being able to tell the difference between data that show school improvement versus student progress. NCLB and ESSA both emphasize the idea that schools should improve from year to year, as measured by test scores. To see whether this is happening, data users look at test scores for the same grade across multiple years (for example, 3rd graders in 2014–15, 3rd graders in 2015–16, and 3rd graders in 2016–17—all of these students are in the same grade, but there are different individuals in each group). This is different from the type of analysis done to understand how the outcomes of a group of the same students have changed over time—a longitudinal analysis (for example, 3rd graders in 2015–16 who became 4th graders in 2016–17, and so on). Data users should be aware of which type of growth they are most interested in and choose their data accordingly.[39]

Second, it is important to understand what tools were used to collect data, as this may change from year to year. Recent changes in English language proficiency assessments and state content area assessments will make comparisons over time difficult. Most states have adopted new or revised English language proficiency assessments, and some may adjust what is considered a passing score—reflecting revisions to the widely used WIDA consortium assessments[40] and/or efforts to meet the ESSA requirement that each state standardize its process for determining EL status. Many states also revised their academic content assessments in ELA and math after adopting the Common Core State Standards published in 2010. Between 2011 and 2018, some states replaced or significantly revised their ELA and math assessments more than once. Data systems that provide test results for multiple years should include an explanation of whether the scores are comparable over time or not, as it is often not appropriate to directly compare outcomes from different assessments. Additionally, when looking at ELs' results on academic content area assessments over time, data users should be aware of whether, between those years, the state changed how it defined ELs. For example, if the English language proficiency test cutoff score is raised so that students need a higher level of English proficiency to exit EL status, the EL subgroup would likely have more students scoring at or close to proficient on ELA and math; by contrast, if the cutoff score is lowered, the EL subgroup would likely have fewer high scorers in ELA and math.

A final challenge is understanding whether other factors—including some far beyond the reach of school systems—may be at play. Dramatic changes to the demographics, geographic boundaries, or governance of a neighborhood, city, or region may make it inappropriate to compare test scores and other data over time (unless the data user is specifically interested in seeing the effect of these context changes). For example, a hurricane or wildfire that permanently displaces some residents could change the level of poverty or the racial/ethnic make-up of a city. Additionally, districts sometimes combine or break apart, such as in Tennessee when Shelby County Schools merged with Memphis

City Schools in 2013. Data on the state report card website are marked "Shelby County (pre-merger)" for 2012–13 and "Shelby County" for 2013–14 on, and it would not be appropriate to compare the results from the 46,600-student district in 2012–13 to those of the 150,000-student district the following year.[41] The composition of student populations may also change as charter schools proliferate. For example, between 2004–05 and 2014–15, District of Columbia Public Schools enrollment dropped from 62,300 (representing 81 percent of public school students) to 46,200 (57 percent of public school students).[42] In order to accurately compare these years, a data user would need to verify that the composition of the student body in the traditional public schools stayed the same over those ten years (for example, in terms of race/ethnicity or poverty).

Conclusion

The last 15 years have been a particularly dynamic time for school systems and data sharing. Many systems have significantly transformed, responding to evolving ideas about appropriate learning standards, preferences for different forms of school governance, and—of huge importance for ELs—improved understanding of language learning and what kind of support students need to close achievement gaps. Student outcome data have taken on a fundamental role in prioritizing how these systems change, with school administrators increasingly sharing data and seeking input from the communities they serve. While developing an understanding of how to best access and use the wealth of data available may take some time, building these skills will prove crucial in helping families and other members of the public engage with decisionmakers and support evidence-based policies and practices to improve the education of ELs.

Endnotes

1 This brief uses the term "data collection" to describe information that is intentionally gathered for inclusion in a student's permanent school record. As more and more school activities happen online, computer systems may automatically collect and store additional data, including student beliefs and behaviors that are captured as part of everyday interactions with technology, thus increasing what potentially becomes long-term student data. See Audrey Watters, "Student Data vs Student Privacy," Modern Learners, April 24, 2014, https://modernlearners.com/student-data-vs-student-privacy.

2 For example, each year, school districts report to their state the number of identified English Learners (ELs) enrolled in district schools on October 1. The state then uses these numbers to decide how to allocate federal Title III funding for ELs. See Julie Sugarman, *Funding an Equitable Education for English Learners in the United States* (Washington, DC: Migration Policy Institute, 2016), www.migrationpolicy.org/research/funding-equitable-education-english-learners-united-states.

3 In some localities, students new to a district register at a central district office and in others, at their neighborhood school using forms provided by the district. Charter schools typically act as their own district, and students register at the school they will attend. This section uses "the district" to refer to the entity collecting registration information, regardless of who exactly is doing the collecting.

4 All families are asked these questions so as not to discriminate based on perceived race, ethnicity, or language proficiency.

5 U.S. Department of Education, *Tools and Resources for Identifying All English Learners* (Washington, DC: U.S. Department of Education, 2016), www2.ed.gov/about/offices/list/oela/english-learner-toolkit/chap1.pdf.

6 Jason Greenberg Motamedi, Zafreen Jaffery, Allyson Hagen, and Sun Young Yoon, *Getting it Right: Reference Guides for Registering Students with Non-English Names*, 2nd Edition (Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, and Regional Educational Laboratory Northwest, 2017), https://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_2016158.pdf.

7 U.S. Department of Education, "New Race and Ethnicity Guidance for the Collection of Federal Education Data," updated August 2008, www2.ed.gov/policy/rschstat/guid/raceethnicity/index.html.

8 Catherine E. Lhamon, Philip H. Rosenfelt, and Jocelyn Samuels, "Dear Colleague Letter: School Enrollment Procedures" (letter, U.S. Department of Justice and U.S. Department of Education, Washington, DC, May 8, 2014), www2.ed.gov/about/offices/list/ocr/letters/colleague-201405.pdf.

9 This right was established in the 1982 Supreme Court case *Plyler v. Doe*. See ibid.

10 Title III of ESSA says the federal government will provide supplementary funds to states based on the number of EL students and recent immigrants in each state (it does not, however, specify the amount of funds; that is done by Congress each year). Recently arrived immigrant students are defined as students who are 1) between the ages of 3 and 21; 2) were not born in any U.S. state, Washington, DC, or Puerto Rico; and 3) have attended school in the United States for a cumulative total of less than three academic years. See U.S. Department of Education, "Additional Questions & Answers on Enrolling New Immigrant Students" (fact sheet, U.S. Department of Education, Washington, DC, December 2014), www2.ed.gov/policy/rights/guid/unaccompanied-children-2.pdf.

11 Ibid.

12 Title III reporting requires states to indicate which of ten types of EL services its schools provide overall, but they do not have to report how many schools offer each type. See U.S. Department of Education, "Consolidated State Performance Reports," accessed July 25, 2017, www2.ed.gov/admins/lead/account/consolidated/index.html.

13 ESSA leaves it to the states to decide what constitutes "long term," but most practitioners consider students to be long-term ELs if they are identified as ELs for six or more years.

14 An IEP, or Individualized Education Program, is a detailed, written document the *Individuals with Disabilities Education Act* requires schools to create for students with one or more of 13 psychologist-diagnosed disabilities. Schools may use a 504 plan to adjust the learning environment to meet the specific needs of a student with any learning difference or disability. See Understood, "The Difference Between IEPs and 504 Plans," accessed July 25, 2017, www.understood.org/en/school-learning/special-services/504-plan/the-difference-between-ieps-and-504-plans.

15 The federal government provides Title I funds to states to support underperforming students in districts with large numbers or high shares of low-income students. The federal Migrant Education Program provides support to students who transfer schools across district or state lines because their parents move to find work in one of a specified set of fields (e.g., farm work). See Sugarman, *Funding an Equitable Education.*

16 Although an imperfect indicator, educators and researchers frequently use students' eligibility for free or reduced-price meals as a proxy to understand poverty levels in student populations. There has been some criticism of combining data on both eligibility levels (free and reduced-price) because the circumstances of a student who is eligible for free meals could differ significantly from those of a student eligible for reduced-price meals. Eligibility for reduced-price meals is set at 185 percent of the poverty level; for free meals, it is 130 percent. This means, for example, that in 2017–18, when the federal poverty level was $24,600 for a family of four, such a family could make as much as $45,510 in most states and still be eligible for reduced-price meals. See Office of the Federal Register, "Child Nutrition Programs: Income Eligibility Guidelines," updated April 10, 2017, www.federalregister.gov/documents/2017/04/10/2017-07043/child-nutrition-programs-income-eligibility-guidelines.

17 Many test companies do not allow non-test takers to see the questions, either during or after the test, so the company can reuse them from year to year. Usually, test makers release a selection of questions each year that are being "retired," as an example of the types of questions the test includes.

18 ACCESS for ELLs stands for Assessing Comprehension and Communication in English State-to-State for English Language Learners.

19 WIDA Consortium, *Spring 2017 Interpretive Guide for Score Reports. Kindergarten–Grade 12* (Madison, WI: Board of Regents of the University of Wisconsin System, 2017), www.wida.us/get.aspx?id=25.

20 Erika Hall, *Identifying a School Quality/Student Success Indicator for ESSA: Requirements and Considerations* (Washington, DC: Council of Chief State School Officers, 2017), www.ccsso.org/sites/default/files/2017-10/CCSSOIdentifyingSchoolQualityStudentSuccessIndicator1242017.pdf.

21 The U.S. Department of Education released regulations in 2008 requiring states to begin using a four-year adjusted cohort graduation rate that is a consistent formula across states. Using this system, graduation rates are calculated based on the number of students in a ninth-grade cohort who graduate within four years. Students who drop out of school altogether are included in the cohort as nongraduates, but students who transfer to another district, emigrate to another country, transfer to a prison or juvenile facility, or die are removed from the cohort and not counted as nongraduates. See U.S. Department of Education, *Every Student Succeeds Act High School Graduation Rate Non-Regulatory Guidance* (Washington, DC: U.S. Department of Education, 2017), www2.ed.gov/policy/elsec/leg/essa/essagradrateguidance.pdf.

22 Because the graduation rate is used for accountability purposes, some districts have come under suspicion of taking advantage of lax recordkeeping to hide their dropouts by miscoding them as having left the district for a reason that does not count against the graduation rate. See, for example, Sarah Karp and Becky Vevea, "Emanuel Touts Bogus Graduation Rate," WBEZ 91.5 Chicago, June 10, 2015, www.wbez.org/shows/wbez-news/emanuel-touts-bogus-graduation-rate/7891382c-6069-4bdb-ad7c-a34fd67895e2.

23 District and state policies may restrict access to even de-identified datasets. Individual student data are protected under the *Family Educational Rights and Privacy Act* (FERPA), which gives parents the right not to have personal information about them or their children disclosed. For more information, see the U.S. Department of Education, "Protecting Student Privacy," accessed September 26, 2017, https://studentprivacy.ed.gov/.

24 The Leadership Conference Education Fund, "Data Provisions in the Every Student Succeeds Act," updated April 2016, http://civilrightsdocs.info/pdf/education/ESSA-Data-Fact-Sheet.pdf.

25 If an average is calculated for a small handful of students, the average might be pulled dramatically up or down by one unusual score. The average would thus not represent what a typical student can do. ESSA addresses this by having states calculate a minimum subgroup size. For more information on reporting-group sizes, see Marilyn Seastrom, *Best Practices for Determining Subgroup Size in Accountability Systems While Protecting Personally Identifiable Student Information* (Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2017), https://nces.ed.gov/pubs2017/2017147.pdf.

26 As an example of the latter case, Massachusetts has a page for "District Analysis Review Tools" (www.doe.mass.edu/dart/) in addition to one for "School and District Profiles" (http://profiles.doe.mass.edu/), and elsewhere, users can download reports with data and analysis of Massachusetts Comprehensive Assessment System results (www.doe.mass.edu/mcas/results.html).

27 One analysis estimated that an average-sized state would produce more than 306,300 discrete data points annually based on ESSA requirements. See memorandum from Penn Hill Group to Council of Chief State School Officers, *State Report Card Requirements*, January 26, 2016, www.ccsso.org/sites/default/files/2017-10/ESSAStateReportCardRequirementsMemo01262016.pdf.

28 U.S. Department of Education, *Every Student Succeeds Act State and Local Report Cards Non-Regulatory Guidance* (Washington, DC: U.S. Department of Education, 2017), www2.ed.gov/policy/elsec/leg/essa/essastatereportcard.pdf.

29 Some data repositories provide a comparison of ELs to non-ELs, whereas others only make it possible to compare ELs to all students—the latter is not a true comparison group since ELs are also counted among all students. An average score for non-ELs can be calculated if the user has the total number of students, the number of ELs, and the average score for ELs.

30 A Seal of Biliteracy is a designation schools can award at high school graduation to students who demonstrate proficiency in English and another language. For more information, see Seal of Biliteracy, "State Laws Regarding the Seal of Biliteracy," accessed April 2, 2018, http://sealofbiliteracy.org/.

31 Formative assessments are tests used to periodically gauge a student's progress toward year-end goals.

32 Zeke Perez, Jr., *Policy Analysis: Examining SLDS Development and Utility* (Denver: Educational Commission of the States, 2017), www.ecs.org/ec-content/uploads/Examining_SLDS_Development_and_Utility.pdf.

33 Illinois State Board of Education, "Bilingual Education Statistics," accessed November 17, 2017, www.isbe.net/Pages/Bilingual-Education-Statistics.aspx; State of Washington, Office of Superintendent of Public Instruction, "Annual Reports to the Washington State Legislature," updated February 16, 2017, http://www.k12.wa.us/MigrantBilingual/BilingualProgram/AnnualReports.aspx.

34 Martin Carnoy and Richard Rothstein, *What Do International Tests Really Show About U.S. Student Performance?* (Washington, DC: Economic Policy Institute, 2013), www.epi.org/publication/us-student-performance-testing.

35 For additional guidance on interpreting educational data, see Janie Tankard Carnock, *Seeing Clearly: Five Lenses to Bring English Learner Data into Focus* (Washington, DC: New America, 2017), www.newamerica.org/education-policy/policy-papers/seeing-clearly; Patricia A. Lauer, *A Policymaker's Primer on Education Research: How to Understand, Evaluate, and Use It* (Denver: Mid-Continent Research for Education and Learning and the Education Commission of the States, 2004), http://files.eric.ed.gov/fulltext/ED518626.pdf; John W. Young, "Ensuring Valid Content Tests for English Language Learners" (brief, R & D Connections No. 8, Educational Testing Service, Princeton, NJ, December 2008), www.ets.org/Media/Research/pdf/RD_Connections8.pdf.

36 Robert Linquanti, H. Gary Cook, Alison L. Bailey, and Rita MacDonald, *Moving Toward a More Common Definition of English Learner: Collected Guidance for States and Multi-State Assessment Consortia* (Washington, DC: Council of Chief State School Officers, 2016), www.ccsso.org/sites/default/files/2017-10/MoreCommonDefinition-Final_0.pdf.

37 This is the case unless reclassified former ELs are included in the EL subgroup, but even those students would be removed from the subgroup after two to four years. See William M. Saunders and David J. Marcelletti, "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, no. 2 (2013): 139–56.

38 Statisticians refer to the typical model of student achievement as a bell-shaped curve, with a few students doing very poorly, most students scoring in mid-range, and a few students doing very well.

39 EdTrust, "Individual Student Growth: An Overview" (fact sheet, EdTrust, Washington, DC, accessed May 9, 2018), https://edtrust.org/students-cant-wait/individual-student-growth/.

40 Corey Mitchell, "Is a New English-Proficiency Test Too Hard? Educators and Experts Debate," *Education Week*, August 4, 2017, www.edweek.org/ew/articles/2017/08/04/is-a-new-english-proficiency-test-too-hard.html.

41 Tennessee Department of Education, "State Report Card," accessed November 3, 2017, www.tn.gov/education/data/report-card.html.

42 NCES, "ElSi tableGenerator," accessed August 28, 2017, https://nces.ed.gov/ccd/elsi/tableGenerator.aspx.

About the Author

Julie Sugarman is a Policy Analyst at the Migration Policy Institute (MPI) National Center on Immigrant Integration Policy, where she focuses on issues related to immigrant and English Learner (EL) students in elementary and secondary schools. Among her areas of focus: policies, funding mechanisms, and district- and school-level practices that support high-quality instructional services for these youth, as well as the particular needs of immigrant and refugee students who first enter U.S. schools at the middle and high school levels.

Dr. Sugarman came to MPI from the Center for Applied Linguistics (CAL), where she specialized in the evaluation of educational programs for language learners and in dual language/two-way immersion programs. At CAL, she directed comprehensive program evaluations of instruction for ELs in K-12 and contributed to numerous research and evaluation projects, including studies of biliteracy development in two-way immersion programs and the evaluation of the STARTALK program that funds teacher training programs and language instruction for students in grades K-16 in critical languages.

Dr. Sugarman earned a B.A. in anthropology and French from Bryn Mawr College, an M.A. in anthropology from the University of Virginia, and a Ph.D. in second language education and culture from the University of Maryland, College Park.

Acknowledgments

This brief was supported by the Carnegie Corporation of New York as the first in a series of EL Insights on topics relating to K-12 education and English Learners. The author thanks Margie McHugh and Lisa Tabaku for their valuable comments on earlier drafts, Lauren Shaw for editing this brief, and Sara Staedicke for its design and layout.

Suggested citation: Sugarman, Julie. 2018. *A Guide to Finding and Understanding English Learner Data*. Washington, DC: Migration Policy Institute.

The Migration Policy Institute (MPI) is an independent, nonpartisan, nonprofit think tank dedicated to the study of the movement of people worldwide. The Institute provides analysis, development, and evaluation of migration and refugee policies at the local, national, and international levels. It aims to meet the rising demand for pragmatic responses to the challenges and opportunities that migration presents in an ever more integrated world.

WWW.MIGRATIONPOLICY.ORG/INTEGRATION

MPI
MIGRATION POLICY INSTITUTE

1400 16th Street, NW, Suite 300, Washington, DC 20036
202-266-1940 (t) | 202-266-1900 (f)