WIDA™

Technical Report

No. TR-2019-3
March 2019

Less Than Four Domains: Creating an Overall Composite Score for English Learners with Individualized Education Plans

Tameka Porter, H. Gary Cook, Narek Sahakyan
Wisconsin Center for Education Research, University of Wisconsin–Madison

Wisconsin Center for Education Research
SCHOOL OF EDUCATION
UNIVERSITY OF WISCONSIN–MADISON

Less Than Four Domains: Creating an Overall Composite Score for English Learners with Individualized Education Plans

Tameka Porter, H. Gary Cook, Narek Sahakyan, Wisconsin Center for Education Research

The Every Student Succeeds Act (ESSA) of 2015 defines an English Learner (EL) as a student aged 3 years through 21 years enrolled or preparing to enroll in elementary or secondary school, whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual (i) the ability to meet the challenging State academic standards; (ii) the ability to successfully achieve in classrooms where the language of instruction is English; or (iii) the opportunity to participate fully in society (ESSA Title 8, Section 8101(20))[1].

Federal requirements under ESSA stipulate that states, and districts within states, establish long-term and interim goals demonstrating that ELs make annual gains in their progress toward English language proficiency (ESSA Title 1, Part A, Section 1111(b)(1)(F)). ESSA also requires states to identify an indicator for ELs that measures increases in the percentage of students making progress in achieving English language proficiency (ESSA Title I, Part A, Section 1111(C)(4)(A)(ii)). Based on the current interpretations of federal law regarding long-term and interim English language growth, states must select an indicator and do the following:

1. Determine a scoring metric (EL indicator) and growth criterion to be used to measure growth,
2. Set the starting point for growth targets,
3. Set the ending point for growth targets,
4. Determine the amount of time needed for schools to get from the starting to ending targets, and
5. Establish an annual rate of growth.

In order to establish an annual rate of growth, ELs should be annually assessed and scored in the four recognized domains of Listening, Speaking, Reading, and Writing. Almost all states use some form of weighted overall composite score to monitor accountability, attainment, and progress (Linquanti et al., 2016).

ELs with a 504 plan (29 U.S.C. § 701)[2], which addresses federal protections for participating in programs and activities for individuals with disabilities, or an Individualized Education Plan (IEP) that ensures specialized instruction for students with disabilities attending elementary or secondary school may have a disability that prevents them from taking a domain test. In such

[1] The definition of an English learner also includes those who were (i) not born in the United states or whose native language is a language other than English; (ii)(I) is a Native American or Alaska Native, or a native resident of the outlying areas; and (II) who comes from an environment where a language other than English has had a significant impact on the individual's level of English language proficiency; or (iii) is migratory, whose native language is a language other than English, and who comes from an environment where a language other than English is dominant.

[2] A 504 plan refers to Section 504 of the Rehabilitation Act of 1973, which prevents exclusion from programs and activities that receive Federal funding for individuals with disabilities.

instances, an overall composite score cannot be calculated, which may impact the ELs' reclassification status as well as the schools' accountability decisions for these students.

The following analyses illustrate models and procedures that can be used to impute a missing domain score for calculating an overall composite score. These methods apply only to ELs with IEPs or 504 plans requiring that they not be assessed in one or more domains. Analyses herein use the WIDA ACCESS for ELLs (ACCESS) English Language Proficiency (ELP) assessment for the 2017-2018 academic school year. The Appendices include tables and statistical programming code used to compute the different models. Appendix A features a sample of the most commonly applied domain-specific English Language Proficiency exit criteria scale scores throughout the WIDA Consortium. Appendix B illustrates the mean scale score values for each grade and domain. Appendix C provides the programming code to generate domain-specific z-score values to calculate overall composite scores for ELs who are missing one or more domain scores.

Cook (2013) conducted an extensive and complex version of a missing domain score imputation analysis through examining how to establish English proficiency using a logistic regression model approach and shared findings with the WIDA Technical Advisory Committee (TAC). The prior computation requires sophisticated statistical modeling and several assumptions related to the relationship between performance on state content assessments and English language proficiency (ELP), composite score weighting, and performance levels of ELs with disabilities relative to their non-disabled EL counterparts. The current research provides comparable yet simpler models.

The purpose of these models is to highlight how a state can apply them for monitoring, achievement, and accountability determinations. The output shown in this study is not intended to endorse or critique any of the models or their corresponding calculations. There are several limitations to this study. First, states vary considerably in their EL and ELs with disabilities demographic composition. Second, the sample sizes for both ELs and ELs with IEPs or 504 plans differ from state to state. As such, the results in this study are not generalizable across states.

The intent of this study is to provide technical guidance to states on how various models for calculating overall composite scores for ELs with IEPs who are missing one or more domains can offer opportunities for accountability. The analyses and results shown aim to provide clear and concise frameworks to approaching a complex issue. Thus, the focus of the study is more on the methods that can be applied to shape accountability criteria rather than on the results from each model. While the models presented in this study use data from the ACCESS assessment to generate alternate overall composite scores for ELs with directives in their IEPs or 504 plans allowing the student to not take one or more domain tests, the following procedures and calculations can also be applied to compute alternate overall composite scores and proficiency levels for ELs missing one or more domains on the Alternate ACCESS for ELLs assessment (Alt ACCESS), the test designed by the WIDA Consortium to monitor academic language development for ELs with significant cognitive disabilities (SWD).

Less Than Four Domain Imputation Models

Across the WIDA Consortium, a little more than two million ELs took ACCESS in the 2017-2018 academic year. About 250,000[3] or 12% of the identified ELs had IEPs or 504 plans. Of the ELs with IEPs or 504 plans, about 5,000, or 2% of those students had one or more missing domain scores. It is uncertain how many of the ELs with IEPs or 504 plans have specific statements in these documents that exempt them from being assessed in one or more domains.

Three conditions are considered when selecting a model for imputing a scale score for ELs missing one or more domains. First, the criteria should be easy to apply. Second, the model should have the potential to calculate the EL indicator. Third, the approach can identify exit criteria that relate to the overall composite score.

Four approaches for creating an overall composite score for ELs with IEPs who are missing one or more domains are examined in this report. Model 1 explores selecting the minimum domain score to exit English language programs to the missing domain. Model 2 imputes the average observed domain score to the missing domain. Model 3 assigns the average z-score of the each observed domain score and applies the z-score value to obtain the missing domain scale score. The fourth approach explores conducting a standard setting with local stakeholders to create a state-specific option for generating an overall composite score for ELs with IEPs.

Model 1: Assign Exit Score

Figure 1 presents a flowchart of the steps to calculating an overall domain score for ELs with IEPs or 504 plans who have one or more missing domain scores using Model 1: assigning the lowest exit score for the missing domain score(s). Model 1 includes four stages. Model 1 begins by identifying the state's exiting or reclassifying overall composite proficiency level and assigns the lowest exiting proficiency level scale score for the missing domain(s). After the lowest exiting scale score is applied to the missing domain, the overall composite scale score can then be calculated using the overall composite weighting formula[4]. Then the overall composite scale score can be matched with its corresponding composite proficiency level.

[3] This figure is an estimate of the number of ELs classified as having a 504 or IEP plan. Two WIDA states do not report the number of ELs who have 504 or IEP plans. Several WIDA states underreport the number of ELs with IEPs or 504 plans and/or do not disclose that some ELs receive accommodations when taking ACCESS. In addition, some WIDA states list that ELs in their state have a primary disability but do not indicate an IEP status.

[4] Both ACCESS and Alt ACCESS apply the following weighting to calculate the overall composite score: 15% Listening, 35% Reading, 15% Speaking, and 35% Writing.

Figure 1: Model 1 Flowchart – Assign Exit Score

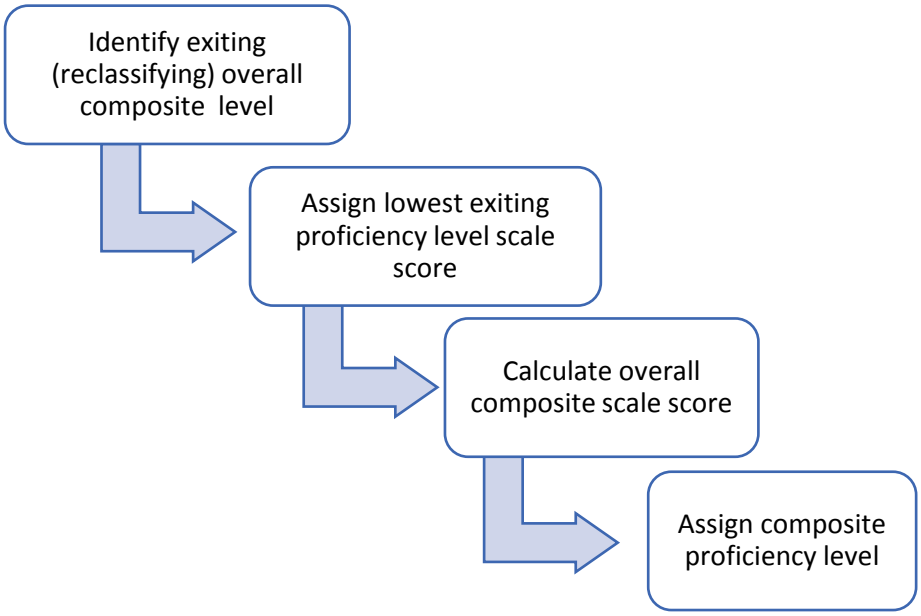


Table 1 provides a sample of the lowest exiting domain score by grade for states with exit criteria of 4.5 and 5.0, the most widely used exit scores across the WIDA Consortium[5].

Table 1: Lowest Exit Scale Score Values by Exit Criteria, Domain, and Grade

Grade	Exit Criteria = 4.5				Exit Criteria = 5.0			
	Listening Scale Score	Speaking Scale Score	Reading Scale Score	Writing Scale Score	Listening Scale Score	Speaking Scale Score	Reading Scale Score	Writing Scale Score
0	282	325	284	339	286	349	289	367
1	297	336	310	360	303	361	315	382
2	322	348	332	365	330	374	337	388
3	340	359	347	370	349	386	352	394
4	353	370	359	376	363	397	364	401
5	365	379	369	382	375	407	373	407
6	374	389	378	387	385	417	382	413
7	382	397	385	393	394	425	389	419
8	390	405	391	398	402	433	395	424
9	397	413	397	404	409	440	401	430
10	402	420	402	414	415	446	406	436
11	407	426	406	416	420	451	410	441
12	412	431	411	423	426	455	414	447

[5] The Appendix includes the lowest exit domain scores for the four most common state exit criteria.

Here is an example of Model 1. Student A in Grade 9 has a statement in her IEP plan that allows her to not participate in the Writing domain test on ACCESS. She attends an elementary school in a state with an exit criteria of an overall composite proficiency level (CPL) score of 4.5. Student A received a 423 Listening Scale Score, a 432 Speaking Scale Score, and a 379 Reading Scale Score. To compute the overall composite score for Student A, the lowest exiting scale score for an overall exit composite of 4.5 on the Writing domain test for Grade 9 is assigned as her Writing score (404). This calculation can be expressed as:

(Listening Scale Score)0.15 + (Speaking Scale Score)0.15 + (Reading Scale Score)0.35 + (Writing Scale Score)0.35

The alternate overall composite scale score for Student A can now be computed:

(423)0.15 + (432)0.15 + (379)0.35 + (404)0.35 = 402.30

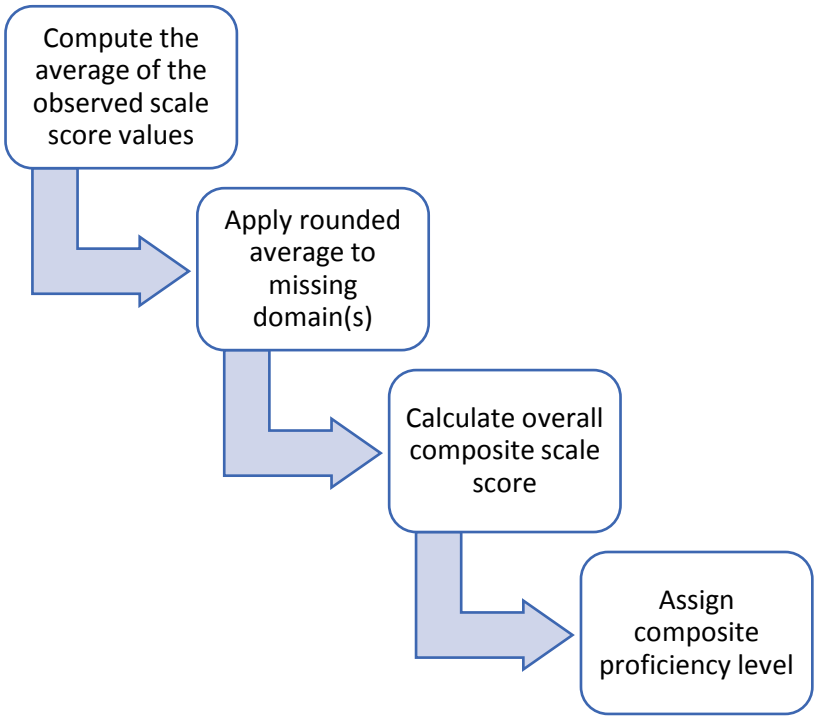
Scale score values are presented as whole numbers. Thus Student A now has as alternate overall composite score of 402. Using a scale score to proficiency level lookup table, this alternate overall composite score corresponds to an overall composite proficiency level of 4.5, which meets the state exit criteria.

The underlying assumption for using this model is that the missing domain score is the same as the lowest exiting or reclassifying score. The benefit to applying Model 1 is that it is the easiest of the four approaches to apply. In addition to the method being conceptually easy to understand, it also provides an assumed gain as the missing score. Challenges to this method include producing higher than observed scores for students and the need for a scale score to proficiency lookup table to apply this approach. Moreover, the exiting or reclassifying score will vary by state.

Model 2: Assign the Average Observed Domain Score(s)

Figure 2 presents a flowchart of the steps to calculating an overall domain score for ELs with IEPs or 504 plans who have one or more missing domain scores using Model 2: assigning the average observed domain score for the missing domain(s).

Figure 2: Model 2 Flowchart – Assign the Average Observed Domain Score



Model 2 begins by computing the average of the observed domain scale scores and applying the rounded average as the missing domain(s). After the rounded average scale score is applied to the missing domain(s), the overall composite scale score can be calculated and then assigned to the associated composite proficiency level.

For example, Student B in Grade 6 has a statement in his IEP plan that allows him to not participate in the Speaking domain test on ACCESS. Student B attends a middle school in a state with an exit criteria of an overall CPL of 4.8. Student B received a 404 Listening Scale Score, a 319 Reading Scale Score, and a 316 Writing Scale Score. To compute the overall composite score for Student B we need to calculate the mean of observed scale scores to impute the missing Speaking score, which is 346 $[(404 + 319 + 316) \div 3]$.

This value is then applied to calculate the overall composite scale score and CPL for Student B:

(404)0.15 + (346)0.15 + (319)0.35 + (316)0.35 = 334.75

Since scale score values are presented as whole numbers, Student B now has an alternate overall composite score of 335. Using a scale score to proficiency level lookup table, this alternate overall composite score corresponds to an overall composite proficiency level of 3.2, which does not meet the state exit criteria.

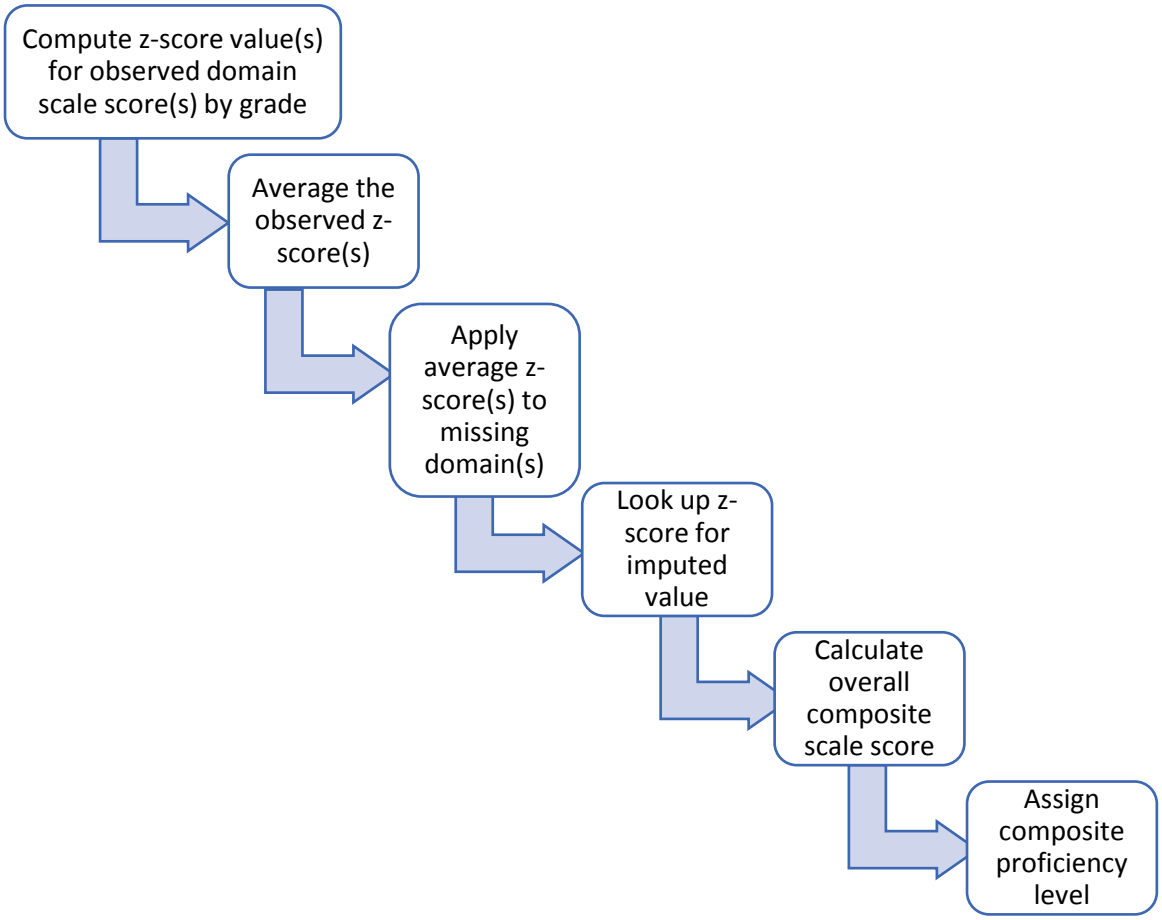
The underlying assumption of Model 2 is that the correlation between the domain scores is high. The benefit to this model is that it is easy to apply and is not conceptually difficult to

understand. A state can calculate this model with its own data. The drawbacks to applying this model are that scale scores may skew too high or low if the correlations between the domain scores are not high. Model 2 also requires a scale score to proficiency level lookup table in order to apply this method.

Model 3: Assign the Average Observed Z-Score Value

Figure 3 presents a flowchart of the steps to calculating an overall domain score for ELs with IEPs or 504 plans who have one or more missing domain scores using Model 3: assigning the average observed z-score value to the missing domain(s). Model 3 includes six stages. This method begins by establishing the z-score value for all observed domain scale scores at each grade. Then the average of observed domain z-scores is used for the missing domain. The scale score value associated with the average observed z-score is used for the missing domain. Once the missing scale score is determined, the overall composite scale score can be generated and then assigned its corresponding composite proficiency level.

Figure 3: Model 3 Flowchart – Assign the Average Observed Z-Score Value



Two lookup tables are needed to apply this method: a domain-specific z-score to domain scale score lookup table and an overall composite scale score to proficiency level lookup table. The underlying assumption of Model 3 is that the distribution of the observed domain scores is normal. As with Model 2, this approach assumes a high positive correlation between domains. The benefit to applying this model is that domain score distributions are often observed to be normal. The drawbacks to using this model include its complexity and the need to have multiple lookup tables to compute values. Moreover, given the constraints for applying this model to specific subgroups (e.g., ELs with IEPs or 504 plans with statements requiring them to be exempt from one or more domain test), the number of students eligible for this analysis may be small.

Model 3 was applied to a WIDA state, herein State A, to compute alternate composite scores for ELs with 504 plans or IEPs. In State A, just over 50,000 ELs took ACCESS in the 2017-2018 academic year. About 15% of ELs had IEPs or 504 plans. Of the ELs with IEPs or 504 plans, 142, or less than 2% had one or more missing domain score. Of the 142 ELs with IEPs or 504 plans who were missing one or more domain scores, about 40, or 0.5% had IEPs or 504 plans that included statements requiring the EL to be exempt from specific domain tests. The number of ELs that meet the qualifications for having a missing domain score imputed and an overall composite score calculated using this model for this state is small. State A has an exit criteria of 4.5.

To apply an example of this model for students in State A, observed z-score values are first calculated for each observed domain scale score value. (Note that domain z-scores were calculated across the WIDA Consortium, not just in State A.) The domain z-scores have a mean of 0 and have a standard deviation of 1.0. Here is an example from State A: Student C in Grade 2 attending an elementary school has an observed Listening scale score of 404, which corresponds to a z-score of 1.35, meaning that the observed Listening scale score for Student C is over one standard deviation above the mean Listening scale score in Grade 2. (Before proceeding, it is important to note that while domain-specific z-scores will have an exact scale score match, the averages will not. As such, states should determine their rounding procedure when selecting a scale score value.) Continuing this example, Student C has an observed Speaking scale score of 263 (z-score = -0.94), observed Writing scale score of 295 (z-score = -0.10), and a missing Reading domain score. The average z-score value for Student C is 0.10, which is calculated as $\{[1.35 + (-0.94) + (-0.10)]/3 = 0.10\}$. This z-score value corresponds to a Reading domain scale score of 321.

Now, the alternate overall composite scale score for Student C can now be computed:

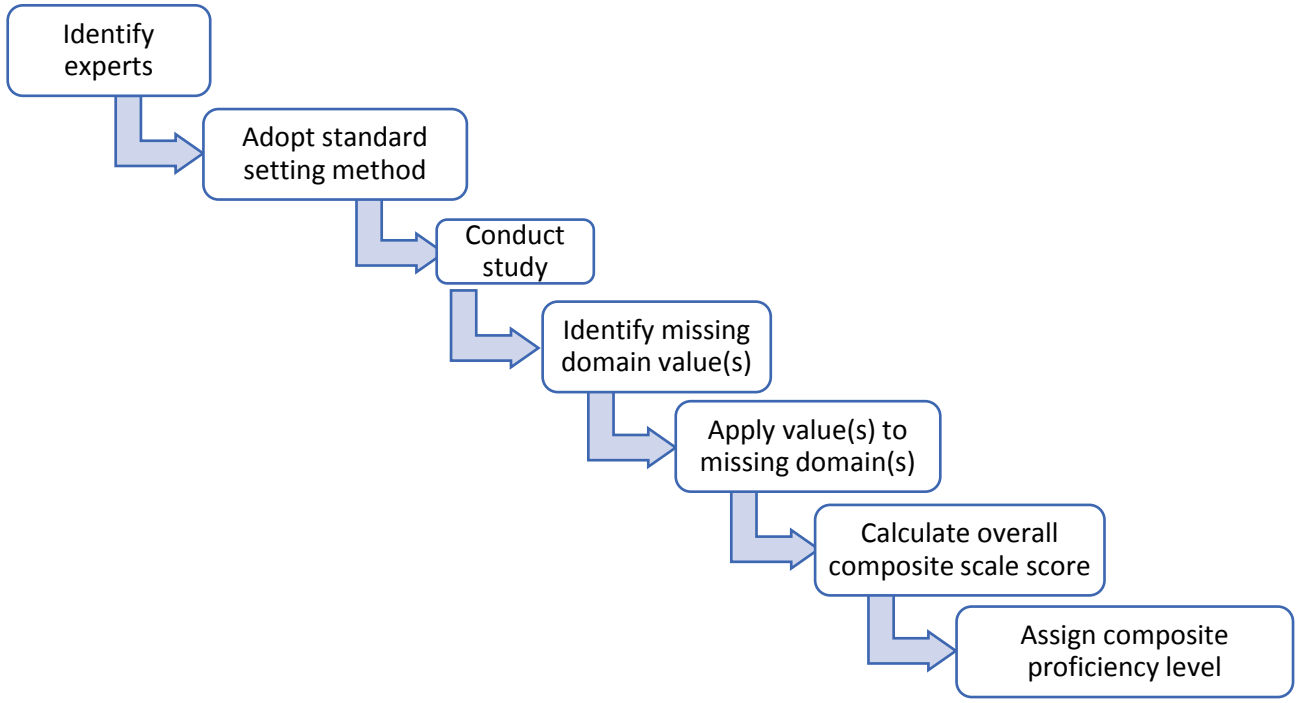
(404)0.15 + (263)0.15 + (321)0.35 + (295)0.35 = 315.65

Since scale score values are presented as whole numbers, Student C now has as alternate overall composite score of 316. Using a scale score to proficiency level lookup table, this alternate overall composite score corresponds to an overall composite proficiency level of 3.6, which does not meet the state exit criteria.

Model 4: Conduct a Standard Setting

Figure 4 presents a flowchart of the steps to calculating an overall domain score for ELs with IEPs or 504 plans who have one or more missing domain scores using Model 4: conducting a standard setting to create a procedure for imputing missing domain scores.

Figure 4: Model 4 Flowchart – Conduct a Standard Setting



Model 4 involves conducting a standard setting activity that assigns missing domain scores by convening experts in the field to create a state or district-specific procedure. The state or district may identify experts who specialize in education for ELs and students with disabilities. These experts may then adopt a standard setting method, examine the WIDA proficiency levels, and conduct one or more studies to identify missing and apply missing domain values. After the studies are conducted, and multiple rounds of review are concluded, the overall composite scale scores for the missing domains are calculated and assigned to their associated composite proficiency level. The underlying assumption for Model 4 is that in-state experts are better decision-makers than score distributions because they are aware of the needs of the student population. This model is beneficial because it uses local, state-specific experts and local decision-making criteria. In addition, this method is often accepted by peer reviewers. The challenges to applying this model are that it is time-consuming to convene panels of experts that have standard setting expertise. Moreover, this activity may be expensive for the state.

Summary

The models presented here are intended to support state monitoring, achievement, and accountability determinations. The benefits and drawbacks to applying the four models illustrated in this report are summarized in Table 4.

Table 4: Imputation Model Assumptions, Benefits, and Drawbacks

Model	Assumptions	Benefits	Drawbacks
Designate exit score(s)	Missing domain score should be assumed as met	Easiest model to apply. Easy concept to understand. Provides assumed missing score benefit.	Likely to provide an inflated higher score.
Apply average observed domain score(s)	High correlation between domain scores.	Easy to apply. Easy concept to understand. State can easily calculate with its own data.	If correlations are not high, scores may be too high or low. Requires lookup table.
Assign average z-score	Normal observed domain score distribution and high correlation between domain scores	Frequently observed assumption about domain score distributions	Complex. Requires multiple look-up tables.
Conduct a standard setting	State experts are better decision-makers than score distribution calculations	Often accepted by peer reviewers. Uses local, state-specific experts and criteria.	Requires standard setting expertise. Labor intensive. Requires substantial financial investment.

Again, the purpose of the models presented in this report is not to endorse or critique any particular one but rather to provide ideas to states on how to calculate overall composite scores for ELs with IEPs who are missing one or more domains and to highlight how a state can apply one or more of these frameworks for decision-making purposes.

References

Cook, H.G. (2013). *Examining how to establish English proficiency using alternate composite scores in a WIDA state: A logistic regression approach*. Madison, WI: Wisconsin Center for Education Research – WIDA Research Report. Retrieved from https://wida.wisc.edu/sites/default/files/resource/Report-ExaminingHowToEstablishEnglishProficiencyUsingAltCompositeScores.pdf

Cook, H.G. (2014). *Examining relationships between Alternate ACCESS and state alternate assessments: Exploring notions of English language proficiency*. Madison, WI: Wisconsin Center for Education Research – WIDA Research Report. Retrieved from https://wida.wisc.edu/sites/default/files/resource/Report-ExaminingRelationshipsBetweenAlternateAccessandStateAlternateAssessmentsAL.pdf

Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).

Linquanti, R., Cook, H. G., Bailey, A. L., and MacDonald, R. (2016). *Moving toward a more common definition of English learner: Selected guidance for states and multi-state assessment consortia*. Washington, D.C.: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/sites/default/files/2017-10/MoreCommonDefinition-Final_0.pdf

Office of Special Education and Rehabilitative Services (2000). *A guide to the Individualized Education Program*. Washington, D.C.: U.S. Department of Education. Retrieved from https://www2.ed.gov/parents/needs/speced/iepguide/iepguide.pdf

Section 504, Rehabilitation Act of 1973, Pub. L. No. 93-112, 87 Stat. 394 (1973).

Suggested Citation

Porter, T., Cook, H. G., Sahakyan, N. (2019). *Less Than Four Domains: Creating an Overall Composite Score for English Learners with Individualized Education Plans* (WIDA Technical Report No. TR-2019-3). Madison, WI: WIDA at the Wisconsin Center for Education Research.

Appendix A
Sample of WIDA State Domain-Specific English Language Proficiency Exit Criteria Scale Scores

Grade	Domain															
	Listening				Speaking				Reading				Writing			
	4.0	4.5	4.8	5.0	4.0	4.5	4.8	5.0	4.0	4.5	4.8	5.0	4.0	4.5	4.8	5.0
K	278	282	285	286	301	325	340	349	279	284	287	289	311	339	356	367
1	291	397	301	303	311	336	351	361	304	310	313	315	337	360	373	382
2	314	322	327	330	322	348	364	374	326	332	335	337	341	365	379	388
3	331	340	346	349	332	359	376	386	342	347	350	352	346	370	385	394
4	343	353	359	363	342	370	386	397	354	359	362	364	351	376	391	401
5	354	365	371	375	350	379	396	407	364	369	372	373	356	382	397	407
6	363	374	381	385	360	389	406	417	373	378	381	382	361	387	403	413
7	370	382	390	394	369	397	414	425	380	385	388	389	367	393	409	419
8	377	390	397	402	377	405	422	433	386	391	394	395	372	398	414	424
9	383	397	404	409	385	413	429	440	392	397	400	401	378	404	420	430
10	389	402	410	415	393	420	436	446	397	402	405	406	385	414	426	436
11	394	407	415	420	400	426	441	451	402	406	409	410	391	416	431	441
12	398	412	421	426	407	431	446	455	407	411	413	414	398	423	438	447

Appendix B
Mean Scale Score Values by Domain and Grade

Grade	Mean Listening Scale Score	Mean Reading Scale Score	Mean Speaking Scale Score	Mean Writing Scale Score
0	269	188	374	205
1	325	293	259	260
2	332	317	269	299
3	353	334	279	318
4	409	354	313	336
5	412	355	313	342
6	387	344	322	324
7	395	351	322	328
8	400	359	324	334
9	387	373	307	353
10	392	380	313	361
11	394	384	318	366
12	394	385	320	368

Appendix C

Programming Code for Z-Score Analysis

```
*THE FOLLOWING CODE APPLIES DOMAIN-SPECIFIC Z-SCORE VALUES TO CALCULATE
COMPOSITE SCORES FOR ELS WITH LESS THAN FOUR DOMAIN SCORES;

* You must first run the programming call to acquire the State's dataset; The current dataset
looks at scale score values for the 2017-2018 school year.

*The sql statement creates the dataset for you to analyze ACCESS. The table created for your
dataset is designated as A. Please rename dataset to fit your state's needs;

********** Calculating z-scores;**********
proc sort data = A; by grade; run;
proc standard data = A mean=0 std=1 out=Alt_stand;
format _numeric_ 9.3;
by grade;
var scale_score_listening scale_score_reading scale_score_speaking scale_score_writing;
run;
data Alt_stand1; set Alt_stand
        (keep =
                drc_student_identifier
                grade
                scale_score_listening--pl_writing
                reported_mode
                iep_status
        );
rename scale_score_reading=zread
                scale_score_listening=zlist
                scale_score_speaking=zspeak
                scale_score_writing=zwrit;
run;
data alt_comp1; set A
        (keep =
                drc_student_identifier
                scale_score_reading
                scale_score_listening
                scale_score_speaking
                scale_score_writing
        );
run;
proc sort data = alt_comp1; by drc_student_identifier; run;
proc sort data = Alt_stand1; by drc_student_identifier; run;

********** Creating a combined z-score file;**********
```

```
data z_score;
merge Alt_stand1 alt_comp1;
by drc_student_identifier;
run;

* Creating domain z-score lookup tables;
data z_list; set z_score (keep = grade zlist scale_score_listening); run;
proc sort data = z_list nodup; by grade zlist; run;
data z_read; set z_score (keep = grade zread scale_score_reading); run;
proc sort data = z_read nodup; by grade zread; run;
data z_speak; set z_score (keep = grade zspeak scale_score_speaking); run;
proc sort data = z_speak nodup; by grade zspeak; run;
data z_writ; set z_score (keep = grade zwrit scale_score_writing); run;
proc sort data = z_writ nodup; by grade zwrit; run;

data z_score; set z_score;
if scale_score_listening = . then domain_miss = 1;
if scale_score_reading = . then domain_miss = 1;
if scale_score_speaking = . then domain_miss = 1;
if scale_score_writing = . then domain_miss = 1;
run;

********** Creating missing domain iep file;**********
data z_iep; set z_score;
avg_z = mean (of zread zlist zspeak zwrit);
where domain_miss = 1;
run;

data z_iep_l; set z_iep;
where zlist = .;
run;

data z_iep_s; set z_iep;
where zspeak = .;
run;

data z_iep_r; set z_iep;
where zread = .;
run;

data z_iep_w; set z_iep;
where zwrit = .;
run;
```

```
data z_iep_miss;
merge z_iep_l z_iep_s z_iep_r z_iep_w;
by drc_student_identifier;
run;

**********Creating z-score lookup tables by domain**********
data z_score_lookup;
merge z_read z_list z_speak z_writ;
run;
```