WISCONSIN DEPARTMENT OF PUBLIC INSTRUCTION

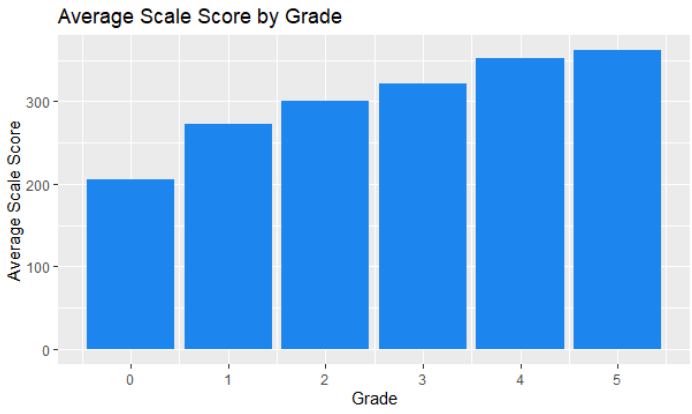
Analyzing EL Growth Data in WISEdash for Districts

Introduction

The Wisconsin Department of Public Instruction (DPI) is committed to every student receiving an equitable education — regardless of learning environment — so all students have the knowledge, skills, and habits to graduate college- and career-ready. To better support our English Learners (ELs), DPI has been adding additional data elements related to the ACCESS for ELLs assessment to WISEdash for Districts.

When it comes to educational data, EL data is some of the most complicated that we deal with. We have two growth measures which have different meanings and different uses, and what these growth measures convey is strongly dependent on the underlying circumstances.

EL growth to proficiency has both high annual variability at the individual student level as well as a generally logarithmic trend across many ELs over time. Growth tends to be more rapid at lower grades and proficiency levels, while leveling off to a plateau at higher grades and proficiency levels. Creating growth metrics that can accommodate these features and create meaningful data upon which to base decisions is thus challenging.



A complicating factor is that ELs who are successful stop being ELs, as they become proficient and exit EL status. Thus the impact of successful EL programming is a reduction of the number of successful ELs later in the program. It is quite easy to take the 30,000 foot view of EL growth data and miss critical details such as this.

Because of this, ELs are a unique demographic cohort which sees significant change in the population of the cohort as it moves up grades. When examining growth data, this can't be forgotten. When comparing schools, districts, or programs, or comparing year-on-year within a

cohort, it's critical to understand exactly who the underlying students are, and how they are distributed across grade levels.

EL Growth Measures in WISEDash for Districts

The two growth indicators we include in WISEdash for Districts are the On-Track measure and Student Growth Percentiles (SGPs). They both relate to student growth, but are complementary measures that each provide different information about student growth towards English proficiency, as measured by the ACCESS for ELLs. This document is designed to explain what each is and how each can be used to inform programming decisions and inform student supports and interventions.

To access these data, you will need to have *student detail analyst* permissions in WISEdash for Districts. If you do not have this level of access, you will need to work with your district administrators to get it. The data in WISEdash for Districts are largely the same as those found in the district Student Information System (SIS). However, they include some additional calculated values, along with students' educational history from around the state. For an overview of the EL related data elements, please see the English Learner Data in WISEdash document.

Student Growth Percentiles

To calculate annual Student Growth Percentiles (SGPs), all ELs in the state with an identical grade and ELP level are pooled, and their Overall Composite Scale Score growth from the previous year is ranked from 1 (lowest of the group) to 99 (highest of the group). A student with a SGP of 50 thus represents the average growth of a student in that ELP level and grade within the state of Wisconsin. SGPs are divided into Low (1-35), Typical (36-65), and High (66-99) growth categories.

The average SGP across all ELs in a school is a reliable indicator of how that school is doing in supporting their ELs. While individual students' SGP may vary by year, as their growth varies and as the other students in that SGP pool vary, school average SGPs tend to be more stable.[1] Average school SGP is the ELP Progress indicator used for federal (ESSA) accountability reports.

SGPs provide a relative measure of growth as compared to very similar students in the state. This allows districts to understand what normal growth looks like, and compare groups of students to this normal growth. As they differentiate based on grade and ELP level, SGPs can be more meaningful than just looking at scale score growth or proficiency level growth, as those vary widely across grades and ELP levels.

With these benefits in mind, SGPs do not tell you if a student is on track to becoming proficient or if they have become proficient. A student can have made low growth for years and not be on-track to become proficient, but they might have very high growth one year resulting in a high SGP. Or a

[1] For sufficiently large EL populations that are sufficiently stable.

student may have made very high growth for several years and be on-track, but make low growth one year resulting in a low SGP.

Thus care must be taken when analyzing SGPs. They are meaningful measures of how well a school is doing in aggregate, and they show how a student is growing relative to their peers, but they don't tell the full story for individual students.

On-Track Measure

Under ESSA, states are required to set long-term goals for ELs. Wisconsin chose to create an on-track measure for ELs' growth to proficiency, with an annual goal of increasing the percent of students on track to become proficient.[2]

To create an on-track measure, we started by calculating the average amount of time it has historically taken ELs to reach full English proficiency, based on their starting English Language Proficiency (ELP) level and grade. These historical averages are used to set a target exit year for each new EL. We then find the Scale Score which aligns with an Overall Composite ELP of 5.0 in the expected grade of exit, and set this as their target exit score.

Each year the student's current Overall Composite Scale Score on the ACCESS for ELLs is subtracted from their target exit score and divided by the number of years left to their expected exit year. This results in a target scale score increase for the next year.[3] Their actual scale score increase is then compared against that target score increase to determine whether they are on track to become proficient by their target exit year.

A major benefit of this annual recalculation is that it takes into account past performance when setting the growth target for the next year. If a student needs to grow 120 points to reach their target exit score within five years, their first year target will be **24 points**. (120/5 = 24).

- If they grow **20 points** the first year, their goal for the next year will be **25 points**.
    - (120 - 20 = 100; 100/4 = 25)
- If they grow **40 points** the first year, their goal for the next year will be **20 points**.
    - (120 - 40 = 80; 80/4 = 20)

The annual recalculation of growth targets within this model allows us to accommodate the tendency for ELs to grow faster at younger ages and lower ELP levels and slower at older ages and higher ELP levels. This allows us to fairly set growth targets as a student becomes more proficient and their growth slows, as we incorporate past growth into future goals.

On Track Status Definitions are as follows:

[2] https://dpi.wi.gov/esea/wisconsin-consolidated-state-plan (p. 29)
[3] Ibid, (p. 31-32)

1. Baseline Year: First year of ACCESS data
2. Not Enough Data: Missing prior year test score
3. Progress Goal Met: Met or exceeded scale score growth target
4. Achieved Proficiency: Met or exceeded ELP 5.0 at or before expected exit year/grade
5. Additional Support Needed: Did not meet scale score growth target
6. Further Support Needed to Exit: Scored below ELP 5.0 at or past the expected exit year/grade
7. Proficient Past Target Year: Scored ELP 5.0 or greater past expected exit year/grade

Creating Ad-Hoc Cohorts

While the most common disaggregation of ELs will be by demographics such as special education status, gender, home language, etc., it is possible to create ad-hoc cohorts of students within WISEdash for Districts. This allows for grouping students based on other characteristics, such as ESL program type, intervention tier or strategy, newcomer status, long-term EL status, etc.

If you are interested in creating such cohorts for analysis, please take special note of the section below titled *Minimum N or Cell Size* before doing so. Disaggregating ELs into groups that do not contain enough students may produce an invalid analysis.

For more information on creating ad-hoc cohorts, please see WISEdash for Districts > Topics > Resources > Group Management User Guide.

Using SGPs

As SGPs are a normative measure, the average SGP of a school can be compared directly with the SGP of other schools. In general, a school with a higher average SGP has higher growth in their EL population.[4]

Within a district, this may allow for the identification of program strength or a lack thereof in the various schools. Such a comparison can identify areas where additional staffing or professional development is needed, or identify successful program implementations which can be leveraged to increase the performance of other programs in the district.

Mean SGPs are ideal for comparing different cohorts of students[5], to determine if there is a systematic difference in their growth to proficiency. Such an analysis can suggest cohorts of students who are in need of additional or differentiated support.

It is important to remember that SGPs are calculated annually from the statewide pool of students with the same grade and ELP level. This means that the students in each SGP cohort change each

[4] Assuming a stable EL population in each, and in consideration of a minimum N size as noted below.
[5] If they meet the minimum N size. See below.

year, and thus changes in mean SGPs over time likely represent a mixture of programmatic changes along with changes in the underlying cohort.

Using the On-Track Measure

Unlike SGPs, the On-Track calculation is not normed for similar students, and we see significant changes in the distribution of students who fall into the *Progress Goal Met* and *Additional Support Needed* categories as grade levels increase. As might be expected, students consistently meeting their progress goals become proficient and drop out of the calculation, as they are no longer ELs and no longer take ACCESS.

While new ELs dilute the pool of remaining students somewhat, in general, most schools will see a higher percent of students falling into the *Additional Support Needed* and *Further Support Needed to Exit* categories in higher grades. This is substantially different than with SGPs, where growth is normed on a 1-99 scale, regardless of how much underlying growth within that pool of students there actually is.

Because the On-Track calculation isn't normed, comparing schools based on the number or percent of students meeting progress goals is inadvisable. The distribution of students across grades in each school will likely provide as much or more variability than any of the underlying causes.

Where the On-Track calculation shines is in determining the individual support needs for students. However, care should be taken to understand where within a growth category a student lies. If they are close to falling into another category, the supports they need may differ from those needed if they are not. In particular, students in the category *Progress Goal Met* who are only meeting their goal by 10 or less scale score points are likely to fall out of that category in future years as language demands increase.

Early Intervention

Student language growth is fastest at lower grades, and grades 1-3 represent the area where small interventions can make the largest long-term impact on ELs' growth to proficiency. For this reason it is critical to design processes to identify students in these grades who are in need of support, to ensure that support is given, and to evaluate the effectiveness of the support over time.

Across the state, we see increasing numbers of students not meeting their progress goals as they progress from grade 1 to grade 3. Eventually, a large percent of these students end up in the population of middle-schoolers struggling to reach proficiency. While it's clear that we need an increased focus on supporting language development in middle school, much of the growth stagnation we see there begins in elementary school.

On-Track Category: Progress Goals Met

While on its face it would seem that if a student is meeting their progress goals they likely don't need additional support, it is important to understand how much they are exceeding their goal by. Dig into the data to find the Target Score Change and the Actual Score Change, and see by how much students in this category have exceeded their goals.

The table below is a rough rule of thumb relating ELs' progress above and beyond their target growth score to the potential need for an intervention to stay on track to become proficient.

Grade	Exceeded Target By	Intervention Need
1	More than 50	None
1	30 to 50	Unlikely
1	10 to 30	Possible
1	Less than 10	Likely
2	More than 30	None
2	10 to 30	Unlikely
2	Less than 10	Likely
3	More than 25	None
3	10 to 25	Unlikely
3	Less than 10	Likely
4th and up	More than 20	None
4th and up	10 to 20	Unlikely
4th and up	Less than 10	Likely

On-Track Category: Additional Support Needed

After 1st grade, state-wide, 15%-20% of our ELs are already falling into the *Additional Support Needed* category. These students, if not supported at this critical time when their language growth tends to be the most rapid, will most likely struggle to get back on track to become proficient.

More than 40% of our ELs completing 2nd grade are not on-track to become proficient. Around half of these students were in the *Progress Goals Met* category the previous year. Thus the need for scrutiny of that category, rather than assuming that students who fall in it will be fine in the future.

By the following year, this number rises to over 50% of ELs who are completing 3rd grade.

It is important to note that these are *remaining* ELs, as some have been reclassified over this period. Still, a large percent of the ELs needing additional support after 3rd grade are the students who, in 6th and 7th grade, remain in the program past their expected exit date.

Middle-School Challenges

It is not uncommon (although problematic) for middle-school students to plateau or even show slight negative growth. This is due to many factors, not least the extensive social and physical changes students undergo at this age. Compounding these factors is a significant increase in the language demands of our college- and career-ready standards, along with a change to a new ACCESS for ELLs grade cluster designed to test these increased language demands.

While we can mitigate many of these challenges with early interventions, middle-school remains an area where the vast majority of the ELs in the state could use extra support. Program analysis and a deep dive into differences in demographic groups may be necessary to begin to understand the needs of ELs in middle school. Transition planning may also be very beneficial, to ensure a continuity of services for students changing schools.

Reclassification

On-Track Category: Further Support Needed to Exit

This category is composed of students who are at or past their expected exit year and who have not yet reached an Overall Composite of 5.0.

If a student who falls into this category has scored 4.5-4.9, they may be ready to exit. Consider administering the Multiple Indicator Protocol (MIP) described in Chapter 15 of the EL Policy Handbook to provide a secondary measure of proficiency, to determine whether or not reclassification would be warranted.

Most students are English proficient somewhere within this range. We have chosen 5.0 as our mandatory exit point because students scoring this high are almost certainly proficient. (Students scoring 5.0 and higher tend to slightly outperform native English speakers on the Forward ELA and Math tests.)

If a student in the *Further Support Needed to Exit* category is below an Overall Composite of 4.5, they are in need of significant interventions. If the student appears to be fluent on the surface, ensure that you are differentiating between conversational/social English and Academic English.

Many students have extensive opportunities to hone their conversational English skills, but limited changes to hone their Academic English skills.

Students in this category should be strongly considered for entry into the district MLSS, if not already there.

ELs with Disabilities

In Wisconsin, the largest demographic of students who remain EL after 6 years are students with disabilities. Ensuring that these students are supported for both needs is critical to their success. If an analysis shows extensive differences between the growth of ELs with and without disabilities, the first consideration should be whether or not all of their needs are being coherently and comprehensively met.

Students with disabilities and students who are ELs have very well defined legal rights to access an education equal to that of their peers. ELs with disabilities retain all rights from membership in both protected groups, and this includes all of the services and supports each classification allows the student. It is not legal to deny services to a student under the assumption that the other program will fill that need. Students with disabilities who are ELs must be enrolled in a language instruction education program, and ELs with disabilities must have an IEP or 504 Plan and appropriate Special Education services to meet their needs[6].

Language plans and IEP/504 Plans are not legally interchangeable, but they should not be viewed as separate pieces of information about the student. An EL does not walk into a Special Education classroom and stop being an EL, nor does a student with disabilities suddenly lose those when they walk into an ESL classroom. Both the Language Instruction Education Plan (LIEP) and the IEP/504 Plan should describe the umbrella of services the student requires, while being reflective of the legal requirements of each document.

A disability may cause a student to struggle in a related domain. This is a large reason we created the MIP process to allow students to exit between 4.5 and 4.9, as that tool only requires a demonstration of proficiency in three domains. This creates a structure to identify a student who is mostly proficient, save for a disability related challenge in one domain.

Zeroing In on Root Causes

Our growth metrics use the Overall Composite score, but it's important to remember that it's made up of the four domains of Listening, Reading, Writing, and Speaking. The exact supports a student needs should be determined after looking at their performance in each of the underlying domains. A student may only need extra support in Writing if the other three domains are strong, or they may need extra support in all four domains.

[6] https://dpi.wi.gov/sped/program/english-learners

This is also true if looking for the root cause of a cohort of students with a low SGP. Be sure to examine their individual domain scores and look for trends. This may help you identify targeted improvements you can make to your curriculum and lesson plans, to better support a particular domain that most students are struggling with. If almost all of the students are demonstrating low Speaking growth, e.g., you may be able to work on this by building 'turn and talk' activities into lesson plans, oral question and answer time, a weekly debate, etc.

Reading and Writing each account for 35% of the Overall Composite, and thus 70% of the growth metrics together. Small improvements in either can have a significant impact on a student's growth trajectory, as well as a school's mean SGP. In particular, Writing is the lowest domain score for most students, and thus it represents the area with the most potential for growth for most students.[7]

Students have more frequent opportunities to use their Oral abilities than their Literacy ones, and thus students' Oral abilities tend to grow faster than their Literacy abilities. But this is not universal, and it's important to dive into the distribution of each student's domain scores to ensure that you're targeting the right areas for each student.

Cautions and Words of Wisdom

Minimum N or Cell Size

In any analysis, please be cautious about interpreting the results for small numbers of students. In general, the individual student variability significantly reduces the reliability of any measure when under about 10 students. When working on state and federal accountability measures, we use a minimum N size of 20.

Thus robust, meaningful interpretations can be made using data sets and cohorts which consist of 20 or more students. Less robust and less meaningful interpretations may be made as the student count gets closer to 10. For less than 10 students, an analysis may be somewhat informative, but should be tempered with the understanding that the reliability is questionable.

One method of boosting the N count is to use more than one year of data. Provided the student population hasn't significantly changed, two or more years of data can be combined for analysis, and this will provide more reliable information than a single year alone will.

[7] Most students in Wisconsin struggle with Writing, whether or not they are an EL. A promising opportunity for growth is a closer tie between ELA and EL staff to include better supports for ELs in writing lessons, as well as to better support native English speakers who are also struggling with Writing.

Relationship between SGP and On-Track

The SGP measure is a current-year measure, and it isn't significantly impacted by the historical growth trajectories of the students in the cohort. It's common to see an individual student's SGP change significantly from one year to the next.

The On-Track measure is recalculated annually, and as it's based on the remaining scale score points to proficiency, it takes into account the past growth of students. The longer a student's growth history, generally the more stable their On-Track categorization becomes. (Unless they are consistently right at the inflection point of being on-track.)

Older students tend to grow less than younger students, which means the spread of scale score points in SGP cohorts in higher grades tends to be lower. Thus a small change in scale score growth in higher grades can be enough to change their SGP category, but not enough to impact the On-Track calculation.

Consider a student who falls behind for several years, making far less progress than desired. If that student receives an intervention which helps them move back into Typical SGP growth, they may never be able to be on-track to becoming proficient. Indeed, for students who fall too far behind, even a High SGP might not be sufficient for them to fall into the *Progress Goals Met* category of the On Track measure.

Similarly, a student who makes growth far above average the first few years may have their target scale score growth for future years set so low that they could fall into the Low SGP category and still be on track to become proficient. If they only need 40 more scale score points to exit within the next 4 years, the On-Track target is 10 scale score points per year. In most SGP cohorts, such growth would fall into the Low category, yet it could still reach *Progress Goals Met* for the on-track measure.

When analyzing EL growth data, it's imperative that both measures be looked at, and in particular, how they change across grades and across ELP levels needs to be examined. Unique cohorts of students (ELs at 4.0 and higher transitioning off of intensive EL supports, e.g.) may have somewhat undesirable growth metrics (Low SGP) but still be On-Track to exit and indicative of a successful EL program.

Using EL Growth Data for Teacher Evaluation

There are numerous and significant challenges with using EL data for teacher evaluation, to the extent that it's generally not advisable.

The first major barrier is the extensive lag between when the data is analyzed and when the corresponding instruction occurred. As growth requires two years of assessment data, assessment happens in the winter, and our growth metrics are generally not available until August, at best growth might be related to instruction between 7 and 15 months ago. The later the growth data is analyzed, the further in the past the instruction happened. If doing an analysis in June, for example, the data would relate to instruction which happened between 17 and 25 months ago!

Unlike content area instruction, language instruction has incredibly varied implementations. Language instruction can happen in environments as different as sheltered instruction for newcomers, push-in or pull-out models, dual language instruction or bilingual instruction, among many others. The speed of growth towards proficiency differs based on the program model, so care must be taken to separate the impacts of the program model from the impacts of instruction.

It's important to understand that this is not a bad thing! The goals of a sheltered instruction classroom are much different than the goals of a bilingual program, and we can see some of that difference in the variance of typical growth to proficiency between the two programs. Language instruction does not need to be designed solely to get a student to English proficiency as fast as possible.

Students can also learn English in many places outside of a language instruction program. They may have language instruction support from other teachers, aids and paraprofessionals, caregivers, their peers, siblings, and family. Understanding the contributions of all of these individuals is needed to parse out what a language instructor is contributing to each student's language growth. These potentially widespread contributions to student learning is not something usually seen in academic content areas.

The non-teaching duties of language instructors must also be taken into consideration. Many are tasked with parent outreach, translation and interpretation, supporting limited English proficient parents in IEP and other parental meetings, etc. It is not uncommon for these duties to infringe on time supporting ELs, and this must be accounted for. In addition, it's necessary to understand how many students the instructor supports, and how much time they get with each one. This is often much more complicated than a classroom teacher with a set schedule, given the high variance in language instruction program types.

Complicating this is whether or not they share a language with the students they teach. It's not unreasonable to think that an instructor who shares a language (and/or culture) with their students may be better able to teach their students than one who does not. When looking at teacher effectiveness, knowing the student makeup and native languages spoken combined with the language abilities and cultural understanding of the instructor is critical.

As diverse as the home and school environments can be for ELs, so too can be their backgrounds. The growth of new ELs in the lower grades differs dramatically from those in the upper grades. Refugees dealing with trauma and trying to adapt to a new life have significantly more challenges

than second generation students anchored in a stable community. Socioeconomic status impacts language growth, as does a related access to enrichment activities and parental support. Understanding the background of an instructor's ELs informs the context in which the language instructor is delivering services.

In sum, accounting for all of these (and other) variables in order to use language growth as a teacher evaluation method is nearly impossible. Districts are advised to choose other methods to evaluate effectiveness.

Wisconsin Department of Public Instruction
Office of Educational Accountability
125 S. Webster Street, P.O. Box 7841
Madison, WI 53707-7841
https://dpi.wi.gov/osa-oea

May, 2023