



State Support Network
Partnering for School Improvement

Guidebook on State Approaches to Incorporating English Learners Into Title I Accountability Systems

Pete Goldschmidt
Robert Linqanti
Diane August
Rachel Slama

June 2020

1000 Thomas Jefferson Street, NW
Washington, DC 20007-3835
202.403.5000
Statesupportnetwork@air.org

This document was produced by American Institutes for Research under U.S. Department of Education (Department) contract number ED-ESE-15- A-0006/0001. The content of this document does not necessarily reflect the views or policies of the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This document may contain the views and recommendations of various subject matter experts as well as hypertext links, contact addresses and websites to information created and maintained by other public and private organizations. The inclusion of links to resources and examples do not reflect their importance, nor is it intended to represent or be an endorsement by the Department of any views expressed, or materials provided. The inclusion of information in this document is not intended to reflect a determination by the Department that any specific aspect of an accountability system mentioned in the document necessarily complies with statutory requirements. The Department does not control or guarantee the accuracy, relevance, timeliness, or completeness of any outside information included in this document. Authorization to reproduce this document in whole or in part for educational purposes is granted.

This document may contain the views and recommendations of various subject matter experts as well as hypertext links, contact addresses, and websites to information created and maintained by other public and private organizations. The inclusion of information in this document is not intended to reflect a determination by the Department that any specific aspect of an accountability system mentioned in the document necessarily complies with statutory requirements. Additionally, the content of this document and opinion expressed in this document do not necessarily reflect the views or positions or policies of the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. The Department does not control or guarantee the accuracy, relevance, timeliness, or completeness of any information included in these materials. In addition, this document does not place any specific requirements on any agencies, organizations, or other entities.

This document does not address requirements under laws enforced by the Office for Civil Rights at the U.S. Department of Education, including Title VI of the Civil Rights Act of 1964 (Title VI), section 504 of the Rehabilitation Act of 1973 (section 504), and Title II of the Americans with Disabilities Act (Title II). This document also does not include information about recipients' affirmative obligations to English learners (ELs) under Title VI. For additional information about legal requirements under Title VI, section 504, or Title II, please see the website of the Office for Civil Rights (OCR), U.S. Department of Education at www.ed.gov/OCR.

Contents

	Page
Acknowledgments.....	i
Introduction.....	1
Intent and Organization of This Document	1
Section 1: Research Related to Incorporating ELs Into State Accountability Systems.....	3
ELP Goals and Measurements of Interim Progress	3
The ELP Indicator.....	7
Translating Student-Level Progress Into a School Performance Indicator	8
Weighting the ELP Indicator in the Accountability System.....	9
Section 2: ELP Goals and Measurements of Interim Progress	11
ESEA Requirements	11
Challenges for States	12
Considerations: ELP Goals and Measurements of Interim Progress	23
Section 3: Constructing and Evaluating the ELP Indicator	24
ESEA Requirements	24
Challenges for States	24
Technical Approaches.....	25
Considerations: The ELP Indicator.....	31
Section 4: Translating Student-Level Progress Into a School-Level ELP Indicator	33
ESEA Requirements	33
Challenges for States	33
Technical Approaches.....	33
Section 5: Integrating the ELP Indicator Into the State Accountability System for Annual Meaningful Differentiation	35
ESEA Requirements	35
Challenges for States	35
Technical Approaches.....	36
Next Steps for SEAs	43
Glossary	44
References.....	45

Appendix A. ELP Community of Practice Goals, Activities, and Membership.....	A-1
Appendix B. ELP CoP Subject Matter Expert Profiles	B-1
Appendix C. Additional Select ESEA Statutory and Regulatory Requirements Related to English Learners	C-1

Tables

	Page
Table 1. Mock Example: Average ELP Performance-Level Gains by Initial ELP Level and Years in Program.....	14
Table 2. Mock Example: Average ELP Performance Level Gains by Initial ELP Level, Years in Program, Exit Status.....	15
Table 3. Delaware EL ACCESS Growth Targets—Annual Calculation Method	17
Table 4. ADE Proxy Exit Criteria.....	22
Table 5. Student-Level ELP Growth Index Score by Student Outcome (Rules for Years Up to and Including the Designated Attainment Year)	28
Table 6. Student-Level ELP Growth Index Score by Student Outcome (Rules for Years After the Designated Attainment Year)	29
Table 7. Sample Display of Percentage on Track by Student Group, Domain, and Combined	30
Table 8. Sample Display of Median Growth Percentiles by Student Group, Domain, and Combined.....	31
Table 9. Impact of Normalization on the ELP Indicator Using an Equal Weight Check.....	37
Table 10. Impact of Different Point Assignment to Same Model	38
Table 11. Assessment Area and Weighting in Example State Accountability System	40
Table A1. ELP CoP Members	A-1
Table A2. Number and Percentage of ELs Served by ELP CoP Member States (from lowest to highest total number of ELs).....	A-4

Figures

	Page
Figure 1. Aggregate Percentage of Students Gaining One or More Proficiency Levels (2005–07).....	4
Figure 2. Comparison of ELP Growth Targets and Actual Performance (Rapid Early Growth, Slower Later Growth).....	18
Figure 3. Sample Worksheet for Documenting Progress to Target Approach	21
Figure 4. Example of Effect of Linear (Constant) ELP Growth Expectations Over Time.....	26
Figure 5. Effect of Value-Added Model Expectations for ELP Growth Over Time	27
Figure 6. Calculating Index Scores in the Delaware Model	42

Acknowledgments

The State Support Network would like to acknowledge the support of Rachel Slama (Massachusetts Institute of Technology), Pete Goldschmidt (California State University, Northridge), Robert Linqanti (WestEd), and Diane August (American Institutes for Research) in preparing this guidebook.

The State Support Network thanks Gary Cook from the WIDA Consortium (formerly World-Class Instructional Design and Assessment), Mark Hansen and Margaret Ho from the English Language Proficiency Assessment for the 21st Century (ELPA21), and Delia Pompa from the Migration Policy Institute for their contributions to this report. The State Support Network also thanks Brenda Calderon, Danielle Smith, and Mario Nuñez¹ from the U.S. Department of Education and Fen Chou from the Council of Chief State School Officers for their ongoing guidance. The State Support Network would also like to acknowledge the contributions of the 11 state members of the State Support Network community of practice for their contributions to the examples detailed in this guidebook and related discussion (please see Appendix A for a list of community of practice members).

Finally, the State Support Network thanks the state educational agency (SEA) staff who contributed to the state examples featured throughout this guidebook. All SEA examples featured in this guidebook were reviewed and vetted by SEA staff and describe the state's approved consolidated state plan² at the time of publication. States interested in implementing similar accountability practices and strategies to those described in this document must follow U.S. Department of Education (Department) requirements related to amendments to state plans.

¹ Please note that at the time of publication, Mario Nuñez is no longer an employee of the U.S. Department of Education.

² Consolidated state plans developed and approved under the Elementary and Secondary Education Act of 1965 (ESEA), as amended by the Every Student Succeeds Act (ESSA)(Section § 8302).

Introduction

The Elementary and Secondary Education Act of 1965 (ESEA), as amended by the Every Student Succeeds Act (ESSA),³ ushered in important changes related to the inclusion of English learners (ELs) in state accountability systems. These changes included moving state English language proficiency (ELP) accountability provisions from Title III to Title I and tracking progress for ELs at the school level rather than the district level. These changes in statute place new attention on peer review of the ELP assessment and necessitate collaboration among state staff leading Title I and Title III implementation. The collective expertise of these state educators is required to set up robust assessment, accountability, and monitoring systems that determine and report on ELs' English language proficiency progress and attainment and provide support for ELs who are not making sufficient progress toward meeting state ELP standards. In short, the ESEA now has an increased focus on ELs and the accountability system requirements related to EL progress.

Improved data collection and analysis over the past decade has helped state educational agencies (SEAs) examine longitudinally, and better understand the pace at which ELs develop English language proficiency. Many states are currently leveraging these data to develop empirically grounded expectations for meeting the requirements of the ESEA. As such, consolidated state plans under the ESEA include more attention to the heterogeneity among ELs than was possible in previous accountability systems.

Yet, states still confront challenges in implementing consolidated ESEA state plans, including provisions on ELP accountability. For example, since recent content and ELP assessments are geared to measure new standards and contain substantial innovations in design, construction, and administration, research evidence suggests that results and growth patterns across old and new assessments are not generalizable and should be used and interpreted with caution (Linn, 2000). Therefore, although it is possible to use historical data spanning two assessments with significant shifts in design, analyzing results using data collected from the newer ELP assessments can more reliably reflect current data trends. Most relevant analyses can be undertaken with two years of results but should be continuously monitored (Goldschmidt, 2020).

Intent and Organization of This Document

This guidebook provides an overview of technical approaches for designing accountability systems for ELs' ELP progress and attainment under the ESEA. The approaches detailed in this guidebook draw on empirical research from published reports, current guidance from subject matter experts, and examples⁴ from SEA consolidated state plans featured by SEA community of practice (CoP) members who participated in the State Support Network CoP. The approaches detailed in this guidebook surfaced during three years of monthly meetings of the Department-sponsored ELP CoP and during a Department-sponsored convening in 2018 that brought state teams together to discuss EL assessment and accountability issues and solutions. Appendix A provides additional information related to ELP CoP goals, activities, and membership.

³ In this document, the term "ESEA" will be used to refer to the ESEA, as amended by ESSA.

⁴ The examples are of technical approaches for designing accountability systems for ELs' ELP progress and attainment under the ESEA.

This guidebook focuses on the technical approaches in state accountability systems related to ELP goals and measurements of interim progress,⁵ the ELP indicator,⁶ and annual meaningful differentiation of school performance related to ELP progress and attainment.⁷ The intention of this guidebook is to stimulate ideas and discussions within and among SEAs as they continue to implement and refine accountability systems with respect to ELs under the ESEA. The guidebook is organized into five sections:

- Section 1 summarizes the empirical research base related to incorporating ELs into state accountability systems.
- Section 2 focuses on ELP goals and measurements of interim progress.
- Section 3 focuses on the ELP indicator.
- Section 4 focuses on technical approaches related to translating student-level progress targets into school-level indicator results.
- Section 5 focuses on incorporating the ELP indicator into the broader state system of annual meaningful differentiation of school performance.

Within each of these sections, subsections offer SEA examples and technical checks.

- **State examples:** These subsections present excerpts of ELP CoP participants' consolidated state plans. Additional state examples are included to show a range of approaches to incorporating ELs into state accountability systems. It is important to note that the approaches used in these examples are not required; rather, they are only illustrative approaches to consider and do not reflect an exhaustive list of possibilities or exemplars.
- **Technical checks:** These subsections provide questions or actions that an SEA can consider for each of the topic areas.

This guidebook also includes three appendices. Appendices A and B provide relevant information on the ELP CoP. Appendix C includes select statutory and regulatory requirements related to ELs.

⁵ Each state must establish long-term goals for increases in the percentage of ELs making progress in achieving English language proficiency, as measured by the statewide ELP assessment. A state must establish a timeline by which it expects ELs to attain English language proficiency. In addition, each state must establish measurements of interim progress toward meeting the long-term goal (ESEA section 1111(c)(4)(A)(ii)).

⁶ Each state must establish an accountability system that includes at least five indicators for all schools, including an indicator of the Progress in Achieving ELP. This "ELP indicator" must be based on the state's definition of English language proficiency, within a state-determined timeline, and measured by the statewide ELP assessment required under ESEA section 1111(b)(2)(G). The ELP indicator must be measured for all ELs in Grades 3–8 and once in high school (in the grade in which reading/language arts and mathematics tests are given) (ESEA section 1111(c)(4)(B)(iv)).

⁷ Each state must establish a system of meaningful differentiation that includes all required indicators (for all students and each subgroup) and that meaningfully differentiates, on an annual basis, among all public schools in the state. The system must give "substantial weight" to each of the academic achievement, other academic, graduation rate, and progress in achieving ELP indicators and "much greater weight" to these indicators in aggregate than the school quality or student success indicator or indicators. The system must also differentiate schools with consistently underperforming subgroups (ESEA section 1111(c)(4)(C)).

Section 1: Research Related to Incorporating ELs Into State Accountability Systems

ELP Goals and Measurements of Interim Progress

States must establish long-term goals for increases in the percentage of ELs making progress in achieving English language proficiency, as measured by the statewide ELP assessment. In addition, states must establish measurements of interim progress toward meeting long-term goals (ESEA section 1111(c)(4)(A)(ii)). In their state plans, states must establish a time frame by which ELs are expected to attain English language proficiency. Throughout this report, a common terminology is used to denote a series of technical steps that states must take to ensure that their statewide systems hold schools accountable for ELs. Student-level progress targets indicate how much progress individual EL students are expected to make from one year to the next. Measurements of interim progress denote the amount of progress the state expects ELs in the state and in a given school to make toward the statewide goal.

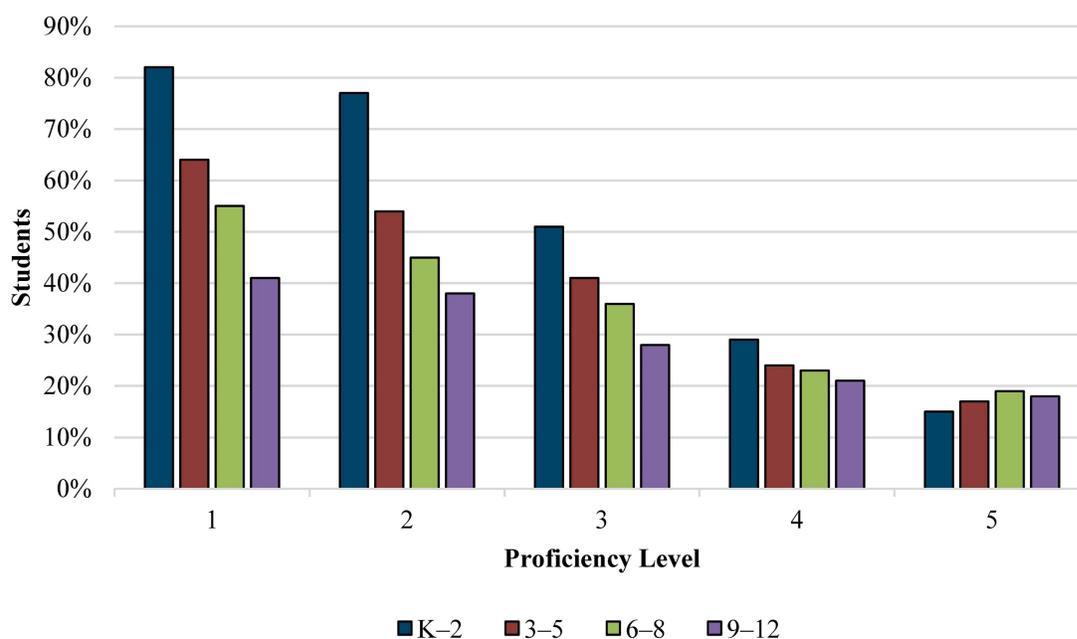
In setting realistic timelines for the length of time it takes ELs to attain English language proficiency, states can take into account individual and contextual characteristics that influence time to attainment. Research studies conducted during the past decade, including many rigorous longitudinal studies, have determined several individual and contextual characteristics that relate to EL students' progress toward and attainment of English language proficiency.

Although method, context, and student samples differ, these studies can help SEAs to consider the following key factors in determining student-level progress targets for attaining ELP because they have been shown to play a meaningful role in determining the pace at which ELs attain English language proficiency:

- Initial level of ELP as of the first ELP assessment (Cook, Boals, Wilmes, & Santos, 2008; Cook, Linquanti, Chinen, & Jung, 2012; Kieffer, 2011; Slama et al., 2017)
- Grade/age at school entry (Conger, 2009; Cook et al., 2008; Cook et al., 2012)
- Level of native language proficiency
- Special education status
- Proportion of students in a school from high-poverty families
- Type of language instruction educational program (LIEP)

The characteristics listed above are explained in greater detail here. Cook and colleagues (2008) coined the phrase *lower is faster, higher is slower*—that is, ELs at lower levels of English language proficiency or lower grade levels show more rapid growth than ELs at higher levels of English language proficiency or higher grade levels.

Figure 1. Aggregate Percentage of Students Gaining One or More Proficiency Levels (2005–07)



Source: Reproduced based on Cook et al. (2008), p. 10

As displayed in Figure 1, ELs’ initial levels of English language proficiency and grade level influence the rate of attaining English language proficiency. ELs who enter schools less proficient in English and in the lower grades make faster gains in English language proficiency than ELs who begin more proficient at school entry or enter schools in later grades. Thus, if the same progress and attainment targets are set across proficiency levels and grade bands, EL students at lower grades and proficiency levels would show progress at higher rates than their higher grade- or proficiency-level peers. As such, setting growth expectations without taking these factors into consideration could cause accountability systems to overidentify middle and high schools as not meeting progress and attainment goals (Cook et al., 2008; Hakuta & Pompa, 2017).

Another student characteristic shown to influence the pace at which students develop English language proficiency is the level of native language proficiency (García-Vázquez, Vázquez, López, & Ward, 1997). Students with higher levels of native language proficiency attain English language proficiency before their respective EL peer comparison groups.

Recently, several empirical studies have focused on special education status as a factor related to time needed to gain English language proficiency. Employing discrete time survival analysis (a longitudinal data analysis methodology), the studies demonstrate that some students who are dual-identified as EL students with disabilities take significantly longer to attain proficiency on statewide ELP assessments (Haas, Huang, Tran, & Yu, 2016a, 2016b; Haas, Tran, Huang, & Yu, 2015; Kieffer & Parker, 2016; Slama, Haynes, Sacks, Lee, & August, 2015). In fact, many of these students never attain English language proficiency during their time in K–12 public schools (see Liu, Thurlow, Press, & Lickteig, 2018).

According to data collected under the Individuals with Disabilities Education Act (IDEA) during the 2013–14 school year, the majority of ELs with disabilities have specific learning disabilities (51 percent) and speech or language impairment (22 percent). Recent research (Kieffer & Parker, 2016) found that ELs with specific learning disabilities took four years longer to be reclassified than their EL peers without such disabilities, and ELs with speech or language impairments took two years longer to be reclassified than EL students without these disabilities.

The proportion of students in a school who are living in poverty has been shown to play a role in students' attainment of English language proficiency and academic achievement (Hakuta, Butler, & Witt, 2000; Kieffer, 2011; Kim, Curby, & Winsler, 2014; Lindholm-Leary & Borsato, 2006; Slama et al., 2017).

A final, program-related contextual factor—the language instruction educational program (LIEP) model—has recently been shown to influence the time frame and rate in which EL students progress in attaining ELP as well as demonstrate academic progress and proficiency using English (Steele et al., 2017; Umansky & Reardon, 2014; Valentino & Reardon, 2015). These three studies examine longitudinal administrative data on ELs in large urban districts and report trends in the time it takes ELs to reach key educational milestones, such as exit from EL status (i.e., reclassification), or meet state performance standards in reading/language arts (R/LA) or mathematics. All three studies drew similar conclusions: EL students in dual-immersion instructional program models (which aim to simultaneously develop and sustain language proficiency and academic achievement in two languages) tend to exhibit slower progress toward ELP and academic progress in early grades but accelerate their growth during middle school grades and ultimately attain both English language proficiency and academic grade-level achievement using English at greater rates than EL students receiving English immersion approaches.

More specifically, Umansky and Reardon (2014) employed discrete time survival analysis to examine time to English language proficiency attainment, reclassification, and content mastery for multiple cohorts of ELs enrolled in four different EL programs: English immersion, transitional bilingual, maintenance bilingual, and dual immersion.⁸ The authors found that Hispanic ELs enrolled in dual-immersion programs had lower reclassification rates in elementary school than students enrolled in other program types but higher overall reclassification rates and higher English and content-area proficiency by the end of high school.

Valentino and Reardon (2015) compared EL students' growth in mathematics and R/LA for students who were enrolled in different EL instructional programs (English immersion, transitional bilingual, developmental bilingual, and dual immersion) but whose parents had similar program

⁸ In this study, English immersion programs are defined as classrooms in which ELs are instructed solely in English with the goal of promoting English language proficiency and providing academic content that is accessible to ELs; transitional and maintenance bilingual programs are classrooms designed specifically for ELs and provide instruction in English and in a target home language (transitional bilingual programs focus on using the home language to support English acquisition and access to curricular content, while maintenance bilingual programs prioritize full bilingualism in English and the home language); and dual-immersion classrooms, like bilingual classrooms, offer instruction in English and a target language but student composition includes both ELs and English-only speakers (EOs) with the goal that both groups develop proficiency in both languages. Within each of these four broad models, there is extensive heterogeneity, including program goals, class composition, instructional techniques, program structure, fidelity to program design, and program quality (August & Shanahan, 2006).

preferences prior to enrolling their child in language instruction educational LIEPs.⁹ The study found EL students in the dual-immersion program showed the strongest positive long-term growth in R/LA achievement. Although EL students in the dual-immersion program showed negative short-term growth in R/LA achievement in Grades 2–4, by Grade 7 dual-immersion EL students outperformed the EL students in the other three programs in R/LA achievement.

Steele and colleagues (2017) capitalized on a lottery system to estimate the impact of randomized enrollment in a dual-language program on EL student achievement in a large, urban school district. They found that EL students randomly enrolled in dual-language immersion programs in kindergarten outperformed their non-dual-language-enrolled EL peers by about seven to nine months' worth of reading gains in fifth and eighth grades, respectively.

Although no studies were located that have systematically examined the role of interruptions in a student's formal education in time to reclassification, this is an attribute that might be considered in future studies. This EL subpopulation is usually referred to as students with interrupted formal education (SIFE) or students with limited or interrupted formal education (SLIFE).¹⁰ The definition of SIFE and SLIFE students is locally defined and may vary by state or district. One example of a district-level definition for a SLIFE student is a student who has attended school in the United States for less than 12 months and performs two or more years below grade level in literacy and/or mathematics in their home language because of inconsistent or interrupted schooling in their home country (New York City Department of Education, 2016). In this example, SLIFE students enter U.S. schools beyond the primary grades and with lower levels of initial English language proficiency and native language proficiency. Given these entry-level characteristics, it is reasonable to expect that these students will need additional instructional supports, and possibly time, to attain key educational milestones. States such as New York and Texas have explicitly factored these characteristics into their progress expectations.

Explicitly factoring initial ELP level and other factors that predict time needed to gain proficiency into accountability plan design with respect to time-to-proficiency and growth expectations might be considered because it begins to address the issue of systematically advantaging schools that have students with one set of initial characteristics over schools with different initial student characteristics. SEAs might also take care not to over-adjust expectations in light of school aggregate characteristics or other school variables. School-level aggregate adjustments assume that (1) ELs are equally represented by the school variable (e.g., percent poverty), (2) EL students' ELP progress is impacted by the school variable, and (3) each EL is equally impacted by the school variable. Further, adjusting results by school variables makes it difficult to interpret performance and can obfuscate results because the lower/higher performance is effectively factored out. SEAs should be mindful of how their accountability design choices interact with school composition (Goldschmidt, 2020).

⁹ The authors (Valentino & Reardon, 2015) capitalize on a complex EL program assignment process in the study district that considers parental preference. Data on parental preference allows the authors to estimate the effects of different EL instructional programs, comparing the academic outcomes of ELs whose parents preferred the same school and program but who attended different programs because of oversubscription.

¹⁰ The term "SLIFE" is used to refer to students with limited formal education, those with interrupted education, and those with both limited and interrupted formal education.

The ELP Indicator

The ELP indicator¹¹ must be calculated for all ELs in at least Grades 3–8 and the grade in high school in which the state administers the statewide reading/language arts and mathematics assessments under ESEA section 1111(b)(2)(B)(v)(I). States can include the ELP assessment results of ELs in earlier grades, which may be appropriate given that younger students tend to show the most growth in English language proficiency (Lyons & Dadey, 2017). As long as the indicator meets the requirements established in the statute,¹² each state makes its own decisions, in conjunction with stakeholders, regarding how the ELP indicator is operationalized.

It is important to consider validity evidence as it relates to an ELP indicator. In this case, validity evidence relates to whether assessment results appropriately and adequately support claims about ELs' English language skills and whether changes in performance allow for reasonable claims that an EL in fact has made progress in learning English. The progress determination (as a transformation of multiple ELP assessment results) should provide evidence to support the interpretation that an individual student is making *sufficient* progress; moreover, evidence should support the claim that, in aggregate, the progress determinations from the ELP indicator are facilitated *by schools* and not a concomitant factor (Goldschmidt, 2020).

There are a variety of models to consider in establishing an ELP indicator (see Goldschmidt and Hakuta [2017] for an extended review). There is no “best model,” particularly when considering state context, accountability systems, and state theory of action (Goldschmidt & Hakuta, 2017; Hakuta & Pompa, 2017). Trade-offs to consider in model selection include transparency, link to a state's theory of action, basis in research and historical performance, and degree of connection with long-term goals and measurements of interim progress (see Lyons & Dadey, 2017; McHugh, Pompa, & Lyons, 2017).

For technical reasons discussed below, a key challenge of the ELP indicator as a differentiator of school performance with respect to ELs' English language proficiency progress and attainment is to ensure that individual student-level targets of ELP progress (see sections 2 and 3) are properly aggregated into a school-level progress indicator. In this document, school-level targets are the percentage of EL students meeting their individual progress targets in a given year.

¹¹ Under the ESEA, the ELP indicator measures increases in the percentage of ELs making progress toward English language proficiency as defined by the state, measured by the state ELP assessment, and within a state-determined timeline.

¹² This ELP indicator must be based on the state's definition of English language proficiency and measured by the statewide ELP assessment required in ESEA section 1111(b)(2)(G), within a state-determined timeline. The ELP indicator must be measured for all ELs in Grades 3–8 and once in high school (in the grade in which reading/language arts and mathematics tests are given) (ESEA section 1111(c)(4)(B)(iv)).

Translating Student-Level Progress Into a School Performance Indicator

Goldschmidt and Hakuta (2017) provide a framework for technical considerations in designing the ELP indicator so that it meaningfully differentiates schools by the actual contributions that schools make to EL student progress (as opposed to characteristics of students or some external factor). The ELP indicator should be research- and empirically-based, realistic, and aspirational. Importantly (and as required by statute), it should differentiate among schools (i.e., the indicator meaningfully distinguishes schools by their outcomes on the ELP indicator).

Distinguishing schools by their contribution to progress toward English language proficiency results from aggregating individual student progress results in some way to a school. The results of this aggregation *and the claims about schools* are affected by the individual targets as well as the reliability of the aggregate.

If student-level targets are not set appropriately, then, in the aggregate, results will not meaningfully differentiate among schools. If targets are set too low, most students will meet targets even if ELs are not making sufficient progress toward meeting English language proficiency within the state’s timeframe. If student targets are set too high, then most students would fail to meet the target even if students are generally making sufficient progress toward proficiency. In the aggregate, under either scenario, the indicator would not meaningfully differentiate among schools, and this factor would erode both the credibility and validity of the accountability system.¹³

For example, if the progress target for Year 2 is set too low,¹⁴ then students will easily meet the target in Year 2. Setting the target too low may result in four unintended consequences: (1) the ELP indicator in elementary school could be inflated because the targets are too easily met in the early years to which the ELP indicator applies; (2) there may be little differentiation in the ELP indicator results by elementary school because ELP results are inflated toward the maximum score (most schools get good results on the ELP indicator); (3) middle schools might receive students who appeared to have made “good” progress but are still ELs and, if the ELP targets are linear, might be likely to miss targets in higher grades because latter targets are too high—generally making it appear as though middle schools are not providing sufficient support for ELs; and (4) there could be a lack of differentiation in the ELP indicator among middle schools, if the ELP targets are linear, because ELP indicator results might be biased toward the low end of its range.

Precision and reliability of the ELP indicator also impact claims about schools. SEAs with significant similarity in student progress across schools could consider the precision of the ELP indicator (i.e., whether one could make claims about progress uniquely attributable to schools). States with low reliability across ELP data, however, could consider the ELP indicator’s validity (i.e., whether one can make a claim about the amount of progress). This concept is detailed in Goldschmidt (2020).

¹³ See the Glossary section of this guidebook for the definition of *validity*.

¹⁴ A target that is too low might be based on a linear progress trajectory.

Weighting the ELP Indicator in the Accountability System

An important step in developing a comprehensive accountability system that includes an ELP indicator is to determine how the ELP indicator contributes to inferences about schools.

The statutory weighting requirement in ESEA section 1111(c)(4)(C)(ii) states that each of the Academic Achievement, other academic indicator for elementary and secondary schools that are not high schools (the Other Academic), Graduation Rate, and ELP indicators must receive “substantial” weight individually and “much greater weight” in the aggregate than the School Quality or Student Success indicator or indicators.

With respect to weighting the ELP indicator, states might choose a fixed weight, a dynamic weight, a dashboard approach, or other approach to meet the weighting requirement. Each approach has trade-offs that should be considered. A weight is fixed when an SEA selects a single weight for a given indicator across schools that meet the minimum n size for accountability. If a state chooses a fixed weight, it should consider what percentage of the state’s ELs are included in the accountability system. For a school that does not meet the minimum n size, the state should consider how the fixed weight might be redistributed to other school-level indicators. For example, if each of the accountability indicators is worth 25 points (total amount of points a school could earn would be 100, since each school would receive points for the academic achievement, ELP, and school quality or student success indicator and either the other academic indicator for non-high schools or the graduation rate indicator for high schools), and if the minimum n for a state is 15, then for schools that have fewer than 15 EL progress scores, the ELP indicator would be excluded from the total summative accountability score. If the state intends to have all schools’ accountability scores sum to 100 for ease of comparison, then the points attributable to the ELP indicator (25) need to be redistributed to the remaining indicators (e.g., 10 points to the academic achievement indicator and 15 points to the other academic indicator).

A dynamic weight is one that changes based on the proportion of EL students in the school. A state might consider a dynamic weight if its schools have a wide range in the percentage of ELs. One approach to implementing and reporting the results based on dynamic weighting is to establish a business rule in which the proportion of ELs determines the weight of the ELP indicator; a school with a larger percent of ELs includes the ELP indicator as a larger percentage of the accountability system than a school with a smaller percent of ELs.

A dashboard approach—which displays multiple performance indicators in a graphic format—could use either a conjunctive or an implicit fixed weighting scheme. A conjunctive approach requires a school to meet several criteria in order for consequences to accrue. For example, a conjunctive rule might be that an elementary school that scores in the bottom 10 percent in the Academic Achievement indicator, Other Academic indicator, and ELP indicator, taking into account performance on the School Quality or Student Success indicator, is designated for state-required interventions. Implicit fixed weights exist when a state uses a matrix (e.g., value table) to relate its indicators to one another by establishing value judgements to categorize school performance on each indicator of school performance.

The next four sections, respectively, illustrate approaches to (1) setting ELP goals and measurements of interim progress, (2) constructing and evaluating the annual ELP indicator, (3) translating

student-level progress into a school-level indicator of ELP progress, and (4) weighting the ELP indicator within the accountability system for purposes of annual meaningful differentiation. Mock state examples as well as actual SEA approaches appear in the next sections.

Section 2: ELP Goals and Measurements of Interim Progress

This section contains actual state examples and illustrative mock examples that demonstrate different ways to develop ELP goals and measurements of interim progress (MIPs). It should be noted that although MIPs need not be annual, states need to measure student progress annually for purposes of calculating the ELP indicator.

Examples in this section first demonstrate how to define student-level progress targets (i.e., how much progress individual EL students are expected to make from one year to the next) and then demonstrate how performance relative to those targets informs long-term goals and MIPs. The examples use simulated data to illustrate general principles relevant to ELP goals and MIPs.

Importantly, many states are transitioning to new ELP assessments, which affects the ability to empirically analyze and identify proficiency cuts, time frames to English language proficiency, baseline progress, progress toward long-term goals, and MIPs. Some states in transition have considered alternative approaches to setting goals and MIPs that are discussed in this section. These alternative approaches use prior years' data or simulated data to set goals and MIPs for now with the intention of amending them as needed once performance data on ELP progress and attainment are available using the new ELP assessment.

ESEA Requirements

To satisfy the requirements of ESEA section 1111(c)(4)(A)(ii), peer reviewers reviewed each state's consolidated state plan to ensure it addressed the following questions, as outlined in the U.S. Department of Education State Plan Peer Review Criteria:¹⁵

A.4.iii.c.1: Long-term goals¹⁶

- Does the SEA identify and describe the long-term goal for increases in the percentage of English learners making progress in achieving English language proficiency, as measured by the statewide English language proficiency assessment?
- Does the SEA's description include baseline data?
- Does the SEA's description include the state-determined timeline for English learners to achieve English language proficiency?
- Is the long-term goal ambitious?

¹⁵ The U.S. Department of Education State Plan Peer Review Criteria, available at <https://www2.ed.gov/admins/lead/account/stateplan17/essastateplanpeerreviewcriteria.pdf>, was issued in March 2017 and provides the criteria by which state plans were reviewed based on statutory and regulatory requirements.

¹⁶ State Plan Peer Review Criteria, page 10.

A.4.iii.c.2: Measurements of interim progress¹⁷

- Does the SEA provide measurements of interim progress toward the long-term goal for increases in the percentage of English learners making progress in achieving English language proficiency?

Challenges for States

Some challenges that states faced in establishing student-level progress targets and state ELP long-term goals and MIPs include the following:

- Establishing ambitious yet reasonable timelines for ELs in a state to achieve ELP
- Determining what EL characteristics, if any, a state should take into consideration when setting student progress expectations and the long-term goal
- Establishing ELP goals using a new ELP assessment and not using historical ELP assessment data
- Determining models that might be used to establish the pace of growth required for an EL to attain ELP within the state’s timeline

Technical Approaches

The following examples illustrate how some states took into account different student factors in setting ELP goals and how they set goals using baseline data. Some examples also include discussion of issues related to changes in the statewide ELP assessment.

Setting Student-Level Progress Targets to Signal an Expected Time Frame to Attain English Language Proficiency (Mock State Example)

Individual EL student-level progress targets should be set in a manner that clearly communicates the state’s expected time frame for an EL student to attain the English-proficient performance standard on the state ELP assessment. This entails two steps: First, establishing the “finish line” that defines English language proficiency on the state ELP assessment; and second, specifying annual progress expectations for EL students at different levels of English language proficiency so that they attain the English-proficient performance standard within an ambitious and reasonable time. Several resources developed over the past decade offer approaches and examples to illustrate how states can undertake these steps using empirical evidence (e.g., Cook et al., 2008; Cook et al., 2012; Linqunti & George, 2007). As described in Section 1, in setting student-level progress expectations, it is important to take into consideration ELs’ initial level of English language proficiency and time in language instruction educational programs, as well as other factors that a state deems important.

Although covered in great detail in Cook et al. (2012) and Goldschmidt (2020), it is important to note that simply using the average time it has taken ELs to reach a specified score on the ELP assessment will likely underestimate the time needed to obtain the English-proficient score. Although survival analysis is the preferred technique, other methods can also provide an adequate

¹⁷ State Plan Peer Review Criteria, page 10.

indication of when ELs could be expected to attain the English-proficient standard as well as the shape of growth over time. These include cumulatively summing the annual gains of ELs by initial proficiency level, as well as plotting the cumulative percentage of students reaching the proficiency cut by initial ELP level and time (see Goldschmidt [2020], Chapter 1, for example).

Once the SEA has set both the English-proficiency cut score on the ELP assessment and the expected time to reach the cut score for each initial ELP level,¹⁸ then the SEA can develop student-level targets. For example, if the SEA determines that it will take an initial ELP level 1 student six years to reach proficiency, then it can devise a method to develop a progress trajectory for each student that results in student progress targets. This can be based, for example, on composite performance levels, composite scale scores, or domain scores. SEAs should not expect linear growth over time and should explicitly align the student-level progress targets to match existing progress (generally curvilinear—showing steep early progress followed by a slowing of progress over time). Targets can then be illustrated using a value table that specifies in the cells how much growth is expected over a certain amount of time by starting level. Targets can also be set by dividing the composite scale score growth needed to reach the proficiency cut by the number of years to reach proficiency (i.e., “growth to target”). Targets can also be based on student growth models (e.g., value-added models [VAM]) or student growth percentile [SGP] models) in which targets are based on a standardized result that aligns current performance with reaching the proficiency cut in the specified time frame (e.g., an EL whose initial ELP level is 1 should be in the 60th percentile of composite scale score growth to reach the English-proficient cut in the specified time frame).

Once a state has set its student-level targets, it should examine what percentage of students reach those targets. This percentage is the baseline for setting the long-term ELP goal and should inform the MIPs. It should be noted that MIPs are not required to be annual—a state could still determine annual student-level targets but just apply them at whatever interval the state is measuring MIPs. Note that a state must make annual accountability determinations even though the MIPs need not be annual.

Identifying Annual Progress and Time to Proficiency (Mock State Example)

States can create tables such as Tables 1 and 2 and compare tables for available data to see how they differ. Based on second language development research, we know that EL students at lower ELP levels grow faster initially and then more slowly the longer they are in EL status. Research also provides an indication of the shape of growth, cumulative growth, and a general time frame within which EL students at the lowest ELP level on entry typically attain English language proficiency—four to ten years or more (Conger, 2009; Grissom, 2004; Hakuta et al., 2000; Umansky & Reardon, 2014).

Tables 1 and 2 are based on a longitudinal data set generated using two or three years of ELP assessment data. Any given statewide data file contains students of varying years in the state educational system. This can be used to create a longitudinal view of the ELP gains that can be expected over time. Table 1 shows the average ELP gains across EL students with varying numbers of years in the school system. For example, an EL student two years in the system and who started at ELP level 1 gains an average of 1.2 ELP performance levels, compared with

¹⁸ And any other student characteristic the SEA deems necessary (e.g., grade span).

0.9 levels for an EL student who has been in the system for three years and started at ELP level 1. These data generally confirm research-derived expectations about progress: Students who start at lower levels show more growth early on and slower growth over time, while students who start at higher initial ELP levels exhibit less growth over time, and growth decreases over time (Hakuta & Pompa, 2017).

Table 1. Mock Example: Average ELP Performance-Level Gains by Initial ELP Level and Years in Program

Year in Program	Level 1	Level 2	Level 3	Level 4
2	1.2	1.2	0.6	0.1
3	0.9	0.7	0.6	0.5
4	0.3	0.2	0.2	0.3
5	0.5	0.5	0.6	0.6
6	0.4	0.4	0.4	0.3
7	0.3	0.3	0.3	0.3
8	0.1	0.2	0.1	0.2

Note: Numbers in the cells refer to average ELP growth in levels.

Table 2 expands on Table 1 but disaggregates the data by whether a student has exited EL status within the current year. If we examine results only for the students that have exited, we will *underestimate* the time to proficiency because the calculation is based only on the most successful students. Similarly, if we focus only on those not attaining ELP, we will likely *overestimate* time to proficiency expectations. Statistical methods that take into account patterns of growth for both groups of students (those attaining and not yet attaining English language proficiency) provide a more accurate picture of growth expectations in a state.

For example, Table 2 shows that an EL student with two years in an EL program and who entered at ELP level 1 and exited grew 4.0 ELP levels in one year, compared with 1.1 ELP levels for a student who did not exit. The average EL student’s growth falls somewhere in between. It is also important to consider *how many* students exhibit each trajectory. States could take the average and then increase the estimate to establish reasonably ambitious expectations for progress.

A state using a normative measure to set student-level targets should ensure that its timeline allows students to attain English language proficiency within that timeline. Such a state might not want to use the median as a benchmark (e.g., the 65th percentile).

Note that in the inaugural year of a new ELP assessment, states should expect lower performance as examinees and administrators become accustomed to the assessment, and thereafter expect a relative jump in performance in the subsequent year. If the state focuses on the second year gains, long-term expectations could be overestimated. Moreover, results may be confounded with scale changes, differing equating methods, and changes in student population from one year to the next. States can mitigate the impact of updating or changing ELP assessments and should consult their TACs (Technical Advisory Committees) before the new ELP assessment is administered. Unfortunately, there is no way to mitigate against such confounding factors if the assessments are already operational. States with operational assessments can conduct analyses (e.g., examining the effects of potential differences in cohorts taking the assessment, differences in proficiency

cuts, differences in growth expectations and differences in expected time to proficiency) to help inform determinations of the percentage of students making sufficient progress.

Table 2. Mock Example: Average ELP Performance Level Gains by Initial ELP Level, Years in Program, Exit Status

Year in Program	Level 1		Level 2		Level 3		Level 4	
	Not Exit	Exit						
2	1.1	4.0	1.0	3.0	0.4	2.0	-0.3	1.0
3	0.8	1.6	0.5	1.3	0.4	1.2	0.2	1.1
4	0.1	1.2	-0.1	1.1	-0.2	1.1	-0.2	1.0
5	0.3	1.2	0.3	1.1	0.3	1.1	0.2	1.1
6	0.3	1.1	0.1	1.0	0.1	1.0	0.0	1.1
7	0.1	1.1	0.0	1.0	0.0	1.0	-0.1	1.0

Note: Numbers in the cells refer to average ELP growth in levels. Note that exit criterion is assumed to be achieving level 5 on the ELP assessment. The table assumes some degree of concordance between old and new scores (e.g., through equipercentile equating). The table creates artificial longitudinal data by taking advantage of current-year students being at different points along the time continuum. English-proficient status is based on current year only. *Source:* Goldschmidt (2020)

Setting Student-Level ELP Progress Expectations: Differentiating Student-Level Targets for ELP Progress by Initial ELP Level and Grade at Entry (Delaware Department of Education)¹⁹

The following example from Delaware illustrates one way to differentiate student-level ELP progress targets by initial ELP level and grade at entry.

As a member of the WIDA Consortium (formerly World-Class Instructional Design and Assessment),²⁰ the Delaware Department of Education administers the ACCESS for ELLs 2.0 (ACCESS) as its statewide ELP assessment. In collaboration with WIDA consultants, the Delaware Department of Education analyzed EL performance on ACCESS in relation to performance on the state’s R/LA content assessment²¹ and set EL students’ English-proficient performance standard (the exit or attainment target [AT]) as a 5.0 composite proficiency level on the ACCESS.²²

Feedback from various stakeholder groups, including the Governor’s Advisory Committee, the English as a second language (ESL) Coordinator Group, and the Spanish-Language Community Engagement sessions, supported calculating growth in ELP based on the student’s initial ELP level and differentiated by grade level or grade band. As a result of this feedback, the Delaware Department of Education established EL accountability measures that better account for individual differences

¹⁹ Drawn from the Delaware Department of Education Consolidated State Plan (approved June 3, 2019) under the ESEA.

²⁰ The WIDA Consortium is a group of 40 U.S. states, territories and federal agencies dedicated to the research, design and implementation of a high-quality, culturally and linguistically appropriate system to support English language learners in K-12 contexts.

²¹ See Cook et al. (2012), Chapter II, for more information.

²² Delaware has submitted an amendment as part of their consolidated state plan to change the attainment target in their statewide ELP assessment from a 5.0 0 composite proficiency level to 4.7. This amendment is currently under review by the Department.

among ELs with a focus on initial English language proficiency level and entering grade-level scale score.

Starting with the 2016–17 assessment cycle, the Delaware Department of Education defines increases in the percentage of all current ELs making expected progress in ELP as ELs that meet the ELP cut scale score (SS) designating the English-proficient performance standard threshold within the established time frame, consistent with the student’s baseline ACCESS performance level; see Table 3 for student-level interim growth targets. Thus, the state considers a student’s performance level on the first annual ACCESS assessment to determine the number of years a student has to attain English language proficiency and then sets annual interim progress targets based on entering grade-level SS accordingly. Under this model, students achieving a performance level of 5.0 or higher on their first ACCESS annual assessment (Year 1) have met their growth target. EL students are expected to meet targets within six years.²³ This decision was a result of significant stakeholder input to the Delaware Department of Education, including ESL coordinators, the Governor’s Advisory Committee, and empirical research in second-language acquisition.

Under this model, each EL student’s ELP attainment target is the scale score at the 5.0 performance level on ACCESS in the grade level and year that he or she is expected to attain English language proficiency. The number of years in which the state expects an EL to reach the attainment target varies from three to six years depending on the student’s initial ELP level (at Year 1 baseline). As described below, each student’s interim growth targets are calculated annually by subtracting his or her previous year’s scale score from the attainment scale score and dividing the difference by the remaining number of years required to reach attainment. This method allows for a variable trajectory depending on each student’s progress over the years while still requiring that the attainment target be reached in the specified number of years.

²³ Although Delaware has established a timeline in which it expects ELs to meet ELP targets, ELs cannot be exited from EL status until they have demonstrated proficiency consistent with the state’s established criteria.

Table 3. Delaware EL ACCESS Growth Targets—Annual Calculation Method

Year 1 Baseline ACCESS Performance Level	Growth Target				
	Year 2	Year 3	Year 4	Year 5	Year 6
5.0 or Higher					
4.0–4.9	Year 1 SS plus SS progress to reach to AT divided by 2	SS for 5.0 two grades out (AT)			
3.0–3.9	Year 1 SS plus SS progress to reach to AT divided by 3	Year 2 SS plus SS progress to reach to AT divided by 2	SS for 5.0 three grades out (AT)		
2.0–2.9	Year 1 SS plus SS progress to reach to AT divided by 4	Year 2 SS plus SS progress to reach to AT divided by 3	Year 3 SS plus SS progress to reach to AT divided by 2	SS for 5.0 four grades out (AT)	
1.0–1.9	Year 1 SS plus SS progress to reach to AT divided by 5	Year 2 SS plus SS progress to reach to AT divided by 4	Year 3 SS plus SS progress to reach to AT divided by 3	Year 4 SS plus SS progress to reach to AT divided by 2	SS for 5.0 five grades out (AT)

Notes: SS refers to the ELP cut scale score for proficiency. Attainment targets (AT) are highlighted in orange. Students receiving a performance level of 5.0 or higher on their initial ACCESS assessment (Year 1) are considered to have met their growth target. Students scoring below 5.0 on their Year 1 ACCESS assessment are expected to reach attainment in two to five additional years depending on their initial proficiency level. The state considers a student’s proficiency level on the first annual ACCESS assessment when determining the specific number of years that the student has to reach proficiency and then sets targets for interim progress accordingly.

Source: Delaware Department of Education. (n.d.). *Delaware Consolidated State Plan* (Approved June 3, 2019). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/deconsolidatedstateplanfinal.pdf>

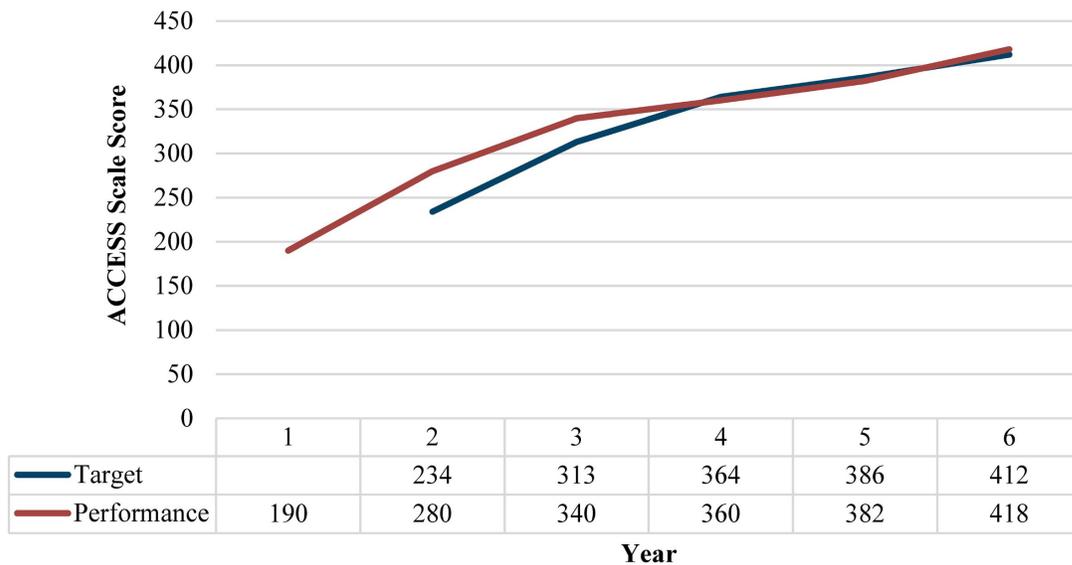
This model is summarized in the following equation:

$$\text{Interim growth targets} = \frac{\text{Scale score}_{\text{attainment}} - \text{Scale score}_{\text{prior year}}}{\text{Number of years to attainment}}$$

Figure 2 provides an example of how the annual target calculation method allows for an ELP development trajectory that matches each student’s progress pattern. The sample EL student featured starts Year 1 in third grade with a scale score of 190. Because the Year 1 proficiency level is in the 1.0–1.9 category, the student has a total of six years to meet her attainment target of 412, the corresponding scale score of a proficiency level 5.0 in eighth grade.

The student demonstrates rapid scale score growth in Years 2 and 3 but slower growth in Years 4, 5, and 6. The student exceeds the attainment target in Year 6. The sample student’s actual performance trajectory (red line) is similar to their expected growth trajectory over the time span.

Figure 2. Comparison of ELP Growth Targets and Actual Performance (Rapid Early Growth, Slower Later Growth)



Source: Delaware Department of Education. (n.d.). Delaware Consolidated State Plan (Approved June 3, 2019). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/deconsolidatedstateplanfinal.pdf>

Setting Long-Term Goals and MIPs With One or Two Years of ELP Assessment Data (Delaware Department of Education)²⁴

Although states were required to assess all ELs’ English language proficiency annually under the ESEA, as amended by the No Child Left Behind Act, many states began the first year under the reauthorized ESEA with a new ELP assessment. Reasons included implementing a new version of a consortium’s assessment, as was the case for WIDA states with ACCESS 2.0; transitioning from a state stand-alone assessment to a consortium ELP assessment; or changing a stand-alone state’s ELP assessment to align with new ELP standards that reflect the language demands of new college- and career-ready academic content standards.

Transitioning to these more rigorous ELP assessments requires EL students to demonstrate higher levels of English language proficiency to attain the equivalent proficiency level scores used to exit them from EL status in previous years. For example, one ELP CoP member state compared its results from the 2016 and 2017 versions of ACCESS and found a 21 percentage point decrease in the proportion of EL students attaining the state’s composite score performance standard of 5.0 or greater needed to exit EL status.

There are methods states can use to compare results across substantially different versions of an ELP assessment. A state may decide to amend its progress expectations based on two years of assessment data (or one year of data for a rescaled assessment). There are several strategies for approximating the gains to expect in the first year of the new assessment. A state can look at successive pairs of years to assess what progress looks like and how year-to-year gains change

²⁴ Drawn from the Delaware Department of Education Consolidated State Plan (approved June 3, 2019) under the ESEA.

comparing old-to-old, old-to-new, and new-to-new ELP assessments (e.g., by comparing matched performance using the assessment two years prior to one year prior, one year prior to current year, and current year to next year’s assessment).

For example, when first establishing long-term attainment and interim progress goals, WIDA states might use the 2015–16 ACCESS 1.0 test results placed on the ACCESS 2.0 scale and compare matched-score results with those of the 2016–17 ACCESS 2.0. When using SGPs, growth-to-target, or value tables, states can examine the relationship of ACCESS to academic content test results in order to inform performance standard-setting for establishing an English-proficient performance standard (if that is desired). The WIDA consortium is continuing to conduct analyses across consortium member states, including historical event (survival) analyses to examine how long it takes EL students of different characteristics to attain the 5.0 performance level on the ELP assessment. The Delaware Department of Education²⁵ determined that the new performance standards set for ACCESS 2.0 significantly affected the trajectory and the amount of time required for ELs in their state to attain English language proficiency. The state conducted an analysis of 2017 ACCESS 2.0 data to establish its long-term goals. Once trend data are available, the state will calculate projections for subsequent years.

The Delaware Department of Education analyzed the differences between the 2015 and 2017 score distributions by performance level to assess differences in performance on the respective 2016 and 2017 ACCESS for ELLs 2.0 scores compared with ACCESS 1.0 scores (2011 to 2015). Analyses revealed increases in the percentage of EL students that scored at performance levels 1, 2, and 3 from 2015 to 2017. The percentage of students scoring at level 4 remained relatively unchanged. The most notable change was a dramatic decrease in the number of students scoring in levels 5/6 (from 26.5 percent in 2015 to 5.4 percent in 2017). In summary, significantly more students scored at lower ELP performance levels in 2017 relative to previous years.

Setting Long-Term Goals and MIPs Assuming Linear Student Progress (Oregon Department of Education)²⁶

Oregon’s long-term goal for ELs making progress toward achieving English language proficiency is 90 percent. The timeline for ELs in Oregon to achieve this long-term progress goal is eight years (i.e., from 2017–18 to 2024–25), which represents an increase of 45 percent to 90 percent performance from the baseline to the 2024–25 school year. The MIPs uniformly increase each school year by 5.625 points.

Year-to-Year Growth and Correlations Across ELP Assessment Administrations (Mock State Example)

ELP assessments tend to be less correlated over time than state content assessments. Correlations tell us about the ranking of students from one year to the next. For an ELP assessment, year-to-year correlations tend to be about .6 to .75. If the year-to-year correlations change dramatically (e.g., from $r=0.7$ to $r=0.2$), those data would raise concerns and instill less confidence in using

²⁵ Drawn from the Delaware Department of Education Consolidated State Plan (approved June 3, 2019) under the ESEA.

²⁶ Drawn from the Oregon Department of Education Consolidated State Plan (approved August 30, 2017) under the ESEA.

the results from one year to the next to establish gains. The first step is to get concordance from the prior year’s assessment and the current assessment.

Technical Check

Questions that states should ask when they examine the year-to-year changes include:

- How difficult is it for ELs to reach the new English-proficient cut score?
- How does English reading/language arts content performance distribution of ELs at various ELP performance levels compare to that of their non-EL classmates?
- What does progress look like on the current assessment from one year to the next?
- What did progress look like on our old assessment from one year to the next?
- Does progress look similar using current and old assessments?
- Are students making less progress from year to year because of implementation of a new, more rigorous assessment?
- Given the same or revised English-proficient cut score, do current time frames seem reasonable?
- Do we observe a lowering of overall scores but annual gains stay the same?
- Is the ELP progress trajectory still curvilinear with generally faster progress at lower ELP levels and grades, and slower progress at higher ELP levels and grades?
- For states using VAM or SGP models: What are the year-to-year correlations over time of our state ELP assessment?

Constructing a Longitudinal Data Set Using Two Years of ELP Assessment Data (Mock State Example)

Results from the first two operational ELPA21 administrations can be used to evaluate the adequacy of each EL student’s progress. Such analyses could help a state characterize the gains in ELP performance levels a student would need to make to be on track to English language proficiency²⁷ within a state-determined number of years and to establish student-level ELP progress expectations. One possible approach is described in the following steps:

Step 1: Determine a level of performance on the overall scale score that is associated with a high probability of attaining English language proficiency the following year for a given proportion of EL students (as determined by the state; e.g., 80 percent). This can be considered the threshold for the current year (in “Grade G”).

Step 2: Establish a “one-year-away threshold”—the point at which these EL students will be expected to attain English language proficiency in the following year (Grade G+1). For example, among EL students with a Year 1 overall scale score greater than X (but who are not yet English proficient in Year 1), a given proportion (e.g., 80 percent, 70 percent, 60 percent) attain English language proficiency in Year 2.

²⁷ ELPA21 consortium assessment decision rules indicate that an EL student less than proficient in any of the four language domains should remain EL and continue to receive English language development and other specialized support services appropriate to their level of English language proficiency.

Step 3: Work backward to establish the point that the EL student should have reached in the previous year to attain English language proficiency in Grade G (Grade G-1). The same procedure can be used to find a G-2 threshold.

Step 4: Combine Steps 1–3 to identify growth targets corresponding to an EL student being on track (i.e., with some fixed probability) to attaining English language proficiency in one year, in two years, in three years, etc. Under this approach, targets can be described in the expected number of years to attain English language proficiency. An EL student is deemed to have met the annual benchmark for progress if they have moved one year closer to attaining proficiency since the prior year’s ELP test.

Figure 3. Sample Worksheet for Documenting Progress to Target Approach

Approach	Worksheet
Step 1a. Specify performance level associated with attaining English language proficiency on the overall scale score the following year (“Grade G”).	_____
Step 1b. Specify a proportion of EL students who will reach this level (e.g., 80 percent).	_____ percent
Step 2. Set one-year-away threshold (Grade G+1) to meet proportion specified in Step 1b.	Among students with a Year overall scale score greater than _____ (but who are not deemed English proficient in Year), a given proportion (___ percent) are deemed English proficient in Year X+1.
Step 3: Establish the point the student should have reached in the previous year in order to demonstrate ELP in previous year (Grade G-1).	Among students with a Year overall scale score greater than _____ (but who are not deemed English proficient in Year), a given proportion (___ percent) are deemed proficient in Year X-1.
Step 4: Identify growth targets corresponding to a student being on track (with some fixed probability) to attaining English language proficiency.	Target: on track to reaching proficiency in one year _____ Target: on track to reaching proficiency in two years _____ Target: on track to reaching proficiency in three years _____

Note: Under this approach, targets might be described in terms of the expected number of years from attaining English language proficiency. A student is deemed to have met the interim progress²⁸ expectation set for students if that number of years has decreased by one year since the prior year’s ELP test.

Source: Mark Hansen, ELPA21 (personal communication)

Transitioning From a State Legacy ELP Assessment to a Consortium ELP Assessment (Arkansas Department of Education)²⁹

In the 2015–16 school year, the Arkansas Department of Education (ADE) transitioned from the English Language Development Assessment (ELDA; 2008–15) to the ELPA21 assessment. This transition limited ADE’s ability to use empirical analysis of student historical data to inform setting of long-term goals, MIPs, and the state-determined timeline for ELs to achieve English language proficiency. To attain English language proficiency on ELDA, students needed to obtain a score of 5 in all domains, which resulted in low percentages of students exiting between 2008 and 2015.

²⁸ The interim progress expectation is a student-level goal and thus different from a MIP, which is not set at the student level and not required to be calculated on an annual basis.

²⁹ Drawn from the Arkansas Department of Education Consolidated State Plan (approved January 16, 2018) under the ESEA.

A student is deemed to have received a “Profile Score of Proficient” on the assessment if the student receives a score of 4 or 5 on each domain.

Because multiyear state- and districtwide patterns in ELPA21 scores were not available, ADE established its ELP goal using baseline data. ADE set the initial long-term goal and MIPs based on the first two years of ELPA21 data, with plans to reevaluate as additional ELPA21 data become available. ADE will evaluate additional metrics for measuring interim progress in the percentage of ELs attaining English language proficiency as they are established at the consortium level.

In the meantime, ADE has simulated English-proficient performance criteria using different domain-score combinations based on its former ELP assessment. Table 4 details two different proxy exit criteria that allow EL students to exit if they score 5 on all domains except writing, which can be 4 (proxy exit 1) or that score 5 in speaking and listening and 4 in reading and writing (proxy exit 2). ADE’s analyses reveal that more than 50 percent of EL students with an initial ELDA level 3 or 4 reclassify within two to four years under each of two different proxy exit criteria. Students with an initial ELDA level 1 or 2 do not meet proxy exit 1 within seven years. Those with an initial level 2 (in grade bands K–2 and 3–5) meet proxy exit 2 within five to six years, and those with initial level 1 do so after seven years (in grade band 3–5). These analyses confirm empirical research that has documented the role of initial ELP level in influencing the likelihood of a student attaining English language proficiency, with students who enter at early grades with higher ELP levels attaining proficiency faster than their peers who enter at lower initial ELP levels.

ADE also highlights how the point at which EL students demonstrate English language proficiency also varies by domain, an important consideration for ELPA21. ADE proposes to set long-term goals for the percentage of students on track to English Language Proficiency. These goals will be based on 2018 ELPA21 results combined with the initial grade level and initial domain level of English Learners from their entry years using ELDA or ELPA21 as applicable for each student’s entry as an English Learner.

Table 4. ADE Proxy Exit Criteria

	Writing	Reading	Speaking	Listening
Proxy Exit 1	4	5	5	5
Proxy Exit 2	4	4	5	5

Source: Arkansas Department of Education. (n.d.). *Arkansas Consolidated State Plan* (Approved January 16, 2018). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/arconsolidatedstateplanfinal.pdf>

ADE has established that ELs can be considered on track to English language proficiency (i.e., making progress) if they satisfy any of the following conditions:

1. Exit EL status.
2. Meet time expectations on three or more ELPA21 domains.

3. Meet time expectations on all nonexempt ELPA21 domains (exempted domains are those for which there are no appropriate accommodations for an EL with disabilities; e.g., a nonverbal EL who because of an identified disability cannot take the speaking portion of the assessment, as determined, on an individualized basis, by the student's individualized education program (IEP) team, 504 team or by the individual or team designated by the LEA to make these decisions under Title II of the Americans with Disabilities Act (ADA)).

After determining whether a student has met student-level targets, ADE will calculate the percentage of students on track to attaining English language proficiency for all ELs as well as for each of the following EL subpopulations: ELs with disabilities and long-term ELs. The percentage of students on track will be calculated by domain and across all domains.

Based on preliminary 2017 ELPA21 results and initial ELDA or ELPA21 data, ADE initially set the long-term goal for the percentages of students on track to attaining English language proficiency. Now that three years of ELPA21 results have become available and been analyzed, ADE plans to revisit these long-term goals due to the data reflecting a continued adjustment to the assessment transition, particularly for the higher grade levels of students. The difference between former EL exit criteria which included the ELDA assessment and those being used with ELPA21 is an artifact of the transition that is anticipated to reduce in impact for high schools in future years.

Considerations: ELP Goals and Measurements of Interim Progress

Proper development of the ELP long-term goal and MIPs requires multiple steps to ensure that all of the components (i.e., goals and measurements) cohere. States need to set an appropriate English-proficient level score and must examine how long it takes ELs to attain English language proficiency and what progress patterns to proficiency look like on the ELP assessment taking into consideration any EL characteristics a state has included in its model when setting student progress expectations. Based on much empirical evidence, ELP progress expectations over time are most reasonably set faster at lower ELP and grade levels and slower as grade and ELP levels increase. The goals should be sensitive to progress from one year to the next over the entire expected time frame (that is, if the ELP assessment has a wide middle performance band, then the model should be sensitive to movement across that band). Importantly, states should check to see what percentage of students are meeting their individual student-level progress targets and determine whether this aligns in aggregate with the school-level baseline percentages used to set long-term goals for the percentage of students at the school level. Depending on current baseline performance and long-term goals, states can then set MIPs.

Section 3: Constructing and Evaluating the ELP Indicator

This section contains state and mock examples that demonstrate different ways to construct and evaluate an ELP indicator. Examples in this section were shared by states participating in the U.S. Department of Education-sponsored ELP CoP. Some examples focus on student-level targets, while others include details related to both student- and school-level targets. The mock examples (Goldschmidt, 2020) illustrate general principles relevant to ensuring that the ELP indicator is properly constructed and checked so that it supports valid and reliable inferences and provides clear and meaningful signals to stakeholders.

ESEA Requirements

To satisfy the requirements of ESEA section 1111(c)(4)(B)(iv), peer reviewers reviewed each state's consolidated state plan to ensure it included the following, as outlined in the U.S. Department of Education State Plan Peer Review Criteria:³⁰

A.4.iv.d: Progress in Achieving English Language Proficiency Indicator³¹

- Does the SEA describe the Progress in Achieving English Language Proficiency indicator used in its statewide accountability system, including that the SEA uses the same indicator across all LEAs in the state?
- Is the indicator valid and reliable?
- Is the Progress in Achieving English Language Proficiency indicator aligned with the state-determined timeline described in A.4.iii.c.1 of the peer review criteria?
- Does the indicator consistently measure the progress of all English learners, statewide, in each of grades 3 through 8 and in the grade for which such English learners are otherwise assessed under ESEA section 1111(b)(2)(B)(v)(I) during grades 9 through 12?
- Does the SEA's description include the state's definition of English language proficiency, based on the state English language proficiency assessment?

Challenges for States

Some challenges that states faced in developing the ELP indicator include the following:

- Selecting a statistical model to calculate the ELP indicator
- Determining state business rules (*n*-size, aggregating data across years) and how these may affect the model

³⁰ The U.S. Department of Education State Plan Peer Review Criteria, available at <https://www2.ed.gov/admins/lead/account/stateplan17/essastateplanpeerreviewcriteria.pdf>, was issued in March 2017 and provides the criteria by which state plans were reviewed based on statutory and regulatory requirements.

³¹ State Plan Peer Review Criteria, page 12.

- Examining how the indicator takes into account state context (e.g., distribution of ELs across grades)
- Communicating to stakeholders (e.g., policymakers, educators, the public) how the ELP indicator functions and how to properly interpret results from it

Technical Approaches

This subsection provides technical examples, based on mock and actual state data, that evaluate assumptions about the ELP indicator by comparing observed and expected growth in English language proficiency. This subsection also illustrates and discusses sample business rules relating to the ELP indicator.

Evaluating the ELP Indicator: Examining the Growth Model (Mock State Example)

Community of Practice Insight

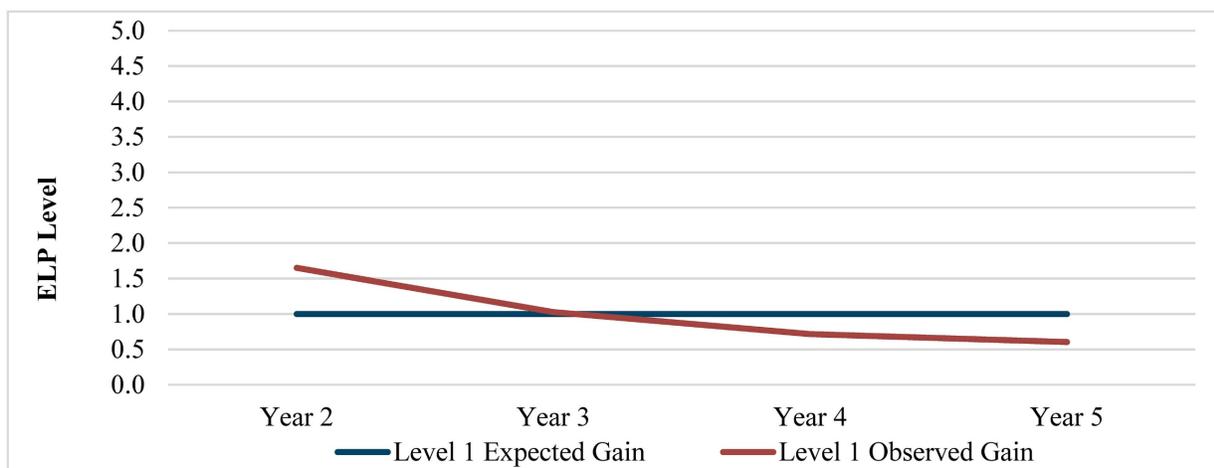
“Given a selected growth model, the SEA needs to decide whether the ELP indicator is based on a single measure (progress in attaining ELP) or multiple measures (attainment of ELP). If the SEA chooses to use attainment of ELP in addition to progress, the SEA must decide how the measures will be combined to create an overall indicator of EL progress. Whether states have a single-measure or multiple-measures indicator, they still need to translate that measure of progress into points, a color, or level that is meaningful within their accountability system.”

—State Support Network Subject Matter Expert

It is important for states to examine their models for expected growth against actual growth. That is, the selected growth model’s results should be plotted against the observed growth to determine whether and to what extent the model adequately represents actual growth. This can provide a check on what results should be expected over time, particularly across grade bands and school segments. Figure 4 displays an example of the effect of constant ELP growth expectations over time. As this figure illustrates, ELs can more easily meet constant growth expectations early on—at lower grade and ELP levels—relative to later on at higher grade and ELP levels. The red line (observed gains) is higher than the blue line (expected gains) at Year 2 but observed gains dip below expected gains in Years 3, 4, and 5. The discrepancy between observed and expected gains will very likely disproportionately appear for middle school ELs, who will generally have been in the program longer. This in turn can disproportionately impact middle schools, which typically have more ELs who have become long term and “off track” (i.e., have missed growth expectations in earlier years).

This would suggest that the state should either adjust its student-level progress targets to reflect the curvilinear nature of ELP growth or (potentially more problematically) adjust its school-level target for the percentage of ELs attaining constant ELP growth downward.

Figure 4. Example of Effect of Linear (Constant) ELP Growth Expectations Over Time



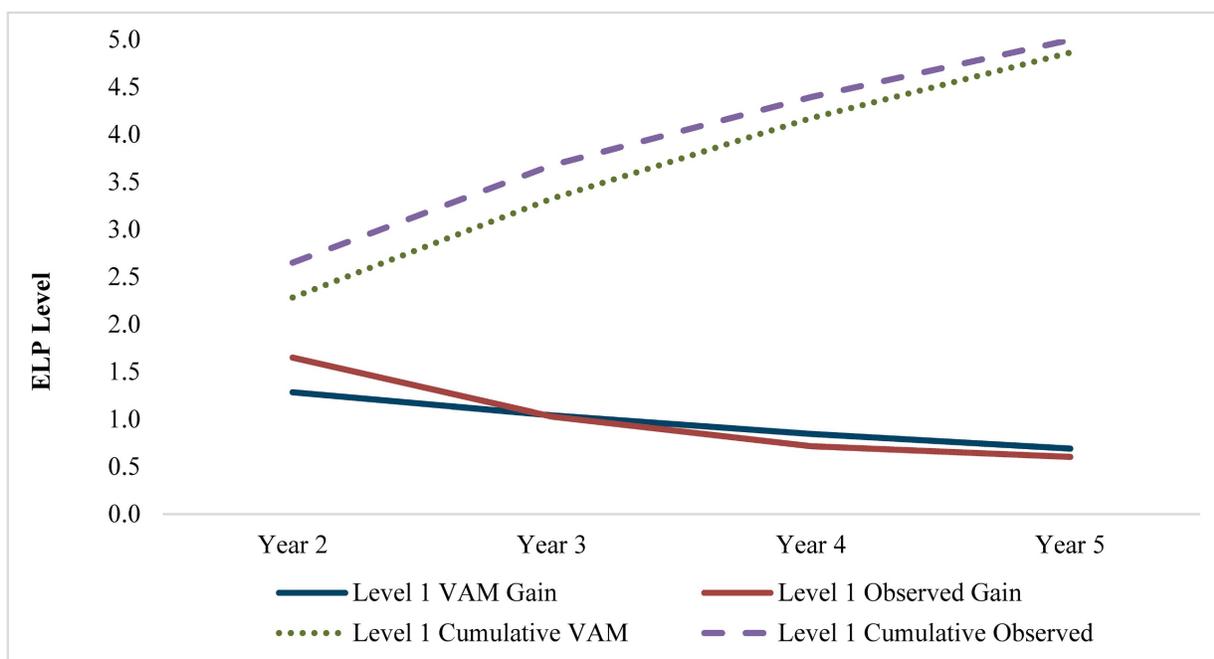
Source: Goldschmidt (2020)

Figure 5 displays the effects of a model of year-over-year growth that better matches observed growth and cumulative growth in this state. Compared to the example in Figure 4, this model appears to better align to actual EL trajectories over time. The model results shown in Figure 5 appropriately capture the decreasing annual gains (and the slowing of cumulative progress), whereas the model in Figure 4 creates expectations that are too low in the early years and too high in the later years. A state could approximate a more realistic expected trajectory by including more student characteristics in their model. However, the more variables included in the model, the more complex and less transparent it is to stakeholders. The aim is to match the observed trajectories with as few variables as possible, which will be easier to explain as it is not based on so many conditions.

An empirical question that the state should examine is whether having additional variables improves precision and reliability. Precision is an estimate of the variation around the progress estimate, while reliability can take on several different definitions. Precision at the individual student level is less critical as the results for accountability are aggregated to the school level.³² Although reliability is often thought of with regard to the assessment (e.g., internal consistency), this conception is not relevant in this situation. That said, aggregating results from even moderately reliable assessments provides sufficiently reliable estimates at the school level. Another conception of reliability is stability (i.e., how stable the results are from one year to the next). Again, at the student level, from one year to the next there may not be as much stability because progress is inherently capturing movement. At the school level, results will be more stable, but there is no specific value that is optimal because a value of 1 implies that schools cannot change their status in terms of ELP progress. Reliability as stability is important to consider but is secondary to considering reliability as an indicator of how well observed school means of the growth model represent true differences among schools in ELP performance. As noted, whether any of these conceptions of precision and reliability are substantively influenced by including additional variables in the model is both an empirical question and a professional judgment.

³² Precision increases with the square root of the school N.

Figure 5. Effect of Value-Added Model Expectations for ELP Growth Over Time



Source: Goldschmidt (2020)

Technical Check

Ensure that the state ELP indicator yields sufficient variation among schools to satisfy the assumption that schools differ in their ability to facilitate English language development for EL students. Without sufficient variation, the ELP indicator cannot meaningfully differentiate among schools.

Examining the Variation in ELP Progress by Grade (Mock Example)

In this example, accountability toward English language proficiency will occur through a single measure, ELP growth targets that index the extent to which students gain ELP over a reasonable amount of time with a timeline of five years after initial classification. In this mock example, ELs are assigned to a cohort based on the students' grade level and English language proficiency at entry as indexed by the state ELP assessment. Students remain in their tracking cohort regardless of migration to different schools or districts, meaning that students are held to the original time and progress expectations irrespective of the school they attend.

As an internal check to determine whether there is in fact sufficient variation among schools in the progress outcome, or the ELP indicator, to allow schools to be meaningfully differentiated, the state in this mock example examined the extent to which ELP progress varied among schools. To determine whether there is sufficient variation among schools, the state followed the steps in Goldschmidt (2020) and calculated the intercluster correlation coefficient (ICC). The ICC is not an indicator of reliability; rather, it is a metric that indicates the proportion of variation that is between schools in the outcome. Examining the ICC is a common method applied to data when the data have a structure in which subjects are clustered within a higher organizational unit (e.g., students within schools). The ICC can be calculated from traditional ANOVA (analyses of variance)

sums of squares but are now more readily estimated in statistical software either automatically or with resulting variance components. Even small ICCs (e.g., .05) indicate that there is a meaningful amount of variation among schools in the outcome. The results in Table 5 demonstrate substantive variation among schools in the progress outcome to allow schools to be meaningfully differentiated. Thus, the results confirm that the ELP indicator will substantively contribute to holding schools accountable for English language progress. Had ICCs in Table 5 been consistently less than .05, then the state would benefit from reconsidering how the ELP indicator was calculated.

Awarding Points to Schools for Meeting Annual Interim Growth Targets and On-Time Attainment of ELP (Delaware Department of Education)³³

In order to set the expectation that all ELs make annual progress toward attaining ELP within the applicable timeline, the Delaware Department of Education plans to award points to schools for students who meet the annual interim growth targets and on-time ELP attainment. Index scores for ELP growth will range from 0.00 to 1.10, with the following rules:

- 0.00 assigned to students who do not participate in the assessment or show no growth
- 0.01 to 0.99 assigned to students who have made growth toward the target³⁴
- 1.0 to 1.10 assigned to students who have reached (1.00) or exceeded the target (1.01 to 1.09), with a maximum bonus for students exceeding the target by 10 percent or more (1.10)³⁵

The following tables illustrate how the index scores are calculated for nonlinear annual growth targets and the on-time attainment of English language proficiency. Table 5 summarizes the accountability rules for the years up to and including the year the student should attain English language proficiency. A bonus of 10 percent is awarded to the EL student’s score when English language proficiency is achieved prior to the specified year of attainment.

Table 5. Student-Level ELP Growth Index Score by Student Outcome (Rules for Years Up to and Including the Designated Attainment Year)

Year	Student Outcome			
	Nonparticipant	No Progress Toward Target	Progress Toward Target but No Grade-Level Attainment	Grade-Level Attainment Target Met or Exceeded
Before designated attainment year	0.00	0.00	0.01–1.10* (CY SS – PY SS)/ (IT SS – PY SS)	1.10
In designated attainment year	0.00	0.00	0.01–0.99 (CY SS – PY SS)/ (AT SS – PY SS)	1.00–1.10* (CY SS – PY SS)/ (AT SS – PY SS)

* Index score of 1.10 is the maximum (10 percent credit for exceeding target by 10 percent or more).

Note: PY SS = previous year’s scale score. CY SS = current year scale score. IT SS = interim target scale score. AT SS = grade-level attainment target scale score. The school-level indicator would be (sum of the student-level

³³ Drawn from the Delaware Department of Education Consolidated State Plan (approved June 3, 2019) under the ESEA.

³⁴ For example, a student whose target is 100 scale score points of growth and progresses 90 points would earn .9 points.

³⁵ For example, a student whose target is 100 scale score points of growth and progresses 105 points would earn 1.05 points. A student with the same target progressing 120 points would earn 1.10 points.

scores)/(the number of participants and non-participants)*100. The possible range of school scores is 0 to 110.
 Source: U.S. Department of Education. (n.d.). *Delaware Consolidated State Plan* (Approved October 31, 2018).
 Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/deconsolidatedstateplanfinal.pdf>

Table 6 highlights the rules that apply if a student does not attain English language proficiency within the designated time frame. As shown, the index score for an EL student who attains English language proficiency one year late decreases by 25 percent; by 50 percent if two years late; and by 75 percent if three or more years late. Such rules are designed to signal to educators the urgency of addressing the needs of ELs requiring greater support to attain the English proficient performance standard.

Table 6. Student-Level ELP Growth Index Score by Student Outcome (Rules for Years After the Designated Attainment Year)

Year	Student Outcome		
	Nonparticipant	Grade-Level Attainment Target Not Met	Grade-Level Attainment Target Met
1 year late	0.00	0.00	0.75
2 years late	0.00	0.00	0.50
3+ years late	0.00	0.00	0.25

States must develop a series of business rules that operationalize the state’s consolidated plan. States must anticipate particular technical challenges to their accountability system design and develop solutions that can be applied systematically to resolve them. The technical checks that follow illustrate a few examples of how to detect and address issues that may arise related to the ELP indicator.

Technical Check

- Examine the distribution of EL students across districts in the state, as well as schools in each individual district, regarding initial ELP levels, grade on entry, recently arrived ELs, long-term ELs, EL students with a disability, and SIFE.
- Establish decision rules for how to attribute ELP growth results for those students moving from one school to another, particularly in the following scenarios (regardless of accountability model):
 - Students who transfer districts
 - Students naturally transitioning school segments (e.g., elementary to middle school, middle to high school)
- Examine the impact of the state’s n-size on the ELP indicator across schools and consider what percentage of ELs are excluded from the accountability model based on minimum n-size.

***Growth on ELP Assessment and Students on Track to English Language Proficiency as Indicators of Progress (Oregon Department of Education)*³⁶**

EL stakeholders in Oregon support the use of two measures for the ELP progress indicator, with the aim to support a comprehensive view of English language progress. These measures are (a) percentage of ELs on track to ELP and (b) ELP growth. The on-track to ELP indicator is based

³⁶ Drawn from the Oregon Department of Education Consolidated State Plan (approved August 30, 2017) under the ESEA.

on the ELPA21 and is criterion-referenced because it measures EL progress compared with a fixed set of expectations for ELP attainment.

The ELP growth indicator is norm-referenced because it measures EL progress on ELPA21 as compared to peers with similar characteristics (e.g., prior achievement, enrolled grade, time identified as EL). Oregon will report each measure by domain (i.e., reading, writing, listening, and speaking) as well as the combination of all four domains. This will be done for all current ELs, as well as for the following subgroups of ELs: SIFE, dual-identified ELs (ELs with disabilities), recently arrived, long-term, and ELs in bilingual programs, all of which evidence suggests may require longer time frames to achieve English language proficiency.

The first measure—percentage of ELs on track to English language proficiency—uses the initial ELP level, current ELP level, and years identified as EL to determine whether the student is on track to English language proficiency. ELs are considered “on track” to English language proficiency if they meet or exceed the trajectory expectations across all four ELPA21 domains given their initial ELP level and years identified as an EL. For example, an EL with an initial ELP level of 1 on all four domains and who has been identified as an EL for four years would need a level 3 or higher on all four domains to be on track to English language proficiency. Oregon intends to calculate the percentage of students on track to English language proficiency for each EL subgroup, domain, and a combination across all domains. Table 7 includes a mock display of the percentage of ELs on track to English language proficiency.

Table 7. Sample Display of Percentage on Track by Student Group, Domain, and Combined

Student Groups	On Track to ELP by Domain				On Track to English Language Proficiency (All Domains)
	Reading	Writing	Listening	Speaking	
All ELs	70%	63%	74%	76%	72%
SIFE	41%	39%	45%	48%	44%
Dual-Identified	34%	33%	38%	39%	35%
Recently Arrived	67%	64%	69%	71%	68%
Long-Term	37%	36%	41%	43%	39%
Bilingual Program	72%	67%	76%	77%	73%

Source: Oregon Department of Education. (n.d.). *Oregon Consolidated State Plan* (Approved August 30, 2017). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/orconsolidatedstateplan.pdf>

The second measure is ELP growth. Oregon plans to measure ELP growth using median growth percentiles under a “modified conditional status model” (see Castellano & Ho, 2013) because of the small number of ELs in high school grades (see Goldschmidt & Hakuta, 2017). The specification of this model includes the current year ELPA21 domain scale score as the outcome and the prior year ELPA21 domain scale score as the covariate while adjusting for time identified as an EL, current enrolled grade, SIFE, and dual-identification (see Hakuta & Pompa, 2017). Oregon will transform the residual (i.e., the difference between the observed and predicted current year ELPA21 domain scale score) to a percentile. This percentile is known as the percentile rank of the residual and is equivalent to a student growth percentile (see Castellano & Ho, 2013).

The student growth percentile is the ranking of the student on the current year ELPA21 domain scale score as compared with academic peers with the same prior year ELPA21 domain scale score and who are in the same grade, have the same time identified as an EL, and are SIFE and/or dual-identified (if the student is SIFE and/or dual-identified). Oregon will calculate the median growth percentile for each student group and report it by domain and the combination of all four domains. Table 8 shows a mock example of sample median growth percentiles.

Table 8. Sample Display of Median Growth Percentiles by Student Group, Domain, and Combined

Student Groups	Median Growth Percentile by Domain				Median Growth Percentile (All Domains)
	Reading	Writing	Listening	Speaking	
All ELs	51	49	54	59	54
SIFE	31	29	32	33	32
Dual-Identified	29	27	30	32	30
Recently Arrived	49	46	51	54	50
Long-Term	32	30	33	36	32
Bilingual Program	56	54	57	61	57

Source: Oregon Department of Education. (n.d.). Oregon Consolidated State Plan (Approved August 30, 2017). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/orconsolidatedstateplan.pdf>

Technical Check

- States that choose to include the percentage of EL students attaining English language proficiency on the state ELP test in the ELP indicator should check whether they are systematically rewarding schools that have poor progress but a higher proportion of students who attain English language proficiency. This pattern may also relate back to the initial ELP level of students, which affects time to English language proficiency.
- Not all states use multiple criteria for exit. States that do use multiple criteria for exiting ELs from EL status as part of their standardized statewide exit procedures—which must include the ELP assessment—should examine the proportion of ELs that attain *each* of the exit criteria. Because ESEA requires all current ELs to be assessed annually using the state ELP assessment, those EL students who score English-proficient on the state ELP assessment but do not attain other criteria needed to exit must continue to be assessed on the state ELP assessment annually until meeting all criteria to exit. States will need to consider how their accountability model will treat those EL students who score English-proficient in the prior and current year (i.e., consider them to have made progress) versus those who score English-proficient in the prior year but not in the current year.

Considerations: The ELP Indicator

Building on Section 2, which discussed how the SEA selects a proficiency cut score and the time to reach that score, Section 3 presented considerations and options for modeling ELP progress over time. There are many options to translate actual progress into an indication of student success. For example, a state can use a VAM or an SGP model, set a specific threshold, and count the number of students meeting the threshold. A state can also aggregate the actual values produced by the model or use year-over-year changes in scale scores or ELP levels to designate

success.³⁷ Using gains, a state may set a threshold and count the number of students meeting the threshold, or it could aggregate the proportion of expected growth (e.g., Delaware).

The results of the modeling option are aggregated in some way to obtain the schoolwide ELP indicator results. This is discussed in detail in Section 4.

³⁷ Although year-over-year changes are often less stable than VAM and SGP estimates, they are also less influenced by small n-sizes.

Section 4: Translating Student-Level Progress Into a School-Level ELP Indicator

This section contains state and mock examples that demonstrate different ways that a state can translate student progress into a school-level ELP indicator. The state examples featured in this section are those that states shared during the ELP CoP and mock state examples.

ESEA Requirements

To satisfy the requirements of ESEA section 1111(c)(4)(B)(iv), peer reviewers reviewed each state’s consolidated state plan to ensure it included the information outlined in the U.S. Department of Education State Plan Peer Review Criteria Section A.4.iv.d (i.e., the same criteria outlined in Section 3 of this guidebook).^{38,39}

Challenges for States

Some of the challenges that states face in translating student-level progress into a school-level ELP indicator include the following:

- Ensuring that differences observed in student performance reflect true differences in student performance
- Ensuring the reliability of the ELP indicator
- Determining appropriate growth targets that lead to a reliable ELP indicator (i.e., the ability to differentiate among schools with respect to inferences about a school’s ability to facilitate progress toward English language proficiency)
- Addressing threats to ELP indicator reliability when a new, more rigorous ELP assessment results in many ELs no longer meeting progress targets

Technical Approaches

This subsection provides technical examples, using mock data and actual state data, that focus on how states translate student progress targets into a school-level ELP indicator.

Generating ELP Indicator Points (Mock State Example)

When states use a point system to award schools points for specified amounts of student progress, how student progress is converted into ELP indicator “points” matters because it affects the likelihood of schools’ movement across the point spectrum.⁴⁰ A state using a point system to credit schools should consider how much actual change in ELP progress is required to

³⁸ The U.S. Department of Education State Plan Peer Review Criteria, available at <https://www2.ed.gov/admins/lead/account/stateplan17/essastateplanpeerreviewcriteria.pdf>, was issued in March 2017 and provides the criteria by which state plans were reviewed based on statutory and regulatory requirements.

³⁹ State Plan Peer Review Criteria, page 12.

⁴⁰ A state may, of course, choose to use a different system, such as a conjunctive model as described previously.

move beyond a cut point on the ELP indicator. Technical considerations are provided in the following Technical Check box.

Technical Check

In translating ELP growth measures into ELP indicator points, states should use nonlinear transformations to check that its identification of schools (e.g., in the bottom 5 percent) is not simply an artifact of model choice.

Nonlinear transformations result in different cuts than linear transformations. Consider, for example, the case of a state measuring ELP progress as the percentage of students in a school meeting their individual progress targets. The school's ELP progress increases from 20 percent to 40 percent meeting target. This state calculates the ELP indicator as $(\text{EL Progress}^2) * 100$. Applying 20 percent and 40 percent to this equation, the results are: $(.2^2) * 100 = 4$ and $(.4^2) * 100 = 16$. Moving from 20 percent to 40 percent of ELs meeting individual student progress targets results in a substantial increase in points assigned (from 4 to 16). Such a substantial increase may distort the signal sent by points assigned versus actual progress. That is, stakeholders may not understand how change in points relates to actual student progress. This is an example why nonlinear transformations might create problematic results.

Moreover, small changes in performance near a critical cut point (e.g., bottom 5 percent) can yield large increases in ranking. Moving above the fifth percentile may require substantial performance gains at the school level, and schools could become stuck in the bottom 5 percentiles. On the other hand, a flat "tail" distribution could mean that measurement error might account for the school's movement out of comprehensive-support-and-improvement school status.

Awarding Points to Schools for Meeting Annual Interim Growth Targets and On-Time Attainment of ELP (Delaware Department of Education)⁴¹

Delaware uses the individual results presented previously in Tables 5 and 6 to create its school-level ELP indicator, which is equal to: $(\text{sum of the student-level scores}) / (\text{the number of participants and non-participants}) * 100$. While the possible calculated range of school scores is 0–110, Delaware caps school scores at 100.

For example, if a school has four students who earned 0, 0.5, 1.0, and 1.1 points, the school score would be 0.65 $((0 + .5 + 1 + 1.1) / 4 = 2.6 / 4 = 0.65)$. If a school has four students who earned 1.0, 1.0, 1.1, and 1.1 points, the school score would be 1.0 $((1 + 1 + 1.1 + 1.1) / 4 = 4.2 / 4 = 1.05)$, which is capped at 1.0).

⁴¹ Drawn from the Delaware Department of Education Consolidated State Plan (approved June 3, 2019) under the ESEA.

Section 5: Integrating the ELP Indicator Into the State Accountability System for Annual Meaningful Differentiation

This section contains state and mock examples that illustrate how some states, including those in the ELP CoP, have incorporated EL progress into their overall Title I accountability systems.

ESEA Requirements

To satisfy the requirements of ESEA section 1111(c)(4)(C), peer reviewers reviewed each state's consolidated state plan to ensure it included responses to the following questions from the U.S. Department of Education State Plan Peer Review Criteria:⁴²

- Does the SEA describe its system of meaningfully differentiating, on an annual basis, all public schools in the state?
- Is the state's system of annual meaningful differentiation based on all indicators in the state's accountability system?
- Does the state's system of annual meaningful differentiation include the performance of all students and each subgroup of students on each of the indicators in the state's accountability system?
- Does the SEA describe the weighting of each indicator in its system of annual meaningful differentiation, including how the weighting is adjusted for schools for which an indicator cannot be calculated because of the minimum number of students (e.g., for the Progress in Achieving English Language Proficiency indicator)?
- Do the Academic Achievement, Other Academic, Graduation Rate, and Progress in Achieving English Language Proficiency indicators each receive substantial weight individually?
- Do the Academic Achievement, Other Academic, Graduation Rate, and Progress in Achieving English Language Proficiency indicators receive, in the aggregate, much greater weight than the School Quality or Student Success indicator(s), in the aggregate?

Challenges for States

Some of the challenges that states face in using the ELP indicator in the state accountability system for annual meaningful differentiation include the following:

- Integrating the ELP indicator in the state's Title I accountability system
- Ensuring that the state's accountability system drives improvement in ELs' progress in achieving ELP

⁴² The U.S. Department of Education State Plan Peer Review Criteria, available at <https://www2.ed.gov/admins/lead/account/stateplan17/essastateplanpeerreviewcriteria.pdf>, was issued in March 2017 and provides the criteria by which state plans were reviewed based on statutory and regulatory requirements.

- Operationalizing “substantial weight” and “much greater weight in the aggregate” within the accountability system to ensure annual meaningful differentiation
- Indicating that some schools are performing better with EL students than other schools

Technical Approaches

States should consider normalizing indicators that are based on different metrics before combining them into a composite (e.g., putting all the indicators on the same scale). Combining indicators in different ways can have differential impacts on schools when the state translates growth monitoring into a point system. The degree of correlation between individual student- and school-level changes will have an impact on how schools cluster in performance over time.

Following are examples of technical checks that states can perform to ensure that their ELP indicator fits into the broader system of accountability.

Technical Check

SEAs should consider the following regarding how the ELP indicator fits in the broader state accountability system:

- Relationship between the ELP indicator and the other indicators as well as between the ELP indicator and the overall accountability system result
- Reliability of the ELP indicator and of the overall results (with and without the ELP indicator)
- Differentiation of schools on the ELP indicator versus on overall results (e.g., using intra-class correlation coefficient as a measure of differentiation)
- Sensitivity of the ELP indicator and of overall results to school input characteristics
- Sensitivity analyses for change in scores, or ranking of schools around the minimum n cut-point (with and without the ELP indicator) for schools that meet minimum n -size

Examining the Relationship Between the ELP Indicator and Other Accountability Indicators: Assessing the Impact of Score Normalization on the ELP Indicator Using an Equal Weight Check (Mock State Example)

Another check on the ELP indicator is to examine how it behaves with other indicators (e.g., Academic Achievement indicator, Other Academic indicator, Graduation Rate indicator, and School Quality or Student Success indicator or indicators) in the state accountability system. Normalizing all indicator scores⁴³ allows them to be more smoothly aggregated. Table 9 examines the impact of including the ELP indicator in its raw versus normalized form. The relationship (correlation) between the overall score and each of the individual scores is different depending on whether it is normalized or not. The results in Table 9 present the correlations between the overall school accountability result and the percentage of ELs in the school (i.e., Whether Student Is an EL),⁴⁴ Content Growth, and ELP Progress. The results indicate that ELP progress is more highly correlated with the overall accountability result than content growth

⁴³ That is, transforming them so that they have a mean of 0 and a standard deviation of 1.

⁴⁴The percentage of students who are ELs in a school is not an accountability indicator; however, it is worth reviewing the correlations in context.

when considering all schools or schools that have at least 20 ELs. The results in Table 9 also indicate that the overall accountability result is more highly correlated to the percentage of ELs in a school than content growth. When scores are normalized, correlations are more similar among the indicators presented in Table 9.

Table 9. Impact of Normalization on the ELP Indicator Using an Equal Weight Check

Indicator	Equally Weighted	
	Raw	Normalized
Overall score (correlation with)		
Whether Student Is an EL	0.66	0.57
Content Growth	0.17	0.16
ELP Progress	0.28	0.53
Min $N < 20$		
Whether Student Is an EL	0.78	0.57
Content Growth	0.19	0.39
ELP Progress	-0.010	-0.050
Min $N \geq 20$		
Whether Student Is an EL	0.57	0.57
Content Growth	0.16	0.26
ELP Progress	0.53	0.46

In addition, how states translate their models into point assignments matters. Table 10 illustrates this effect on the same value-added analysis. The top rows in panels A and B, respectively, display the results using a yes/no approach (whether or not a student met/exceeded expectations). The second rows in panels A and B, respectively, display the same results using a cumulative distribution function (CDF) approach.⁴⁵ The distribution of points differs depending on the selected approach, most apparently at the tails of the distribution. Table 10 lists the average points earned by schools at different points in the distribution (5th percentile, etc.) under the different approaches. Panel A shows that on average, schools at the 5th percentile will earn 0 points under the yes/no approach and 2.82 points under the CDF approach. There is slightly less variation in point assignments under the CDF approach.

States should also consider how much movement is required for schools to advance beyond the 5th percentile of the distribution. Panel B displays the degree of difficulty in advancing beyond different percentile categories. For example, under the yes/no approach, a school would have to grow 0.81 standard deviations to move from the 5th to the 10th percentile, compared to 0.40 standard deviations to make that same move under the CDF model. Clearly it is easier for a school to advance out of the 5th percentile category under the CDF approach.

⁴⁵ The CDF approach uses the area under a normal curve that lies to the left of the normalized score.

Table 10. Impact of Different Point Assignment to Same Model

Percentiles	5 th	10 th	25 th	50 th	75 th	90 th	95 th
Panel A. Weighted Average (10 points possible)							
Yes/No	0.00	2.00	3.59	5.00	6.39	8.33	10.00
CDF	2.82	3.42	4.23	5.03	5.94	6.86	7.49
Panel B. Effect Size (to advance to next percentile category)							
Yes/No	N/A	0.81	0.64	0.57	0.56	0.79	0.67
CDF	N/A	0.40	0.54	0.53	0.61	0.62	0.42

Notes: CDF = cumulative distribution function. Read Table 10 as: The first two rows indicate how many points a school earns at the given percentile. The second two rows indicate how much improvement (in effect size) a school must demonstrate to move to the next percentile category.

Source: Goldschmidt (2020)

Technical Check

- Sensitivity analyses can reveal how changing the weight of the ELP indicator will change the ranks of schools. The SEA must decide whether any resulting change is consistent with expectations, perceptions, and other measures.
- The SEA must decide how to distribute the ELP indicator points to schools that have ELs but do not meet the minimum *n* size. The state should determine whether ELP indicator points are distributed equally (reducing the denominator) or among all remaining indicators, neither of which produces exactly equal results. If the remaining indicators are “easier” to make points on, then schools with more ELs are disadvantaged. “Easier” can refer to higher average levels and/or greater ease of up/down movement.
- If the ELP indicator is more volatile than other indicators, then it will appear that schools are doing a much less consistent job with ELs.

Determining a Weight for the ELP Indicator (Mock State Example)

It is important to examine the impact of changing the weight of the ELP indicator on school rankings. States should consider the following technical checks on weighting the ELP indicator as summarized in the following Technical Check box.

Technical Check

Sensitivity analyses indicate that the correlation between results with an ELP indicator weighted at 10 percent versus 20 percent is 0.9. This finding might suggest that school ranks do not change much. However, if the SEA intends to use a conjunctive model or specific cuts, then the absolute change in scores may have an impact on school ranks.

Dynamic Weighting of the ELP Indicator (Louisiana Department of Education)⁴⁶

The Louisiana Department of Education asserts in its ESEA consolidated state plan that it is committed to measuring progress to achieve English language proficiency for all ELs. It further asserts that this commitment informed its decision to publicly report EL performance for all schools with 10 or more students in the subgroup. Louisiana’s plan ensures a substantial weight for the ELP indicator (as required by statute) but further ensures accountability for all EL students, through a weighted ELP indicator.

Louisiana’s ELP indicator awards points for all ELs making annual progress toward attaining English language proficiency. Progress is counted as either meeting exit criteria and/or meeting or exceeding annual ELP targets based on a student’s baseline proficiency level.

Business rules for the weighting of the ELP indicator in Louisiana include the following:

- Every EL student that has been enrolled in schools in the United States for two or more years⁴⁷ is included in school’s performance score (SPS)⁴⁸ regardless of the number of ELs that a school serves.
- Every EL student’s progress towards English language proficiency as measured by the ELP assessment carries equal weight as the other state content assessments a student is expected to take on average in a given grade. Students in grades 3–8 are expected to take reading/language arts, mathematics, science, and social studies assessments. Reading/language arts and mathematics assessments are double-weighted, so students in grades 3–8 are expected to have six assessment units in a given year: two units from reading/language arts, two units from mathematics, one unit from science, and one unit from social studies). Therefore, the ELP assessment is weighted at six units for students in grades 3–8. Students in high school grades are expected to take six assessments over the course of high school (i.e., English I, English II, Algebra I, Geometry, US History, and Biology), where each assessment is weighted as one unit; however, high school students are only expected to take an average of two assessments per year (e.g., a 9th grade student takes English I and Algebra I assessments). Therefore, the ELP assessment is weighted at two units for high school students.
- Every school with at least 10 EL students has the performance of EL students disaggregated as a subgroup and reported separately, in addition to the SPS, which already includes EL results in weight relative to core academic tests.⁴⁹

As noted previously, ELP assessment results are included in weight relative to the core academic subjects within the Assessment Index⁵⁰ for every EL student. Thus, both ELP and R/LA and

⁴⁶ Drawn from the Louisiana Department of Education Consolidated State Plan (approved August 8, 2017) under the ESEA.

⁴⁷ Because measuring ELP progress requires two data points.

⁴⁸ The SPS refers to the overall rating for a school.

⁴⁹ Schools with fewer than 10 ELs assess EL performance and include assessment results in the SPS, but data are not disaggregated by EL status.

⁵⁰ The Assessment Index is part of the SPS calculation in Louisiana.

mathematics assessments are substantially weighted in the formula. As illustrated in Table 11, points are awarded as follows:

- Award index points for each core academic assessment score based on the achievement level, and to each ELP assessment score based on progress from the baseline.
- Weight each subject index score. ELP is weighted by six or two (depending on the grade level) such that it is equivalent to the number of core academic assessment units that a student would be expected to take.

The Assessment Index including the ELP indicator makes up 65 percent to 70 percent⁵¹ of the elementary/middle school overall school rating and 12.5 percent of the high school overall school rating. Other variables that contribute to a school’s rating at the elementary/middle school levels include ELA and mathematics growth (25 percent) and ninth-grade credit accumulation (only for schools with 8th grade; 5 percent). Other variables at the high school level include ELA and mathematics progress (12.5 percent), ACT and WorkKeys scores (25 percent), graduation rates (25 percent), and strength of diploma (25 percent). Strength of diploma is a college- and career-ready (CCR) index that awards additional points based on credential attainment, such as passing an AP test. Additionally, both elementary/middle and high schools have 5 percent of their formula assessed on an Interests and Opportunities Index, which measures students’ access to enrichment opportunities.

To calculate a school’s Assessment Index, all weighted index scores are summed and divided by the sum of all weights applied from the table (please see Examples 1 and 2 below).

Table 11. Assessment Area and Weighting in Example State Accountability System

Assessment Area	Unit Weight	
<i>Grades K–8</i>		
Core academic content	6	6
English language proficiency	6	6
<i>Grades 9–12</i>		
Core academic content	2	2
English language proficiency	2	2

Source: Louisiana Department of Education. (n.d.). *Louisiana Consolidated State Plan* (Approved August 8, 2017). Retrieved from <https://www2.ed.gov/admins/lead/account/stateplan17/lastateplan882017.pdf>

Assessment Index Calculation: Example 1

- An elementary school has 100 total students in tested grades with two years of test data (i.e., have a baseline), 50 of whom are ELs.
- All 100 students will take all core academic content assessments, weighted at six units per student (100 x 6 = 600).

⁵¹ As of 2019-20 (previously 70 to 75 percent).

- The total number of test units is 600 content test units + 300 ELP test units ($50 \times 6 = 300$) = 900 total test units.

Assessment Index Calculation: Example 2

- Under Louisiana’s plan, even when the specific performance of ELs cannot be publicly reported because of privacy concerns (i.e., when the number of ELs falls below the state’s *n*-size), the school is still accountable for success with EL students.
- An elementary school has 100 total students in tested grades with two years of test data (i.e., have a baseline), nine of whom are ELs. ELP assessment results are included in the calculation but are not reported separately because the *n*-size is less than 10.
- All 100 students will take all core academic content tests, weighted at six units per student ($100 \times 6 = 600$).
- The total number of test units is 600 content test units + 54 ELP test units ($9 \times 6 = 54$) = 654 total test units.

Connecting ELP Goals, ELP Indicator, and Annual Meaningful Differentiation

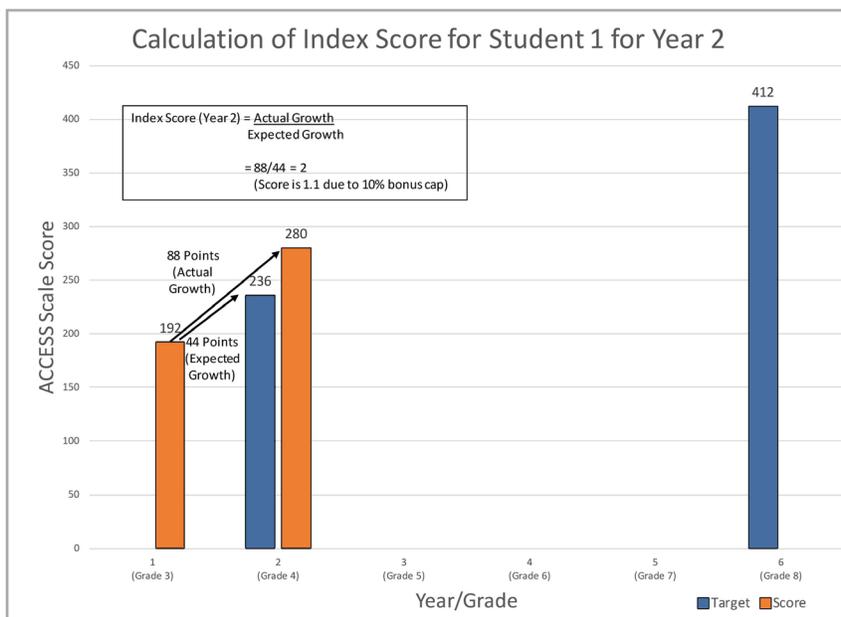
The following examples describe and then apply an ELP indicator model. The aim is to connect ELP goals, the ELP indicator, and annual meaningful differentiation. Note that this example builds on each of the four steps presented in the previous sections of this guidebook:

- Step 1: Specify Individual EL’s ELP Progress Expectation
- Step 2: Calculate School-Level ELP Indicator Results
- Step 3a: Establish Basis for Annual Meaningful Differentiation
- Step 3b: For purposes of this example, Rank Order School Status Results, then Rank Order School Change Results
- Step 4: Identify Schools

State Example of Connecting ELP Goals, the ELP Indicator, and Annual Meaningful Differentiation: Delaware Department of Education

The Delaware model uses a growth-to-target approach that annually calculates a target for each student based on where they are and the attainment score to reach proficiency. This model is depicted in Figure 6. In Figure 6, the orange columns represent observed student scores, and the blue columns represent the target score. The student’s initial score is 192; given this initial score, the expected score in Year 2 is 236, but the student actually scored 280. Although the student made twice the progress that was expected (88 versus 44 points), the student earns 1.1 points due to the cap Delaware has placed on student growth exceeding expectations.

Figure 6. Calculating Index Scores in the Delaware Model



Source: Delaware Department of Education, personal communication.

As noted, the School Score = (Sum of Student Index Scores)/(Count of all students) * 100 and is capped at 100 (range 0–100) (Delaware Department of Education, personal communication with the State Support Network).

The Delaware model of converting growth into points is very straightforward. A key monitoring element for Delaware is the extent to which a student falling behind is able to meet high subsequent growth targets, given the known slowing of English language progress. In the aggregate, monitoring would examine whether the aggregation rules (above) then result in some schools (e.g., middle schools) being disadvantaged because students who are still ELs in middle school have likely missed targets and are generally less likely to meet more aggressive targets to catch up.

Next Steps for SEAs

Incorporating ELs in state accountability systems to meet the requirements of the ESEA presents many technical challenges but also great opportunities. State approaches to meeting these requirements vary widely in both design and implementation. The variability in state approaches highlights the importance of taking into account state context and stakeholder input as well as the need for coherence with other indicators in the accountability system.

As states move fully into the implementation phase, they should conduct ongoing checks to ensure that the accountability system is working as the state intends—that is, that it identifies the correct ELs and schools meriting comprehensive and targeted support and improvement and, consequently, drives improvement in educational practice and in educational outcomes for these ELs. In addition, it may be beneficial for a state to concurrently make efforts to present technical information in ways that make it comprehensible for educators, parents, and students in the state.

Glossary

Conjunctive rules: In the context of growth monitoring, conjunctive rules apply when a school needs to meet two or more distinct criteria in order to receive the full amount of points or “credit.” For example, a student must score a 5 overall *and* score at least a 4 on reading in order to meet the target.

Disjunctive rules: In the context of growth monitoring, a disjunctive rule implies that a student met at least one of the criteria. For example, a school receives credit for student growth if the student either met the growth target *or* achieved proficiency.

Dynamic weight: A dynamic weight is a weight that has a set method for calculation (a fixed equation), and the equation results in different weights based on the results of the equation. For example, if the weight equation for the ELP indicator is $W = \text{Number of ELs} / \text{Total Number of Students}$, then even though W is calculated the same way for each school, W will differ depending on both Number of ELs and the Total Number of Students.

Equipercentile equating: A process through which assessment results of one assessment can be aligned with another assessment. The process includes two steps. The first step entails identifying the percentile rank of each score on the assessment that denotes proficiency (in the case of the ELP assessment). In the second step, these percentiles are used to place the new scores at the same location. For example, if the proficiency cut score was at the 75th percentile on one assessment, proficiency can be placed at the 75th percentile on the other assessment.

Fixed weight: A weight is fixed when the SEA selects a single weight for a given indicator.

Hybrid: A hybrid decision rule combines two rules. For example, a school may earn credit if the composite gain is equal to or greater than a particular value and the domain score did not decrease. This rule combines a composite and a conjunctive rule.

Measurements of interim progress: Targets toward a state’s long-term goal that are expected to increase across years.

Precision: Precision relates to the accuracy of the score and can be represented by the standard deviation. The standard deviation will have different names (and formulas) depending on the data (e.g., individual, group, assessment, etc.).

Reliability: Reliability is generally thought of as internal consistency, but for school accountability, reliability refers to the extent to which the variation in true means is captured by the variation in observed means.

School-level targets: The percentage of EL students meeting their individual progress targets in a given year and, at a state’s discretion, the additional measure of ELP attainment.

Student-level progress targets: How much progress individual EL students are expected to make from one year to the next.

References

- August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Castellano, K. E., & Ho, A. D. (2013). *A practitioner's guide to growth models*. Washington, DC: Council of Chief State School Officers. Retrieved from https://scholar.harvard.edu/files/andrewho/files/a_practitioners_guide_to_growth_models.pdf
- Conger, D. (2009). Testing, time limits, and English learners: Does age of school entry affect how quickly students can learn English? *Social Science Research*, 38(2), 383–396. Retrieved from <http://eric.ed.gov/?ID=ED501661>
- Cook, H. G., Boals, T., Wilmes, C., & Santos, M. (2008). *Issues in the development of annual measurable achievement objectives for WIDA consortium states* (WCER Working Paper No. 2008-2). Madison, WI: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved from http://www.wcer.wisc.edu/publications/workingPapers/Working_Paper_No_2008_02.pdf
- Cook, G., Linquanti, R., Chinen, M., & Jung, H. (2012). *National evaluation of Title III implementation supplemental report: Exploring approaches to setting English language proficiency performance criteria and monitoring English learner progress*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development.
- García-Vázquez, E., Vázquez, I. A., López, I. C., & Ward, W. (1997). Language proficiency and academic success: Relationships between proficiency in two languages and achievement among Mexican American students. *Bilingual Research Journal*, 21(4), 395–408.
- Goldschmidt, P. (2020). *Handbook for developing and monitoring the English language proficiency indicator and English learner progress*. Washington, DC: Council of Chief State School Officers. Retrieved from <https://ccsso.org/resource-library/handbook-developing-and-monitoring-english-language-proficiency-indicator-and>
- Goldschmidt, P., & Hakuta, K. (2017, September). *Incorporating English learner progress into state accountability systems*. Washington, DC: Council of Chief State School Officers. Retrieved from <https://ccsso.org/resource-library/incorporating-english-learner-progress-state-accountability-systems>
- Grissom, J. B. (2004). Reclassification of English learners. *Education Policy Analysis Archives*, 12(36), 1–36.
- Haas, E., Huang, M., Tran, L., & Yu, A. (2016a). *The achievement progress of English learner students in Nevada* (REL 2016–154). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2015098.pdf

- Haas, E., Huang, M., Tran, L., & Yu, A. (2016b). *The achievement progress of English learner students in Utah* (REL 2016–155). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2016155.pdf
- Haas, E., Tran, L., Huang, M., & Yu, A. (2015). *The achievement progress of English learner students in Arizona* (REL 2015–098). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2015098.pdf
- Hakuta, K. Y., Butler, G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (Policy Report No. 2000-1). Los Angeles, CA: The University of California Linguistic Minority Research Institute. Retrieved from <http://escholarship.org/uc/item/13w7m06g#page-1>
- Hakuta, K., & Pompa D. (2017). *Including English learners in your state Title I accountability plan*. Washington, DC: Council of Chief State School Officers.
- Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal*, 48(5), 1187–1225.
- Kieffer, M. J., & Parker, C. E. (2016). *Patterns of English learner student reclassification in New York City public schools* (REL 2017–200). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2017200.pdf
- Kim, Y. K., Curby, T. W., & Winsler, A. (2014). Child, family, and school characteristics related to English proficiency development among low-income, dual language learners. *Developmental Psychology*, 50(12), 2600–2613.
- Lindholm-Leary, K., & Borsato, G. (2006). Academic achievement. In F. Genesee, K. Lindholm-Leary, W. S. Saunders, & D. Christian (Eds.), *Educating English language learners: A synthesis of research evidence* (pp. 176–222). New York, NY: Cambridge University Press.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16.
- Linguanti, R., & George, C. (2007). Establishing and utilizing an NCLB Title III accountability system: California’s approach and findings to date. In J. Abedi (Ed.), *English language proficiency assessment in the nation: Current status and future practice* (pp. 105–118). Davis, CA: University of California.

- Liu, K. K., Thurlow, M. L., Press, A. M., & Lickteig, O. (2018). *A review of the literature on measuring English language proficiency progress of English learners with disabilities and English learners* (NCEO Report 408). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Lyons, S., & Dadey, N. (2017). *Considering English language proficiency within systems of accountability under the Every Student Succeeds Act*. Dover, NH: Center for Assessment and Chicago, IL: Latino Policy Forum. Retrieved from http://www.latinopolicyforum.org/publications/reports/document/Considerations-for-ELP-indicator-in-ESSA_030817.pdf
- McHugh, M., Pompa, D., & Lyons, S. (2017, March). *Strategic opportunities for including English learners in ESSA state accountability plans*. Webinar presented by the Migration Policy Institute. Retrieved from <http://www.migrationpolicy.org/multimedia/strategic-opportunities-including-english-learners-essa-state-accountability-plans>
- New York City Department of Education. (2016). *English language learner policy and reference guide*. New York, NY: Author. Retrieved from <http://steinhardt.nyu.edu/scmsAdmin/media/users/rp1018/ELLPolicyandReferenceGuideFINALOctober2016.pdf>
- Slama, R., Haynes, E., Sacks, L., Lee, D. H., & August, D. (2015). *Massachusetts English language learners' profiles and progress: A report for the Massachusetts Department of Elementary and Secondary Education*. Washington, DC: American Institutes for Research. Retrieved from <https://eric.ed.gov/?id=ED578925>
- Slama, R., Molefe, A., Gerdeman, D., Herrera, A., Brodziak de los Reyes, I., August, D., & Cavazos, L. (2017). *Time to proficiency for Hispanic English learner students in Texas* (REL 2017–280). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2018280.pdf
- Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion programs on student achievement evidence from lottery data. *American Educational Research Journal*, 54(1), 282S–306S.
- Umansky, I. M., & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion and English immersion classrooms. *American Educational Research Journal*, 51(5), 879–912. Retrieved from <http://aer.sagepub.com/content/51/5/879>
- Valentino, R. A., & Reardon, S. F. (2015). Effectiveness of four instructional programs designed to serve English learners. *Educational Evaluation and Policy Analysis*, 37(4), 612–637.

Appendix A. ELP Community of Practice Goals, Activities, and Membership

Objectives

State Support Network facilitators set the following objectives for CoP participants:

- Understand ESEA provisions related to incorporating ELs into state accountability systems.
- Identify and prioritize areas of support needed to develop and implement their ESEA consolidated state plans as they relate to including ELs into state accountability systems.
- Understand extant research and best practices related to EL accountability systems and how they can be used to inform the design and implementation of ESEA consolidated state plans.
- Learn about best methods to obtain and incorporate stakeholder feedback related to ESEA consolidated state plans and implementation.
- Refine current or proposed state plans and their implementation as they relate to ELs.

ELP CoP Membership

This section lists the ELP CoP members (Table A1) and key characteristics of the states they represented during the CoP (Table A2).

Table A1. ELP CoP Members⁵²

Name	Job Title
Arizona Department of Education	
Kate Wright	Deputy Associate Superintendent, Title III Director
Peter Laing	Deputy Associate Superintendent, Title I Director
Tammy McKeown	Director, ELL Accountability and Support
Carol Lippert	Associate Superintendent
Kelly Koenig	Associate Superintendent
Arkansas Department of Education	
Tricia Kerr	ESOL Program Director
Miguel Hernandez	Title III Program Specialist
Stacy Smith	Assistant Commissioner, Division of Learning Services
Jayne Green	Title I Director
Bobby Lester	Federal Programs Director
Louis Ferren	State Systems Administrator
Denise Airola	Director, Office of Innovation for Education
Alan Lytle	EL Assessment Specialist

⁵² This table reflects the name of members submitted by states upon joining the CoP.

Name	Job Title
Hope Allen	Director of Student Assessment
Johnny Key	Commissioner of Education
Tina Smith	Special Projects Director
Kelli Langan	Research Associate, Office of Innovation in Education
Deena Rorie	Research Assistant, Office of Innovation in Education
California Department of Education	
Cindy Kazanis	Director, Analysis, Measurement and Accountability and Reporting Division
Veronica Aguila	Director, English Learner Support Division
Karmina Barrales	Education Programs Consultant
Marcela Rodriguez	Division Consultant, English Learner Support
Jonathan Isler	Research Scientist
Delaware Department of Education	
Maria Paxson	Title III Field Agent
Gregory Fulkerson	Director, Language Acquisition
Ted Jarrell	Education Associate, Title I, Part A, Student Support Team
Kentucky Department of Education	
Tara Rodriguez	Branch Manager, Office of Continuous Improvement and Support
Gary Martin	Title III, EL Program Consultant, Office of Teaching and Learning
Chris Williams	Program Consultant, Office of Assessment and Accountability
Louisiana Department of Education	
Melanie Mayeux	Educational Program Consultant, NCLB Services, Office of Student Programs
Jill Zimmerman	Director of Accountability Policy
Beverly Diaz	Director of English Language and World Languages
Maine Department of Education	
April Perkins	Acting Title III Coordinator
Chelsea Fortin	Director, Federal Programs and Title I
Daniel Weeks	Title V Coordinator & Title I Data Specialist
Janette Kirk	Deputy Director of Learning Systems
Massachusetts Department of Elementary and Secondary Education	
Erica Gonzales	Center for District Support
Paul Aguiar	Director, Office of English Language Acquisition and Academic Achievement
Melanie Manares	Title III Coordinator
Michigan Department of Education	
Chris Janzer	Assistant Director, Accountability
Shereen Tabrizi	Manager, Special Populations Unit (EL)
Mike Radke	Director, Office of Field Services (State Title I Director)
Jen Paul	Assessment Specialist, Office of Student Assessment

Name	Job Title
Chad Bailey	Accountability Specialist, Office of Data Services
Michael Mekhayel	State and Federal Accountability Specialist, Office of Data Services
Kelly Alvarez	EL Consultant
New Mexico Public Education Department	
Mayra Valtierrez	Director, Language and Culture Bureau
Kirsi Laine	English Learner Specialist Language and Culture Bureau
Dr. Yun Yao	Statistician Assessment and Accountability Division
Lisa Chandler	Director, Assessment and Accountability Division
Ryan Tolman	Statistician Supervisor
Judy Harmon	Indian Education Staff
North Carolina Department of Public Instruction	
Christie Lynch Ebert	K-12 Standards, Curriculum and Instruction Division Director
Tammy Howard	Accountability Division Director
Ivanna M T Anderson (lead)	ESL/Title III Staff
Xatli Stox	ESL/Title III Staff
Marshall Foster	ESL/Title III Staff
Wendy Wooten	Education Testing/Accountability Consultant
Talbot Troy	Federal Program Administrator
Puerto Rico Department of Education	
Leonardo Torres	Under Secretary of Academic Affairs
Amaralis Caro	Spanish Learners as Second Language and Immigrant Program Director
Euclides Valentin	Director of Statistics and Data Quality
Maria C. Martinez (Mary Carmen)	Title III State Coordinator in Federal Affairs
Yanira Raices	Auxiliary Secretary of Planning
Ada E. Hernández Guadalupe	Assessment and Accountability Director
Damaris L. Matos Carrillo	Staff
Daileen Carrion Muñoz	Staff
Yanin M. Dieppa Perea	Staff
Maria C. Christian	Staff
Rhode Island Department of Education	
Phyllis Lynch	Director, Office of Instruction, Assessment, and Curriculum
Ana Karantonis	Assessment Specialist
Rachel Peterson	Research Specialist
David Sienko	Director, Office of Student, Community, and Academic Supports
Illinois State Board of Education	
Gil Sanchez	Division Supervisor, DELL
Seng Naolhu	Principal Consultant
Beth Robinson	Principal Consultant
Sonia Serrano	Principal Consultant

Name	Job Title
Jenna Chapman	Principal Consultant
Barry Pedersen	Principal Consultant
Rocio Seda	Principal Consultant
Samuel Aguirre	Principal Consultant
Indiana Department of Education	
Valerie Beard	Assistant Director of English Learners and Migrant Education Programs
Olga Tuchman	Title III Specialist
Adam Pitt	English Learner Specialist
Nicole Leach	English Learner and Dual Language Immersion Specialist
New Jersey Department of Education	
Lori Ramella	Bilingual/ESL Coordinator
Kenneth Bond	Bilingual/ESL Coordinator
Karen Campbell	Title I Director
Jacquelyn Léon	Bilingual/ESL Specialist
Clare Barrett	Accountability Coordinator
Jesse Young	ESSA Federal Liaison
Oregon Department of Education	
Kim Miller	Title III
Jon Wiens	Accountability
Josh Rew	Reporting Title I
Taffy Carlisle	Title III/Equity
Melinda Bessner	Title I-A

Table A2 provides a snapshot of the characteristics of ELP CoP member states with respect to the number and percentage of ELs that they serve. Collectively, the ELP CoP served at least 2,136,200 ELs, or 44.0 percent of the nation’s EL population (based on 2016–17 national statistics). The majority of ELP CoP states served student bodies composed of more than 5.9 percent ELs statewide, with four states serving more than 8.5 percent of ELs.

Table A2. Number and Percentage of ELs Served by ELP CoP Member States (from lowest to highest total number of ELs)

State*	Number (%) of ELs Statewide
Maine (ME)	5,295 (2.9%)
Delaware (DE)	10,831 (7.9%)
Rhode Island (RI)	11,057 (7.8%)
Kentucky (KY)	21,897 (3.2%)
Louisiana (LA)	22,843 (3.2%)
Arkansas (AR)	41,482 (8.4%)
New Mexico (NM)	44,899 (13.4%)
Indiana (IN)	47,676 (4.5%)

State*	Number (%) of ELs Statewide
Oregon (OR)	56,598 (9.8%)
Arizona (AZ)	70,546 (6.3%)
New Jersey (NJ)	70,941 (5.0%)
Massachusetts (MA)	86,658 (9.0%)
North Carolina (NC)	92,388 (6.0%)
Michigan (MI)	94,921 (6.2%)
Illinois (IL)	197,496 (9.8%)
California (CA)	1,260,672 (20.2%)
United States	4,858,377 (9.6%)

Note: Puerto Rico Department of Education (PRDE) is a member of the ELP CoP. However, PRDE does not report the number of ELs because the primary language of instruction on the island is Spanish. Accordingly, almost all students in Puerto Rico are ELs. PR uses Title III to serve limited Spanish proficient students.

Source: National Center for Education Statistics (2016–17)

Appendix B. ELP CoP Subject Matter Expert Profiles

Pete Goldschmidt (California State University, Northridge)

Pete Goldschmidt, Ph.D., is a Professor in the Department of Educational Psychology and Counseling at California State University, Northridge. His teaching interests include social science research methods, program evaluation, and applied statistics for evaluation. His research interests include developing and applying longitudinal methods to programs and school reform evaluation and studies of international education quality, particularly in developing countries. Dr. Goldschmidt participated in several international educational evaluations, including Peru, India, China, and East Timor. Previously, he served as the Assistant Secretary for Assessment and Accountability in the New Mexico Public Education Department. In 2017 and 2018, he worked with the Council of Chief State School Officers to provide technical support to approximately 30 state educational agencies in developing their ELP indicators. Dr. Goldschmidt also serves on several state and district technical advisory committees. Most recently he published the *Handbook for Developing and Monitoring the English Language Proficiency Indicator and English Learner Progress*.

H. Gary Cook (Wisconsin Center for Education Research)

H. Gary Cook directs research for the WIDA consortium and is a research scientist attached to the Wisconsin Center for Education Research. He has served in educational leadership and research positions in the private industry, in an urban public school district, in a state department of education, and at the university level. He is an experienced federal peer reviewer for ESEA and serves on several state and national technical advisory committees. His recent research and publication interests have focused on the relationship between ELP and content assessments, standards alignment, policy issues associated with Title III accountability, and applying growth modeling techniques to address key education questions for ELs.

Mark Hansen (University of California, Los Angeles/CRESST/ELPA21)

Mark Hansen is an Assistant Professor in Residence in the UCLA Graduate School of Education and Research Scientist at CRESST. His work focuses on the use of latent variable models, particularly item response theory and diagnostic classification models, to support the design of educational, psychological, and health-related assessments.

Robert Linqanti (WestEd)

Robert Linqanti is a Senior Researcher at WestEd, specializing in assessment, evaluation, and accountability policies, practices, and systems for ELs. He conducts research, provides technical assistance, and advises education leaders and policymakers on these topics at the national, state, and local levels. In his current work, he supports state and local implementation of new content and ELP standards and assessments and advises states on more common policies and processes for defining ELs. He recently coauthored a U.S. Department of Education study to define and measure EL linguistic and academic progress; several guidance publications for the Council of Chief State School Officers (CCSSO) to help states move toward a more common definition of EL; and a policy primer for CCSSO on supporting formative assessment for deeper learning.

Delia Pompa (Migration Policy Institute)

Delia Pompa is Senior Fellow for Education Policy at Migration Policy Institute's (MPI's) National Center on Immigrant Integration Policy, where her work focuses on research and policy analysis related to improving educational services for immigrant students and ELs. Her previous experience as Executive Director for Bilingual and Migrant Education in the Houston Independent School District and as a bilingual classroom teacher and instructor to prospective teachers at the graduate level has anchored her work. Her influence has been felt widely throughout the field of education policy; she has served as an advisor or board member for many key institutions, including the Chapter I Commission and the Stanford Working Group, the Civil Rights and Business Coalition on the Reauthorization of the Elementary and Secondary Education Act, the American Youth Policy Forum, EdReports, the National PTA, International Baccalaureate, and the Joan Ganz Cooney Center.

Martha Thurlow (National Center on Educational Outcomes)

Martha Thurlow is the Director of the National Center on Educational Outcomes. She addresses the implications of contemporary U.S. policy and practice for students with disabilities and ELs with disabilities. Dr. Thurlow has a broad range of experience and expertise on policy and practice issues that affect students with disabilities and those who are ELs. During the past decade, she has been the principal investigator on more than 20 federal and state projects focused on students with special needs in state and national policies and in large-scale accountability assessments, including graduation exams. Dr. Thurlow has given particular emphasis to how to obtain valid, reliable, and comparable measures of the knowledge and skills of these students while ensuring that assessments are truly measuring their knowledge and skills, not measuring their disabilities or limited English when these are not the focus of the assessment.

Appendix C. Additional Select ESEA⁵³ Statutory and Regulatory Requirements Related to English Learners

Former English Learners

Results of the student’s reading/language arts and mathematics assessments for students previously identified as ELs may be included in the EL subgroup for up to four years after the student ceases to be identified as an EL (i.e., exit) when calculating performance on indicators that use results from those assessments (ESEA section 1111(b)(3)(B)).

Recently Arrived English Learners⁵⁴

- Recently arrived ELs (RA ELs) have been enrolled for less than 12 months in a school in one of 50 states in the United States or District of Columbia (ESEA section 1111(b)(3)(A)).
- A State may include RA ELs in the same manner it includes all ELs, or the State may adopt two additional options for including RA ELs in state assessments and Title I school accountability under ESEA section 1111(b)(3)(A).
 - **Option 1:**
 - Year 1 of RA ELs’ enrollment in U.S. schools:
 - » May exclude RA ELs from one administration of Title I reading/language arts assessment. Assess RA ELs on all other required assessments (i.e., ELP and mathematics).
 - » Exclude RA ELs from Title I accountability for reading/language arts and mathematics.
 - » Exclude RA ELs’ ELP assessment results from Title I accountability.
 - Year 2 of RA ELs’ enrollment in U.S. schools:
 - » RA ELs take all Title I required assessments.
 - » Include RA ELs’ performance on Title I reading/language arts and mathematics for accountability purposes.
 - » Include RA ELs’ performance on the ELP assessment for accountability purposes.
 - **Option 2:**
 - Year 1 of RA ELs’ enrollment in U.S. schools:
 - » Assess and report RA ELs’ Title I (including reading/language arts and mathematics) assessment results.
 - » Exclude RA ELs from Title I accountability for reading/language arts and mathematics.

⁵³ Reference to the ESEA, unless otherwise noted, is the *Elementary and Secondary Education Act of 1965*, as amended by the *Every Student Succeeds Act*.

⁵⁴ Select requirements presented during the April 13, 2017, English Language Proficiency Community of Practice meeting.

- Year 2 of RA ELs’ enrollment in U.S. schools:
 - » Assess and report RA ELs’ Title I (including reading/language arts and mathematics) assessment results.
 - » Include growth from Year 1 to Year 2 on reading/language arts and mathematics for Title I accountability.
- Year 3 of RA ELs’ enrollment in U.S. schools:
 - » Assess and report RA ELs’ Title I (including reading/language arts and mathematics) assessment results.
 - » Include proficiency results on reading/language arts and mathematics for accountability purposes (and in succeeding years).
- An SEA may also choose to use both options (i.e., apply option 1 to a subset of RA ELs that meet certain criteria—such as initial English language proficiency level—and option 2 to a different subset of RA ELs).

Source: ESEA section 1111(b)(3)(A)

ELs With Disabilities⁵⁵

1. State assessments must be valid and reliable for their intended purposes (ESEA 1111(b)(2)(B)(iii); 34 CFR 200.2(b)(4)(i)).
2. Assessments must provide for the participation of all students (ESEA 1111(b)(2)(B)(vii); 34 CFR 200.2(b)(2)(i)).
3. States must provide for an ELP assessment of all ELs in their state (ESEA 1111(b)(2)(G), 34 CFR 200.6(h)(1)).

Accommodations

1. SEAs must ensure that ELs with disabilities are appropriately included in state ELP and content area assessments (34 CFR 200.6(a),(f),(h)(4)).
2. The state must ensure that a student who requires and uses appropriate accommodations on content and ELP assessments is not denied any benefit afforded to a student who does not need such an accommodation (34 CFR 200.6(b)(3) and (f)(2)).
3. ELP assessments must include appropriate accommodations for all ELs with disabilities (34 CFR 200.6(h)(4)(i)).

⁵⁵ Select requirements presented during the October 27, 2017, English Language Proficiency Community of Practice meeting.

4. In cases where the IEP Team under IDEA, the section 504 team, or the individual or group designated by the LEA to make those decisions under Title II of the Americans with Disabilities Act (ADA) determines, on an individualized basis, that an EL's disability precludes their participation in one or more domains of the ELP assessment such that there are no appropriate accommodations for the affected domain(s) (e.g., a non-verbal EL who because of an identified disability cannot take the speaking portion of the assessment), a State must assess the student's English language proficiency based on the remaining domains in which it is possible to assess the student (34 CFR 200.6(h)(4)(ii)).
5. Teachers of ELs are among those who should receive necessary training regarding administering assessments, including ELP assessments, that covers how to administer appropriate accommodations and alternate assessments (34 CFR 200.6(b)(2)(ii)).

Alternate Assessments

1. A state may adopt alternate academic achievement standards for assessing the performance of students with the most significant cognitive disabilities on content assessments provided those standards are aligned with the challenging state academic content standards (ESEA 1111(b)(1)(E)(i)(I)).
2. For assessments in mathematics, reading/language arts, and science, the number of students with the most significant cognitive disabilities whose performance may be assessed with an alternate assessment aligned with alternate academic achievement standards (AA-AAAS) is limited to 1 percent of the total number of students in the State who are assessed in that subject (ESEA 1111(b)(2)(D)(i)(I)).

Alternate ELP Assessments

1. States must provide an alternate ELP assessment for EL students with the most significant cognitive disabilities who cannot participate in the general ELP assessment even with appropriate accommodations, as determined by the IEP team (34 CFR 200.6(h)(5)).
2. The alternate ELP assessment is not subject to the 1 percent cap in ESEA 1111(b)(2)(D); however, it is expected that the vast majority of ELs with disabilities will be able to take the general ELP assessment with or without appropriate accommodations.
3. The Department is conducting peer review for the ELP and alternate ELP assessments (ESEA section 1111(a)(4); 34 CFR 200.2(d)).

Selected Title III Reporting Requirements

- States must report the number and percentage of ELs achieving English language proficiency (ESEA 1111(h)(1)(C)(v)).
- States must report the number and percentage of ELs making progress toward ELP in the aggregate and by ELs with disabilities (ESEA 3121(a)(2); 3122(a)).

- States must report the number and percentage of ELs in the programs and activities funded under Title III attaining English language proficiency based on state ELP standards established under section 1111(b)(1)(G) by the end of each school year, as determined by the state’s ELP assessment under section 1111(b)(2)(G) (ESEA section 3121(a)(3); 3122(a)).
- States must report the number and percentage of ELs who exit the language instruction educational programs based on their attainment of English language proficiency. (ESEA section 3121(a)(4); 3122(a)).
- States must report the number and percentage of ELs meeting the state academic standards for each of the four years after they no longer receive Title III services, in the aggregate and by ELs with disabilities (ESEA section 3121(a)(5); 3122(a)).
- States must report the number and percentage of ELs who have not attained English language proficiency within five years of initial classification as an EL and first enrollment in the LEA (ESEA section 3121(a)(6); 3122(a)).