



theRPgroup

Research • Planning • Professional Development  
for California Community Colleges

SM

# Maximizing Gateway English Throughput for International Students in the California Community Colleges:

Understanding the Predictive Validity of  
Common ESL Tests

Craig Hayward

February 2023

[www.rpgroup.org](http://www.rpgroup.org)

# Table of Contents

<b>Introduction .....</b>	<b>3</b>
Reader’s Guide .....	4
<b>Key Findings .....</b>	<b>5</b>
<b>Methodology.....</b>	<b>6</b>
Sample Description .....	7
<b>Analysis.....</b>	<b>10</b>
Overall Attrition/Pathway Baseline Model .....	10
CELSA Placement Test .....	11
ACCUPLACER Placement Tests .....	14
Summary of Models .....	29
<b>Limitations .....</b>	<b>31</b>
<b>Conclusion.....</b>	<b>32</b>
<b>References .....</b>	<b>35</b>

# Introduction

The RP Group’s Multiple Measures Assessment Project (MMAP) produced the following technical report as part of a series on how California’s community colleges can ensure more English learners (ELs) successfully complete “gateway” English coursework – courses that satisfy the English writing requirements for completion of an associate’s degree as well as for transfer to a four-year college or university – in a timely way. This current research focuses on how specific placement strategies can help international students achieve this outcome.

In the past, assessment and placement practices in the California Community Colleges (CCC) assigned most students to remedial sequences with subsequent low rates of “throughput” – completion of transferable, college-level English and math within a specified timeframe – particularly for students from historically marginalized communities (Bahr, et al., 2017; Meija, Rodriguez, & Johnson, 2016). Passed in November 2017, California’s Assembly Bill 705 (AB 705) requires community colleges to maximize the probability that students will complete these gateway courses expeditiously. As a result, community colleges have reconsidered whether their practices, policies, and structures were conducive to maximizing throughput rates. They have increased access to gateway English courses by minimizing the number of prerequisites in their sequences and placing most students directly into transferable, college-level English using high school coursework, grades, and grade point average (GPA).

These policies have had an especially strong impact on English learners. The MMAP team found that ELs who are US high school graduates achieved much greater throughput when they were allowed to enter directly into transfer-level English composition rather than being placed into either developmental English or ESL sequences (Hayward, 2020). With the advent of AB 705, more ELs who are US high school graduates are directly entering gateway English. For example, in fall 2008 California community colleges directed 41% of students who previously took a high school English Language Development (ELD) class to the ESL pathway. Since then, the percentage has dropped. In fall 2017, only 33% of ELD students entered the ESL pathway, a decline of 8 percentage points in nine years; by fall 2020, just 21% of former ELD students were enrolling in ESL pathway courses, a decline of 12 percentage points in just three years.

However, not all ELs who enter the CCC are US high school graduates. International students represent a distinct subgroup of ELs who do not arrive to their community college with US high school data that the institution can use to facilitate placement. For these students, test scores from ESL assessments and [guided self-placement](#)<sup>1</sup> are the primary options.

The community college entrance experience for international students who are ELs is distinct from their US high school graduate peers in other ways. Most international students possess an F-1 Visa which requires them to be full-time students in an academic program that

---

<sup>1</sup> <https://equitableplacementtoolkit.cccco.edu/guided/>

culminates in a degree.<sup>2</sup> Since international students are overwhelmingly degree-seeking, they do not have the variation in education goals seen in other EL subgroups. Thus, it is safe to assume that (1) international students will need to fulfill a gateway English requirement (i.e., degree-applicable and transferable English composition) in order to complete their program of study, and (2) completion of this course is important to their academic objectives.

Additionally, international students must also demonstrate a baseline level of proficiency in English, typically via achieving or exceeding a threshold score on Test of English as a Foreign Language (TOEFL) or International English Language Testing System (IELTS) assessments, before they can even matriculate at a California community college and take an ESL assessment. These tests are part of an international student's application paperwork and are intended to provide evidence that they possess the English competence needed to successfully navigate a California community college, a requirement that is unique to international students.

## Reader's Guide

This technical report summarizes findings on how well information about international students' initial curricular pathway (English or ESL), level of placement in the pathway, and ESL placement test score(s) can predict their throughput in gateway English. We specifically focus on the value of different commonly used ESL placement tests to this process, including the Combined English Language Skills Assessment (CESLA) and multiple types of ACCUPLACER tests.

We begin with key findings from this research followed by a description of our research methodology. Then, we offer detailed analysis of the predictive validity of different ESL placement tests for students' throughput. We conclude with high-level recommendations based on these findings. We intend for this research to support the California Community Colleges Chancellor's Office (CCCCO) and other educators and policymakers involved in development education reform in their decision-making about ways to continue supporting the success and completion of English learners in the system.

---

<sup>2</sup> The M-1 Visa is an alternative to the F-1 that allows international students to pursue vocational studies, but it is rarely used.

## Key Findings

- We evaluated six commonly used ESL tests. None consistently added meaningfully to our ability to predict which international students would achieve throughput above and beyond predictions based on the pathway a student takes (ESL or English) and the level of the sequence at which the student begins.
  - The ACCUPLACER ESL Reading Skills Test stood out in that inclusion of its test scores consistently improved the accuracy of throughput prediction, even after controlling for starting pathway and course level. However, the gain in accuracy, though statistically significant, was small (66.6% vs. 66.2%).
- Only three tests improved the explained percentage of variation in English composition throughput by 1% or greater: the CESLA, the ACCUPLACER WritePlacer ESL Test, and the ACCUPLACER ESL Listening Test. However, deeper examination of these three tests revealed that each was problematic in its own way.
- For colleges that need to select a placement test, the ACCUPLACER ESL Reading Skills Test may be the best available option; while a relatively weak predictor of throughput, its results were at least consistently positive in terms of improving the prediction of which students would complete transfer-level English.

# Methodology

The dependent variable of interest in our analysis is throughput, generally defined as the percentage of a cohort of students that begin an English or math sequence at a given time who then go on to complete a designated course within a specified timeframe (see Key Terms). For our purposes, the course of interest is usually referred to as English Composition, Freshman Composition, Composition and Reading, or some variant thereof – herein referred to as gateway English). The curricula of all California community colleges include a version of this course. It transfers to the University of California and the California State Universities and satisfies the English writing requirements for completion of an associate’s degree.

The core concept of MMAP’s analytical approach rests on a foundational understanding: knowing a student’s pathway (either English or ESL) and their placement level in that pathway allows us to predict their likelihood of achieving throughput simply by modeling the impact of attrition across these multi-level sequences (Hayward & Willett, 2014; Hern & Snell, 2010). We call this approach “attrition/pathway modeling.”

At the same time, we are interested in understanding how adding students’ placement test scores to this model improves our ability to predict throughput. To be effective in its purpose, an assessment test used for placement must offer additional insight into which students are likely to achieve throughput beyond what the simple attrition/pathway model can provide. That is, a predictive model that includes both the (a) attrition modeling based on students’ English language arts pathway and placement level, and (b) placement test scores and which explains significantly more variance than the attrition model alone, would indicate that the assessment test is adding to our ability to predict student throughput. In turn, this model would be valid for the purpose of placing students in ways that will maximize their completion of gateway English in one year.

For this modeling approach to function, there must be sufficient variability in test scores and actual placements levels; they cannot be collinear. Fortunately for the purposes of this research, such variability exists in the CCC, where we find students of equivalent test scores experiencing placement into different levels of the ESL or English sequence. Multiple factors contribute to this context. The system’s size, combined with colleges’ decentralized, local governance systems and policies (often referred to as “local control”), means each institution implements their own specific placement practices with distinct cut score ranges. There are also differences in placement cut score decisions within an individual college over time. This

## Key Terms

**Throughput: The percentage of students who enrolled at a particular starting level successfully completed the transfer-level English course within a specified time frame (e.g., one year for the English pathway, three years for the ESL pathway).**

**Course success: completion of a course in one term, with a grade of A, B, C, or P.**

**One-term throughput: successful completion of a transfer-level English or math course within one term, which typically requires the student to start at the transfer level.**

variability in the development of test cut scores across the CCC results in students with the same or similar test scores being placed into different pathways and different course levels, creating the opportunity to evaluate and disentangle the degree to which throughput is related to the test scores and to what extent throughput is simply a function of the pathway and course level selection, per se.

In addition to their starting course level, the English language arts pathway ELs take is another critical element in predicting their throughput (Hayward, 2020; Llosa & Bunch, 2011). Research shows that for ELs who have graduated from US high schools, pathway selection is strongly associated with completion of the gateway English. In other words, those who start on the English curricular pathway experience markedly higher throughput rates compared to those who start in the ESL sequence (Hayward, 2020; Hayward et al., 2022).

At present, the selection of the English or ESL pathway is influenced by many factors. Some colleges have a practice of directing international students whose high school instruction was in English to the English assessment process, while all other international students are directed to the ESL assessment and pathway by default. Most colleges include a meeting with a counselor as an element of the placement process (i.e., a multiple measure) where, in addition to the counselor's judgment, student preference for a specific pathway is typically given considerable weight. The lack of consistency across colleges in the process and the relatively vague documentation around pathway selection has led some researchers to refer to this process as a "black box" (Llosa & Bunch, 2011).

The result of these highly variable assessment and placement practices is a natural experiment in which there is a pseudo-random sorting of students with the same or similar test scores into various starting course levels of the English and ESL sequences. The wide array of different placements of students with the same or similar test scores provides useful variability that allows us to isolate the impact of an EL's placement on their likelihood of completing gateway English at the end of the sequence. Like darts flung at a dartboard, similar students may land close to each other or spread far apart in their actual placement level.

In this context, a test can demonstrate true predictive validity by improving our understanding of how to maximize a student's likelihood of achieving throughput over and above knowledge of starting pathway and starting course level. Predictive validity is measured by including students' test scores as well as placement levels in a multivariate statistical model. When added to what we refer to as the "attrition/pathway baseline model," the test must be a significant predictor of throughput, and the model fit should be stronger when the test is included than when it is not. **Our current research presents an example of this approach to test validation, assessing the incremental contribution of commonly used ESL tests to the predictive validity of international students' throughput in gateway English.**

## Sample Description

We derived the data used in this research from a set of international students with active enrollments from 2010 to 2020, identified by a participating cohort of 12 California community

colleges.<sup>3</sup> We started by compiling an initial list of 18 colleges with the greatest number of international students between 2010 and 2020. Each college had an average level of at least 200 international students per year. Of these 18 colleges, 12 were able to provide test scores for the analytical data set. With support from the Educational Results Partnership, MMAP then matched the submitted student assessment data to students' course-taking and demographic information in the CCCC Management Information System (MIS) data system. Two test vendors provided additional data on CESLA and ACCUPLACER test scores, which we also matched to the analytical data file to backfill any missing assessment information where possible.

We used the demographic data to confirm the students' international status, ensuring the data file included only students with an international student visa (F-1 or M-1) who had graduated from a foreign high school ( $N = 28,327$ ). However, not all international students had usable ESL placement test scores. For example, some participating colleges uploaded scores from tests that are no longer available or viable (e.g., the Secondary Level Proficiency (SLEP) test, COMPASS); tests that were not aligned with our purpose (e.g., ESL Writing Challenge); or tests that were not clearly labeled. In a handful of other cases, the test scores provided indicated invalid or out of range results. We also excluded international students who only took an English placement test because there are no longer any English assessment tests that are approved for use by the CCCC.<sup>4</sup>

Finally, AB 705 and related educational code dictates timeframes for tracking throughput for gateway English: three years if entering the ESL pathway, or one year if entering the English pathway. For the purposes of this study, the analytic data set specifically includes data points from summer 2010 through fall 2020. For students entering the ESL pathway, fall 2017 was the last term in which they could join the pathway (to allow three years through fall 2020 to determine the outcome). For students entering the English pathway even after completing ESL placement testing, fall 2019 was the last term possible in which they could join to allow one year for outcome tracking.

**Ultimately, the final sample included 12,117 international students,** further described in Table 1. The sample has a slightly higher percentage of male students. As indicated above, all in the sample had an F1 or M1 student visa and were graduates of a foreign high school. Ethnicity demographics were missing for a large proportion of students (44%). Among those with ethnicity information, Asian students were by far the largest group (43.4% of the overall sample and 80.3% of those with a known ethnicity), while students who identified as White were the next most common group (9.9% of the overall sample and 18.2% of those with a known ethnicity). Among students with known educational goals, the overwhelming majority indicated that they intended to complete a degree and/or transfer to four-year institution. Most of the students started on the ESL pathway (77.2%) and 60.5% of the group achieved throughput.

---

<sup>3</sup> Participating colleges: Citrus, De Anza, El Camino, Foothill, Golden West, Irvine Valley, Mt. San Antonio, Orange Coast, Pasadena, Saddleback, Santa Barbara, and Santa Monica.

<sup>4</sup> AB 705 requires that English placement utilize either high school GPA or guided self-placement.

**Table 1.** Descriptive Statistics for Overall Sample

<b>Student Characteristics</b>	<b><i>M</i></b>	<b><i>SD</i></b>
Age	20.5	3.229
<b><i>Gender</i></b>	<b>%</b>	<b><i>N</i></b>
Female	47.0%	5,693
Male	52.9%	6,412
Unknown	0.1%	12
<b><i>International Student Type</i></b>	<b>%</b>	<b><i>N</i></b>
F1 or M1 Visa	100.0%	12,117
Foreign High School Graduate	100.0%	12,117
<b><i>Ethnicity</i></b>	<b>%</b>	<b><i>N</i></b>
Asian	43.4%	5,261
Black	0.5%	56
Hispanic and Latina/o/x	2.0%	239
White	9.9%	1,194
Other	0.3%	33
Unknown	44.0%	5,335
<b><i>Student Journey Type</i></b>	<b>%</b>	<b><i>N</i></b>
Degree or Transfer	66.5%	8,052
Short-Term CTE	0.5%	63
Adult Ed	0.7%	83
Other	0.7%	90
Unknown	31.6%	3,829
<b><i>Pathway</i></b>	<b>%</b>	<b><i>N</i></b>
ESL	77.2%	9,351
English	22.8%	2,766
Achieved Throughput	60.5%	7,332

As Table 2 below shows, the availability of scores for each type of ESL placement test (i.e., CESLA, different ACCUPLACER test types) varied.

**Table 2.** Count of Students and Colleges with Specific Test Scores

<b>Placement Test Types</b>	<b>Students</b>	<b>Colleges</b>
Any ESL Test Record	12,117	12
<b>ACCUPLACER Test Records</b>		
ESL Language Use	6,247	6
ESL Reading	11,791	7
ESL Listening	1,950	10
ESL Sentence Meaning	10,780	5
WritePlacer ESL Essay	1,858	3
<b>CESLA Test Records</b>		
Form 1 and/or 2	296	2

Therefore, our analysis of each test utilized a specific subset of the overall sample of international students. We include a table of descriptive statistics specific to each subsample analysis, along with the associated baseline attrition/pathway and test regression models, as found in the following section.

## Analysis

### Overall Attrition/Pathway Baseline Model

The first step of the analysis involved developing a baseline attrition/pathway model for each placement test. The performance of this model provides a useful benchmark against which to compare a model that includes placement test scores as additional predictors of throughput. We also evaluated each test instrument with a confusion matrix to determine the accuracy of the placements it would have generated.

We developed a simple attrition and pathway selection model (herein named overall baseline attrition/pathway model) to provide a point of comparison for the specific baseline attrition/pathway models for each test subsample. For the total sample of 12,117 international students with any ESL test score, this overall baseline attrition/pathway model explained 18.6% of the variability in throughput (see Table 3).

*Table 3.* Summary of Overall Baseline Attrition/Pathway Model

<b>-2 Log likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
14466.5039	.137	.186

As Table 4 shows, this research confirmed that both pathway selection and starting course level are significant predictors of throughput for this group of students.

**Table 4.** Variables Included in Overall Baseline Attrition/Pathway Model

<b>Variable</b>	<b>B</b>	<b>SE</b>	<b>Wald</b>	<b>Df</b>	<b>Sig.</b>	<b>Exp(B)</b>
Starting Pathway*	1.028	.070	217.814	1	<.001	2.796
Starting Course Level	.438	.019	545.541	1	<.001	1.549
Constant	.956	.042	514.162	1	<.001	2.601

\*English or ESL

The overall baseline attrition/pathway model accurately classified or predicted throughput status for 66.1% of students (see Table 5), an improvement of 5.6% percentage points relative to predictions based on the overall mean throughput level of 60.5% for the sample.

**Table 5.** Classification Table for Overall Baseline Attrition/Pathway Model

<b>Observed Throughput</b>	<b>Predicted Throughput</b>		<b>Accuracy</b>
	<b>Did Not Complete Gateway English</b>	<b>Completed Gateway English</b>	
Did Not Complete Gateway English	1,922	2,863	40.2%
Completed Gateway English	1,248	6,084	83.0%
Overall %			66.1%

## CELSA Placement Test

The CELSA data comprised two equivalent multiple choice test forms (Form 1 and Form 2). We combined scores from both test forms to maximize available power. Separate analyses of each individual form (not shown) produced similar results.

Students with CELSA scores predominantly came from two colleges. Their demographics were similar to the overall sample, though the percentage of Asian students was somewhat lower (36.8% vs. 43.4%) and the average rate of throughput attainment was much lower (47.6% vs. 60.5%) (see Table 6).

**Table 6.** Descriptive Statistics of CELSA Subsample

<b>Student Characteristics</b>	<b>M</b>	<b>SD</b>
Age	20.1	2.664
<b>Gender</b>	<b>%</b>	<b>N</b>
Female	36.1%	107
Male	63.2%	187
Unknown	0.7%	2
<b>International Student Type</b>	<b>%</b>	<b>N</b>
F1 or M1 Visa	100.0%	296
Foreign High School Graduate	100.0%	296
<b>Ethnicity</b>	<b>%</b>	<b>N</b>
Asian	36.8%	109
Black	1.0%	3
Hispanic and Latina/o/x	4.1%	12
White	11.8%	35
Other	0.3%	1
Unknown	45.9%	136
<b>Student Journey Type</b>	<b>%</b>	<b>N</b>
Degree or Transfer	69.3%	205
Short-Term CTE	1.7%	5
Adult Ed	0.7%	2
Other	1.4%	4
Unknown	27.0%	80
<b>Pathway</b>	<b>%</b>	<b>N</b>
ESL	76.7%	227
English	23.3%	69
Achieved Throughput	47.6%	141
Colleges Using Test	18.1%	2

The baseline attrition/pathway model for the CELSA subsample explained 10.8% of the variance in CELSA test takers’ throughput, lower than the 18.6% of the overall baseline attrition/pathway model (see Table 7).

**Table 7.** Summary of Baseline Attrition/Pathway Model for CELSA Subsample

<b>-2 Log likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
384.790	.081	.108

Unlike the overall baseline attrition/pathway model, entry into the English or ESL pathway was not predictive of throughput for the CELSA subsample. However, starting level was a significant predictor (see Table 8). At the same time, the sample size of students with CELSA test scores was much smaller than for the overall sample ( $n = 296$  vs.  $N = 12,117$ ), leading to lower statistical power to find statistically significant relationships in the CELSA model than in the overall model.

**Table 8.** Variables Included in Baseline Attrition/Pathway Model for CELSA Subsample

Variable	B	SE	Wald	Df	Sig.	Exp(B)
Starting Pathway*	-.570	.394	2.092	1	.148	.565
Starting Course Level	.662	.158	17.604	1	<.001	1.938
Constant	1.695	.474	12.797	1	<.001	5.448

\*English or ESL

The baseline attrition/pathway model for the CELSA subsample was able to accurately classify throughput status for approximately 64% of students (see Table 9).

**Table 9.** Classification Table for Baseline Attrition/Pathway Model for CELSA Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	119	36	76.8%
Completed Gateway English	71	70	49.6%
Overall %			63.9%

As Table 10 shows, the model with CELSA test scores explained 11.8% of the variability in throughput, an increase over the baseline attrition/pathway model’s explanatory power for this subsample (10.8%).

**Table 10.** Summary of CELSA Test Model

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
382.156	.089	.118

After we included the CELSA test scores in the model, starting level remained a statistically significant predictor of throughput. However, the CELSA test scores alone were not a statistically significant predictor of international student throughput (see Table 11).

**Table 11.** Variables Included in CELSA Test Model

Variable	B	SE	Wald	Df	Sig.	Exp(B)
Starting Pathway*	-.545	.397	1.886	1	.170	.580
Starting Course Level	.505	.183	7.629	1	.006	1.657
CELSA Score	.023	.014	2.644	1	.104	1.023
Constant	-.048	1.161	.002	1	.967	.953

\*English or ESL

Performance on the classification table was slightly improved for the CELSA test model relative to the baseline attrition/pathway model for the CELSA subgroup (64.5% vs. 63.9%) (see Table 12).

**Table 12.** Classification Table for CELSA Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	113	42	72.9%
Completed Gateway English	63	78	55.3%
Overall %			64.5%

## Summary: CELSA Placement Test

- **Two colleges provided cases of students with CELSA test scores.**
- **CELSA scores were not a significant predictor of throughput after controlling for starting pathway and starting level.**
- **Accuracy of predicted throughput (classification model) is slightly improved when CELSA scores are included (64.5% vs. 63.9%).**
- **Results were not consistently supportive of the utility of the test.**

## ACCUPLACER Placement Tests

Many of the colleges participating in the study used one or more types of ACCUPLACER ESL assessments, including the WritePlacer ESL Essay Test, ESL Reading Skills Test, ESL Language Use Test, ESL Sentence Meaning Test, and ESL Listening Test. We considered each type of ACCUPLACER test separately and then evaluated whether their joint use of the different tests could improve predictive accuracy.

### WritePlacer ESL Essay Test Model

The ACCUPLACER WritePlacer ESL Essay Test assesses the ability to compose an organized, coherent, and grammatically correct response to a topic prompt. Essays are scored by computer<sup>5</sup> based on the quality of the composition's organization, development, support, sentence structure, use of mechanical conventions, and focus.

Three colleges in the study cohort used the WritePlacer ESL Essay Test. Students at colleges completing this assessment achieved a higher average throughput than the overall sample (67.2% vs. 60.5%). Otherwise, the sample's demographics were similar to those of the overall sample (see Table 13).

---

<sup>5</sup> Student essays were scored using ACCUPLACER's IntelliMetric artificial intelligence system.  
Maximizing Gateway English Throughput for International Students in the California Community Colleges  
The RP Group | February 2023 | 14

**Table 13.** Descriptive Statistics of WritePlacer ESL Essay Test Subsample

<b>Student Characteristics</b>	<b>M</b>	<b>SD</b>
Age	19.8	2.907
<b>Gender</b>	<b>%</b>	<b>N</b>
Female	44.1%	819
Male	55.9%	1,038
Unknown	0.1%	1
<b>International Student Type</b>	<b>%</b>	<b>N</b>
F1 or M1 Visa	100.0%	1,858
Foreign High School Graduate	100.0%	1,858
<b>Ethnicity</b>	<b>%</b>	<b>N</b>
Asian	60.5%	1,125
Black	0.9%	17
Hispanic and Latina/o/x	1.7%	31
White	7.9%	146
Other	0.2%	3
Unknown	28.8%	536
<b>Student Journey Type</b>	<b>%</b>	<b>N</b>
Degree or Transfer	58.0%	1,077
Short-Term CTE	0.6%	12
Adult Ed	0.5%	10
Other	0.7%	3
Unknown	40.2%	746
<b>Pathway</b>	<b>%</b>	<b>N</b>
ESL	83.2%	1,545
English	16.8%	313
Achieved Throughput	67.2%	1,238
Colleges Using Test	27.3%	3

The baseline attrition/pathway model for the WritePlacer ESL Essay Test subsample explained 25% of the variance in throughput for students who took this assessment, higher than the 18.6% of the overall baseline attrition/pathway model (see Table 14).

**Table 14.** Summary of Baseline Attrition/Pathway Model for WritePlacer ESL Essay Test Subsample

<b>-2 Log Likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
1978.536	.182	.254

Both pathway and placement level were significant predictors of throughput in the baseline attrition/pathway model for the WritePlacer ESL Essay Test subsample (see Table 15).

**Table 15.** Variables Included in Baseline Attrition/Pathway Model for WritePlacer ESL Essay Test Subsample

Variable	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>Df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Starting Pathway*	.883	.210	17.646	1	<.001	2.417
Starting Course Level	.738	.050	220.976	1	<.001	2.091
Constant	1.616	.095	289.761	1	<.001	5.032

\*English or ESL

The baseline attrition/pathway model for the WritePlacer ESL Essay Test subsample accurately predicted the throughput status of 73.3% of students, higher than the overall baseline attrition/pathway model (66.1%) (see Table 16).

**Table 16.** Classification Table for Baseline Attrition/Pathway Model for WritePlacer ESL Essay Test Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	216	394	35.4%
Completed Gateway English	103	1,145	91.7%
Overall %			73.3%

As Table 17 shows, when including WritePlacer ESL Essay Test scores, the model explained 26.8% of the variability in throughput, an increase over the baseline attrition/pathway model’s explanatory power for this subsample (25.4%).

**Table 17.** Summary of WritePlacer ESL Essay Test Model

-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
1954.542 <sup>a</sup>	.193	.268

When including the WritePlacer ESL Essay Test scores in the model, all predictor variables – pathway, starting course level, and WritePlacer ESL Essay Test score – were significant predictors of throughput for this subsample (see Table 18).

**Table 18.** Variables Included in WritePlacer ESL Essay Test Model

Variable	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>Df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Starting Pathway*	.724	.213	11.585	1	.001	2.063
Starting Course Level	.565	.060	88.244	1	<.001	1.759
WritePlacer ESL Essay Score	.306	.063	23.586	1	<.001	1.358
Constant	.145	.315	.214	1	.644	1.157

\*English or ESL

Despite the improved explanatory power of the regression model, the confusion matrix for the WritePlacer ESL Essay Test model did not show a higher level of accuracy in predicting which students would and would not achieve throughput relative to the baseline attrition/pathway model (72.7% vs 73.3%).

**Table 19.** Classification Table for WritePlacer ESL Essay Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	288	322	47.2%
Completed Gateway English	185	1,063	85.2%
Overall %			72.7%

SUMMARY: WRITEPLACER ESL ESSAY TEST

- **Three colleges provided student cases with WritePlacer ESL Essay Test scores.**
- **WritePlacer ESL Essay Test scores were significant predictors of throughput even after controlling for starting pathway starting level.**
- **However, the accuracy of the predicted throughput (classification table) was lower when the WritePlacer ESL Essay test scores were included.**
- **Results for the WritePlacer ESL Essay Test were not consistent.**

ESL Reading Skills Test Model

The ACCUPLACER ESL Reading Skills Test presents students with short passages of 50 to 90 words to assess their ability to comprehend material by responding to both factual questions as well as to questions that require inference.

Seven colleges in the study cohort used this test. Students completing this assessment were similar to the overall sample, although they experienced a somewhat higher average throughput rate (63.6% vs. 60.5%, see Table 20).

**Table 20.** Descriptive Statistics of ESL Reading Skills Test Subsample

<b>Student Characteristics</b>	<b>M</b>	<b>SD</b>
Age	20.5	3.242
<b>Gender</b>	<b>%</b>	<b>n</b>
Female	47.3%	5,576
Male	52.6%	6,205
Unknown	0.1%	1
<b>International Student Type</b>	<b>%</b>	<b>n</b>
F1 or M1 Visa	100.0%	11,791
Foreign High School Graduate	100.0%	11,791
<b>Ethnicity</b>	<b>%</b>	<b>n</b>
Asian	43.5%	5,218
Black	0.4%	53
Hispanic and Latina/o/x	1.9%	226
White	9.8%	1,156
Other	0.3%	31
Unknown	44.1%	5,197
<b>Student Journey Type</b>	<b>%</b>	<b>n</b>
Degree or Transfer	66.4%	7,831
Short-Term CTE	0.5%	58
Adult Ed	0.7%	81
Other	0.7%	86
Unknown	31.7%	3,735
<b>Pathway</b>	<b>%</b>	<b>n</b>
ESL	77.2%	9,103
English	22.8%	2,688
Achieved Throughput	67.2%	7,171
Colleges Using Test	63.6%	7

The baseline attrition/pathway model for the ESL Reading Skills Test subsample explained 18.8% of the variance in throughput for students who took this assessment, nearly the same as the overall baseline attrition/pathway model (see Table 21).

**Table 21.** Summary of Baseline Attrition/Pathway Model for ESL Reading Skills Test Subsample

<b>-2 Log Likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
14026.575 <sup>a</sup>	.139	.188

Both pathway and starting course level were significant predictors of throughput in the baseline attrition/pathway model for the ESL Reading Skills Test subsample (see Table 22).

**Table 22.** Variables Included in Baseline Attrition/Pathway Model for ESL Reading Skills Test Subsample

Variable	B	SE	Wald	Df	Sig.	Exp(B)
Starting Pathway*	1.090	.072	229.211	1	<.001	2.974
Starting Course Level	.431	.019	510.649	1	<.001	1.539
Constant	.942	.042	491.208	1	<.001	2.565

\*English or ESL

The baseline attrition/pathway model for the ESL Reading Skills Test subsample accurately predicted the outcomes of 66.2% of students in this group, virtually the same as the overall baseline attrition/pathway model (66.1%) (see Table 23).

**Table 23.** Classification Table for Baseline Attrition/Pathway Model for ESL Reading Skills Test Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	1,821	2,799	39.4%
Completed Gateway English	1,185	5,986	83.5%
Overall %			66.2%

When including the ESL Reading Skills Test scores in the model, the percent of variability in throughput explained by the model improved relative to the baseline attrition/placement model for this subsample (19.6% vs. 18.8%, see Table 24).

**Table 24.** Summary of ESL Reading Skills Test Model

-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
13943.035a	.145	.196

In addition, the ESL Reading Skills Test score was a significant predictor of throughput when added to the regression model (see Table 25). Pathway and starting course level remained significant predictors as well.

**Table 25.** Variables Included in ESL Reading Skills Test Model

Variable	B	SE	Wald	df	Sig.	Exp(B)
Starting Pathway*	1.080	.072	224.267	1	<.001	2.946
Starting Course Level	.306	.023	170.288	1	<.001	1.357
ESL Reading Skills Test Score	.015	.002	82.396	1	<.001	1.015
Constant	-.723	.188	14.861	1	<.001	.485

\*English or ESL

The ESL Reading Skills Test model showed a slight improvement in overall accuracy of predicted throughput relative to the baseline attrition/pathway model for this subgroup (66.6% vs. 66.2%) (see Table 26).

**Table 26.** Classification Table for ESL Reading Skills Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	1,992	2,628	43.1%
Completed Gateway English	1,316	5,855	81.6%
Overall %			66.6%

SUMMARY: ESL READING SKILLS TEST

- **Seven colleges contributed cases of students with ESL Reading Skills Test scores.**
- **ESL Reading Skills Test Scores were significant predictors of throughput even after controlling for pathway and starting level.**
- **Accuracy of predicted throughput was slightly higher when ESL Reading Skills Test scores were included in the model (66.6% vs. 66.2%).**
- **The results were consistent in that both the regression analysis and the confusion matrix analysis indicated a small but positive improvement in prediction of throughput when the ESL Reading Skills Test scores were included.**

ESL Language Use Test Model

The ESL Language Use Test assesses use of subject-verb agreement, pronoun case structure, sentence structure, adverbs/adjectives, and subordination/coordination. Test-takers demonstrate their skills and knowledge by either selecting a word or phrase that best completes a sentence (i.e., fill in the blank) or by selecting a sentence that represents the best combination of two other sentences.

Six colleges in the study cohort used this test. Students completing this assessment had a demographic profile similar to that of the overall sample (see Table 27), though the percentage achieving throughput was somewhat higher than for the overall sample (64.0% vs. 60.5%).

**Table 27.** Descriptive Statistics of ESL Language Use Test Subsample

<b>Student Characteristics</b>	<b>M</b>	<b>SD</b>
Age	20.3	3.282
<b>Gender</b>	<b>%</b>	<b>n</b>
Female	46.4%	2,897
Male	53.5%	3,343
Unknown	0.1%	7
<b>International Student Type</b>	<b>%</b>	<b>n</b>
F1 or M1 Visa	100.0%	6,247
Foreign High School Graduate	100.0%	6,247
<b>Ethnicity</b>	<b>%</b>	<b>n</b>
Asian	45.2%	2,825
Black	0.4%	28
Hispanic and Latina/o/x	1.9%	121
White	10.4%	652
Other	0.2%	12
Unknown	41.8%	2,609
<b>Student Journey Type</b>	<b>%</b>	<b>n</b>
Degree or Transfer	59.7%	3,728
Short-Term CTE	0.4%	24
Adult Ed	0.2%	13
Other	1.0%	64
Unknown	38.7%	2,418
<b>Pathway</b>	<b>%</b>	<b>n</b>
ESL	73.9%	4,617
English	26.1%	1,630
Achieved Throughput	64.0%	3,396
Colleges Providing Test Data	54.5%	6

The baseline attrition/pathway model for the ESL Language Use Test subsample explained 19.4% of the variability in throughput for students who took this assessment, higher than the 18.6% of the overall baseline attrition/pathway model (see Table 28).

**Table 28.** Summary of Baseline Attrition/Pathway Model for ESL Language Use Test Subsample

<b>-2 Log Likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
7211.842a	.142	.194

Both pathway and starting course level were significant predictors of throughput in the baseline attrition/pathway model for the ESL Language Use Test subsample (see Table 29).

**Table 29.** Variables Included in Baseline Attrition/Pathway Model for ESL Language Use Test Subsample

Variable	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>Df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Starting Pathway*	1.090	.096	129.514	1	<.001	2.974
Starting Course Level	.427	.026	271.656	1	<.001	1.532
Constant	1.020	.058	308.850	1	<.001	2.774

\*English or ESL

The ACCUPLACER ESL Language Use baseline pathway/attrition model correctly predicted outcomes for 67.5% of students in this subsample, higher than the overall baseline attrition/pathway model (66.1%) (see Table 30).

**Table 30.** Classification Table for Baseline Attrition/Pathway Model for ESL Language Use Test Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	839	1,412	37.3%
Completed Gateway English	619	3,377	84.5%
Overall %			67.5%

When including the ESL Language Use Test scores in the model, the percent of variability in throughput explained by the model improved slightly relative to the baseline attrition/placement model for this subsample (19.6% vs. 19.4%) (see Table 31).

**Table 31.** Summary of ESL Language Use Test Model

-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
7203.938 <sup>a</sup>	.143	.196

When adding the ESL Language Use test scores to the regression model, all three predictors (pathway, starting course level, and ESL Language Use Test score) were significantly associated with completion of gateway English (see Table 32).

**Table 32.** Variables Included in ESL Language Use Test Model

Variable	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Starting Pathway*	1.095	.096	130.646	1	<.001	2.990
Starting Course Level	.375	.032	141.168	1	<.001	1.455
ESL Language Use Score	.006	.002	7.915	1	.005	1.006
Constant	.297	.263	1.268	1	.260	1.345

\*English or ESL

Despite the slight increase in the explanatory power of the regression model when including the Language Use Test scores, the confusion matrix analysis had the same level of accuracy at predicting outcomes as the baseline confusion matrix for this subgroup (67.5%) (see Table 33).

**Table 33.** Classification Table for ESL Language Use Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway	Completed Gateway English	
Did Not Complete Gateway English	860	1,391	38.2%
Completed Gateway Course	642	3,354	83.9%
Overall %			67.5%

SUMMARY: ESL LANGUAGE USE TEST MODEL

- Six colleges contributed cases of students with ESL Language Use Test scores.
- ESL Language Use Test scores were significant predictors of throughput even after controlling for starting pathway and starting level.
- Accuracy of predicted throughput was not improved by the inclusion of ESL Language Use Test scores in the model.
- Results were not consistently supportive of the utility of the ESL Language Use Test.

### ESL Sentence Meaning Test

The ACCUPLACER ESL Sentence Meaning Test assesses understanding of word meanings in the context of one or two sentences. Students must choose the best way to complete the sentence, either by changing it or keeping it the same. This assessment focuses on use of idioms, nouns, verbs, particles, phrasal verbs, prepositions, adverbs, adjectives, and connective sequences.

Five colleges in the study cohort provided ESL Sentence Meaning Test data. Students completing this assessment were similar to the overall sample, though the percentage of Asian students was somewhat lower (39.9% vs. 43.4%, see Table 34).

**Table 34.** Descriptive Statistics of ESL Sentence Meaning Test Subsample

<b>Student Characteristics</b>	<b>M</b>	<b>SD</b>
Age	20.5	3.165
<b>Gender</b>	<b>%</b>	<b>n</b>
Female	47.1%	5,079
Male	52.8%	5,691
Unknown	0.1%	10
<b>International Student Type</b>	<b>%</b>	<b>n</b>
F1 or M1 Visa	100.0%	10,780
Foreign High School Graduate	100.0%	10,780
<b>Ethnicity</b>	<b>%</b>	<b>n</b>
Asian	39.9%	4,304
Black	0.4%	38
Hispanic and Latina/o/x	1.7%	186
White	9.9%	1,071
Other	0.3%	29
Unknown	47.8%	5,152
<b>Student Journey Type</b>	<b>%</b>	<b>n</b>
Degree or Transfer	59.7%	7,145
Short-Term CTE	0.4%	46
Adult Ed	0.2%	21
Other	1.0%	76
Unknown	32.4%	3,492
<b>Pathway</b>	<b>%</b>	<b>n</b>
ESL	75.8%	8,175
English	24.2%	2,605
Achieved Throughput	61.5%	6,634
Colleges Providing Test Data	45.5%	5

The ESL Sentence Meaning Test baseline attrition/pathway model explained 18.9% of the variance in throughput for students who took this assessment, slightly higher than the 18.6% of the overall baseline attrition/pathway model (see Table 35).

**Table 35.** Summary of Baseline Attrition/Pathway Model for ESL Sentence Meaning Test Subsample

<b>-2 Log Likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
12746.640 <sup>a</sup>	.139	.189

Both pathway and starting course level were significant predictors of throughput in the baseline attrition/pathway model for the ESL Sentence Meaning Test subsample (see Table 36).

**Table 36.** Variables Included in Baseline Attrition/Pathway Model for ESL Sentence Meaning Test Subsample

Variable	B	SE	Wald	df	Sig.	Exp(B)
Starting Pathway*	1.082	.073	217.827	1	<.001	2.951
Starting Course Level	.420	.020	458.696	1	<.001	1.523
Constant	.932	.044	450.998	1	<.001	2.540

\*English or ESL

The ESL Sentence Meaning Test baseline pathway/attrition model correctly predicted outcomes for 66.7% of students in this subsample, higher than the overall baseline attrition/pathway model (66.1%) (see Table 37).

**Table 37.** Classification Table for Baseline Attrition/Pathway Model for ESL Sentence Meaning Test Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	1,659	2,487	40.0%
Completed Gateway English	1,108	5,526	83.3%
Overall %			66.7%

When including the ESL Sentence Meaning Test scores in the model, the percent of variability in throughput explained by the model remained the same as the baseline attrition/placement model for this subsample (18.9%) (see Table 38).

**Table 38.** Summary of ESL Sentence Meaning Test Model

-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
12746.586 <sup>a</sup>	.139	.189

ESL Sentence Meaning test scores were not a significant predictor of throughput for this subsample (see Table 39).

**Table 39.** Variables Included in ESL Sentence Skills Test Model

Variable	B	SE	Wald	df	Sig.	Exp(B)
Starting Pathway*	1.082	.073	217.528	1	<.001	2.949
Starting Course Level	.424	.026	274.853	1	<.001	1.528
ESL Sentence Meaning Scores	.000	.002	.053	1	.817	1.000
Constant	.977	.198	24.325	1	<.001	2.656

\*English or ESL

We found no difference in the overall accuracy of the ESL Sentence Meaning Test model relative to the baseline attrition/placement model for this subsample (both 66.7%) (see Table 40).

**Table 40.** Classification Table for ESL Sentence Meaning Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	1,659	2,487	40.0%
Completed Gateway English	1,108	5,526	83.3%
Overall %			66.7%

SUMMARY: ESL SENTENCE MEANING TEST

- **Five colleges contributed cases of students with ESL Sentence Meaning Test scores.**
- **ESL Sentence Meaning Test scores were not significant predictors of throughput after controlling for starting pathway and level.**
- **Accuracy of predicted throughput was not improved by the inclusion of ESL Sentence Meaning Test scores in the model.**
- **Results were not supportive of the utility of the ESL Sentence Meaning Test.**

ESL Listening Test

The ACCUPLACER ESL Listening Test assesses a student’s ability to understand one or more people speaking English. The recorded conversations cover a range of topics from academic topics such as science, math, and literature, to conversations that take place at home, at work, and while shopping or eating.

While 10 colleges in the study cohort provided ESL Listening Test data, just three colleges submitted most of the cases (95.9%). Students completing this assessment were younger on average than in the overall sample (19.5 years old vs. 20.5 years old) and more likely to be Asian (51.5% vs. 43.4%, see Table 41). The throughput rate for this sample of students is higher than the throughput rate for the overall sample (69.3% vs. 60.5%).

**Table 41.** Descriptive Statistics of ESL Listening Test Subsample

<b>Student Characteristics</b>	<b><i>M</i></b>	<b><i>SD</i></b>
Age	19.5	2.906
<b>Gender</b>	<b>%</b>	<b><i>n</i></b>
Female	41.9%	818
Male	58.0%	1,131
Unknown	0.1%	1
<b>International Student Type</b>	<b>%</b>	<b><i>n</i></b>
F1 or M1 Visa	100.0%	1,950
Foreign High School Graduate	100.0%	1,950
<b>Ethnicity</b>	<b>%</b>	<b><i>n</i></b>
Asian	51.5%	1,005
Black	0.9%	18
Hispanic and Latina/o/x	1.2%	23
White	6.9%	134
Other	0.1%	2
Unknown	39.4%	768
<b>Student Journey Type</b>	<b>%</b>	<b><i>n</i></b>
Degree or Transfer	54.7%	1,066
Short-Term CTE	0.4%	7
Adult Ed	0.3%	5
Other	0.2%	3
Unknown	44.6%	869
<b>Pathway</b>	<b>%</b>	<b><i>n</i></b>
ESL	87.7%	1,711
English	12.3%	239
Achieved Throughput	69.3%	1,352
Colleges Providing Test Data	90.9%	10

The ESL Listening Test baseline attrition/pathway model explained 18.4% of the variance in throughput for students who took this assessment, slightly lower than the 18.6% of the overall baseline attrition/pathway model (see Table 42).

**Table 42.** Summary of Baseline Attrition/Pathway Model for ESL Listening Test Subsample

<b>-2 Log Likelihood</b>	<b>Cox &amp; Snell R Square</b>	<b>Nagelkerke R Square</b>
2130.830a	.131	.184

Both pathway and starting course level were significant predictors of throughput in the baseline attrition/pathway model for the ESL Listening Test subsample (see Table 43).

**Table 43.** Variables Included in Baseline Attrition/Pathway Model for ESL Listening Test Subsample

Variable	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>Df</i>	<i>Sig.</i>	<i>Exp(B)</i>
Starting Pathway*	.867	.251	11.941	1	.001	2.379
Starting Course Level	.663	.051	170.297	1	<.001	1.942
Constant	1.805	.107	286.470	1	<.001	6.077

\*English or ESL

The ESL Listening Test baseline pathway/attrition model correctly predicted outcomes for 70.5% of students in this subsample, higher than the overall baseline attrition/pathway model (66.1%) (see Table 37).

**Table 44.** Classification Table for Baseline Attrition/Pathway Model for ESL Listening Test Subsample

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	181	417	30.3%
Completed Gateway English	158	1,194	88.3%
Overall %			70.5%

The model with ESL Listening Test scores included explained 19.6% of the variation in international student throughput, an increase over the baseline attrition/pathway model for this subsample (18.4%).

**Table 45.** Summary of ESL Listening Test Model

-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
2112.632 <sup>a</sup>	.139	.196

Interestingly, we found that ESL Listening Test scores were a significant negative predictor of throughput, meaning that lower scores were associated with higher throughput after controlling for the effect of pathway and starting course level.<sup>6</sup>

<sup>6</sup> This finding may be due to college-level influences on throughput, which are not included in the current models, though the zero order correlations of the listening test scores with other test scores are only in the 0.30 range and are low relative to the intercorrelations among the other test scores.

**Table 46.** Variables Included in ESL Listening Test Model

Variable	B	SE	Wald	Df	Sig.	Exp(B)
Starting Pathway*	.905	.251	12.941	1	<.001	2.471
Starting Course Level	.788	.060	171.379	1	<.001	2.199
ESL Listening Test Scores	-.019	.004	18.103	1	<.001	.982
Constant	3.624	.446	66.136	1	<.001	37.472

\*English or ESL

The accuracy of the ESL Listening Test model is higher than the baseline attrition/pathway model for this subsample (71.2% vs 70.5%). However, the Beta (*B*) for the ESL Listening Test is negative (i.e., higher listening scores are associated with decreased likelihood of attaining throughput after controlling for the influence of starting course level and pathway). One would expect higher ESL Listening Test scores to predict higher throughput; however, the negative relationship between higher test scores and lower predicted throughput calls into question the overall validity (e.g., construct, concept, criterion) of the ESL Listening Test. Given that the ESL Listening Test score is significant negative predictor of throughput, the test fails to show fundamental criterion validity of being positively associated with a higher likelihood of completing the gateway English composition class.

**Table 47.** Classification Table for ESL Listening Test Model

Observed Throughput	Predicted Throughput		Accuracy
	Did Not Complete Gateway English	Completed Gateway English	
Did Not Complete Gateway English	172	426	28.8%
Completed Gateway English	136	1,216	89.9%
Overall %			71.2%

#### SUMMARY: ESL LISTENING TEST

- **Ten colleges contributed cases of students with ESL Listening Test scores.**
- **Although predicted throughput accuracy rates were higher (71.2% vs. 70.5%) when ESL Listening Test scores were included in the model, ESL Listening Test scores were negatively associated with throughput after controlling for starting pathway and level. This failure to demonstrate a positive relationship between test scores and throughput indicates an issue with test validity.**
- **Results were not supportive of the utility of the ESL Listening Test.**

## Summary of Models

Table 48 provides a quick summary of the key statistical elements from each of the models presented in this research report. Each row summarizes and compares the baseline and the test models for each type of ESL placement test, with the two ‘Delta’ columns showing the change in the model statistics when test scores are included in addition to the baseline factors (i.e.,

starting course level and pathway). A positive Delta value indicates that when including the placement test scores, the model performed better than the baseline model; a Delta of zero indicates no difference in performance; and a negative Delta indicates that the model with the test scores performed worse than the baseline model.

**Table 48.** Summary of Strength and Accuracy of Models with and without Test Scores

Test	Baseline Model	Test Model	Model Delta	Baseline Accuracy	Test Model Accuracy	Accuracy Delta
CELSA	10.8%	11.8%	1.0%*	63.9%	64.5%	0.6%
ACCUPLACER WritePlacer ESL Test	25.4%	26.8%	1.4%**	73.3%	72.7%	-0.6%
ACCUPLACER ESL Reading Test	18.8%	19.6%	0.8%**	66.2%	66.6%	0.4%
ACCUPLACER ESL Language Use Test	19.4%	19.6%	0.2%**	67.5%	67.5%	0.0%
ACCUPLACER ESL Sentence Meaning Test	18.9%	18.9%	0.0%	66.7%	66.7%	0.0%
ACCUPLACER ESL Listening Test	18.4%	19.6%	1.2%**	70.5%	71.2%	0.7%

\*  $p < .05$  \*\*  $p < .01$

Only three test models improved the percentage of variation in throughput explained by 1% or greater: CELSA, ACCUPLACER WritePlacer ESL Test, and ACCUPLACER ESL Listening Test. Deeper examination of these three tests revealed that each was problematic in its own way. The CELSA Test model only improved accuracy of throughput prediction by 0.6%, and the test score itself was not a significant predictor in the test model. The accuracy of the ACCUPLACER WritePlacer ESL Test model declined relative to the baseline model, indicating a lack of robustness in the ACCUPLACER WritePlacer ESL Test model. Finally, the ACCUPLACER Listening Test scores were negatively associated with throughput once we controlled for starting course level and pathway, potentially indicating fundamental issues with the validity of the test for the purposes of predicting throughput.

A fourth test, the ACCUPLACER ESL Reading Skills Test, had perhaps the best overall performance. The inclusion of ACCUPLACER ESL Reading Skills scores was associated with both improved explanatory power and accuracy though the gain in model fit was less than 1% (0.8%) and the improvement in accuracy was also small (0.4% improvement).

## ACCUPLACER Omnibus Test

Colleges typically will use more than one ESL test as part of their placement system, so we developed additional models that included multiple ACCUPLACER tests. None of these omnibus models showed enhancements in predictive validity or accuracy greater than the results for single

tests. When we included more than one test score, only the strongest predictor would attain statistical significance,<sup>7</sup> and the overall variance explained by the model was not enhanced.

## Limitations

The sample sizes available for analyzing throughput for EL international students were modest. It is possible that the additional power of a larger sample would allow for the detection of significant relationships that were not apparent in this analysis, particularly for the CELSA test. Additionally, we performed this analysis using summary test scores delivered by colleges and the test vendors. These scores represent the students' final scores, but they do not allow for item-level analysis or weighting. It is not clear to what extent non-linear item weighting might enhance or even already be included in the summary scores. To the extent that non-linear item weighting is an important element of the test's performance, the models developed here will underperform relative to modeling that includes proper weighting of test items. Finally, the data set in this paper covered a 10-year period. It is possible that an instrumentation bias may be present in the data if there were changes to the testing instruments over this time, such that the scores reported at one time are not equivalent to scores at other times.

The available ACCUPLACER test score data are derived primarily from the ACCUPLACER Classic, which was discontinued in 2019 and replaced with a similar but updated set of tests. Given this change in the lineup of available tests, the analysis of ACCUPLACER test scores in this paper provides a general historical context for the performance of test instruments from ACCUPLACER. There are not sufficient data to provide direct evidence of the performance of their more updated tests, particularly since many California community colleges discontinued ESL placement testing in the 2019-2020 academic year with the implementation of AB 705, replacing it with guided self-placement and multiple measures-based placement.

An important element of the approach taken in this analysis is variation in the starting course levels experienced by students with the same or similar test scores. Due to variations across colleges, students did experience different placement systems, but some tests were not as widely used as others, which may have restricted some of the desirable variation in placements. Additionally, there may be other differences among the colleges that contributed data to this project that are associated with variations in throughput rates. We did not include potentially relevant college-level differences in the current analysis. For example, meeting with a counselor is an important part of determining both the pathway and starting course level of international students for most of the colleges in the sample. However, the impact of any counselor meetings is not systematically captured in the available assessment data.

Finally, this analysis focused specifically on international students. As a group, international students are quite distinct from other English learner populations. For instance, only international students have an English proficiency requirement that must be met to be

---

<sup>7</sup> Most likely due to multicollinearity of the test scores.

accepted at a California community college. International students are also typically required to be degree-seeking and are required to enroll full-time and stay on track toward degree completion as a condition of their Visa status in the US. These additional contextual factors are unique to international students and likely contribute to their relatively high overall rate of English throughput. Additionally, since international students typically will not have a US high school GPA that can be used for placement, as required by AB 705, they have only two placement options available to them: placement test or guided self-placement. Given these unique characteristics of international students, the result of this analysis will not necessarily generalize to other English learner populations.

## Conclusion

To determine whether a given test enhances the ability to predict which students would achieve throughput, we developed baseline attrition/pathway models that used just two predictors – starting course level and pathway (i.e., ESL or English) – and then compared these baseline models to augmented models that also included students’ test scores. Since students with the same test scores can start at different levels and can take different pathways (due to variations in practice across colleges), the test scores should, in theory, improve our ability to accurately predict which students would ultimately achieve throughput. Additionally, each model’s accuracy was assessed with a confusion matrix (or accuracy table) that compared predicted throughput to actual throughput. To be considered a good option, a test must improve both the overall predictive model and the demonstrated accuracy of predicted placements.

The ACCUPLACER WritePlacer ESL Essay Test, an AI-graded assessment provided the greatest improvement in model performance (1.4%). However, the accuracy of the WritePlacer ESL Essay model as measured by the confusion matrix analysis was lower than the baseline model’s accuracy, so this test could not be recommended based on the current data and results. Further research with a larger sample size could provide additional power and the ability to more definitively assess the performance attributes and potential benefits of using the WritePlacer ESL Essay Test as part of an assessment and placement system for EL international students.

Inclusion of ACCUPLACER ESL Listening Test scores improved model fit nearly as much as the WritePlacer ESL Essay Test (1.25%). Unfortunately, ACCUPLACER ESL Listening Test scores were negatively correlated with throughput, indicating an issue with the validity of the test for the purpose of predicting throughput.

The CELSA improved model fit relatively well (1%) and also demonstrated an improvement in accuracy of predictions (0.6%). However, the CELSA test scores were not significant predictors in the augmented model, even though the overall fit was improved when they were included. However, data limitations may be hampering the ability to draw stronger conclusions. For

example, CELSA data were available for only two of the participating colleges<sup>8</sup> resulting in a relatively small sample size. Further research with a larger sample would allow for a more definitive understanding of the CELSA test's performance attributes.

Another test with a positive gain in model fit (0.8%) that also improved the accuracy of predictions was the ACCUPLACER Reading Skills Test. This test was used by over half of the colleges in the sample ( $n = 7$ ). While accuracy of predicting throughput improved when the ACCUPLACER Reading Skills Test scores were included, the gain was relatively small (i.e., from 66.2% accuracy to 66.6% accuracy). It is not clear that the magnitude of the gain is large enough to be considered an operationally meaningful improvement in accuracy. Still, the ACCUPLACER Reading Skills Test did have the most consistently positive results of any of the six tests evaluated in this paper.

It is important to note that there is not a central repository for assessment and placement test data that CCCC researchers can access. Placement test scores and cut scores are not collected as part of the standard MIS data file submissions gathered by the system office in Sacramento. There is centralized data collection of assessment instrument names or codes, but the collection of raw scores was suspended in the early 1990s when the MIS data system was relatively new.<sup>9</sup> We recommend that the MIS data collection re-activate the gathering of raw placement test scores along with the date and name/code of the assessment test. While collection of raw assessment test scores is not a panacea for the challenges of conducting research into the efficacy of assessment and placement systems that utilize tests, it would go a long way toward increasing available samples sizes and making such research more tractable.<sup>10</sup> Additionally, it would be beneficial for the MIS assessment data collection to be expanded to include collection of proficiency test scores and test name, as these data elements were not systematically available at any of the participating community colleges with high enrollments of international students.

This research has demonstrated that the structural elements of assessment and placement, namely starting pathway (English or ESL) and starting course level, exert a powerful influence on the likelihood of EL international students successfully completing the gateway English composition course. Taking these structural factors into account, test scores added relatively little to our ability to predict which students will achieve throughput (i.e., complete gateway English). Based on the available data, none of the tests we evaluated can be strongly recommended for use in placing international students.

Given the relatively weak predictive utility of several commonly used ESL placement tests for predicting student throughput, future research should rigorously evaluate the use of guided self-placement models for international students. During the pandemic, most colleges shifted

---

<sup>8</sup> In Llosa & Bunch (2011), the CELSA was the most used ESL placement test in the CCC, with 47 colleges using it.

<sup>9</sup> See [https://webdata.cccco.edu/ded/sa/sa\\_all.pdf](https://webdata.cccco.edu/ded/sa/sa_all.pdf).

<sup>10</sup> With access to consistent and reliable test scores, the supplemental data gathering tasks could focus on how campuses implement multiple measures which, for international students at least, often includes the input of counselor and of the student themselves in response to the results from the test (Bracco et al., 2014).

from administering in-person placement tests to using guided self-placement processes that were amenable to remote administration. This relatively sudden shift in placement processes creates the conditions for a natural experiment in which it should be possible to determine if guided self-placement processes resulted in a markedly different institutional throughput profile for those colleges that made this shift relative to those that did not.

**In summary, we make the following recommendations:**

- **While none of the tests had very strong performance, for colleges that need to select a placement test, the ACCUPLACER ESL Reading Skills Test may be the best available option as it had the most consistent results in terms of predicting English throughput.**
- **To the extent that colleges are continuing to use ESL placement tests, it would be beneficial to reactivate MIS data collection of raw placement scores to facilitate and improve placement test and cut-score validation work in the future.**
- **Collection and reporting of data and test scores related to colleges' initial assessment of English proficiency (e.g., TOEFL scores, IELTS scores) as well as international students' native language would be valuable to future research into the assessment and placement of international students.**
- **Future research should analyze the effect of the sudden shift to using guided self-placement processes during the pandemic and evaluate whether overall institutional throughput improved as a result relative to the institutional throughput achieved using placement tests in the period just before the pandemic.**

# References

- Buenrostro, M. & Maxwell-Jolly, J. (2021). *Renewing our promise: Research and recommendations to support California's long-term English learners*. Californians Together. [https://californianstogether.org/wp-content/uploads/2021/10/Renewing\\_Our\\_Promise\\_to\\_LTEs.pdf](https://californianstogether.org/wp-content/uploads/2021/10/Renewing_Our_Promise_to_LTEs.pdf)
- Bahr, P. R., Fagioli, L. P., Hetts, J., Hayward, C., Willett, T., Lamoree, D., Newell, M. A., Sorey, K., & Baker, R. B. (2019). Improving placement accuracy in California's community colleges using multiple measures of high school achievement. *Community College Review*, 47(2), 178-211.
- Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, 36(1), 9-48.
- Bracco, K. R., Dadgar, M., Austin, K., Klarin, B., Broek, M., Finkelstein, N., Mundry, S., & Bugler, D. (2014). *Exploring the use of multiple measures for placement into college-level courses: Seeking alternatives or improvements to the use of a single standardized test*. San Francisco, CA: WestEd.
- Bunch, G. C., Endris, A., Panayotova, D., Romero, M. & Llosa, L. (2011). *Mapping the terrain: Language testing and placement for US-educated language minority students in California's community colleges*. Report prepared for the William and Flora Hewlett Foundation. <http://escholarship.org/>
- California Department of Education. (2021a). *English learners in California schools*. <https://www.cde.ca.gov/ds/sg/englishlearner.asp>
- California Department of Education. (2021b). *Facts about English learners in California*. <https://www.cde.ca.gov/ds/ad/cefelfacts.asp>
- Hayward, C. (2020). *Maximizing ELL completion of transferable English: Focus on US high school graduates*. The Research and Planning Group for California Community Colleges. [https://rpgroup.org/Portals/0/Documents/Projects/MultipleMeasures/AB705\\_Workshops/Maximizing-English-Language-Learners-Completion\\_September2020.pdf](https://rpgroup.org/Portals/0/Documents/Projects/MultipleMeasures/AB705_Workshops/Maximizing-English-Language-Learners-Completion_September2020.pdf)
- Hayward, C. & Willett, T. (2014). *Curricular redesign and gatekeeper completion: A multi-college evaluation of the California Acceleration Project*. The Research and Planning Group for California Community Colleges. [https://rpgroup.org/Portals/0/Documents/Projects/California%20Acceleration%20Project%20\(CAP\)%20Evaluation/CAP\\_Summary\\_Final.pdf](https://rpgroup.org/Portals/0/Documents/Projects/California%20Acceleration%20Project%20(CAP)%20Evaluation/CAP_Summary_Final.pdf)
- Hayward, C., Kanno, Y., Rios-Aguilar, C., & Vo, D. (2022). *English learners' pathways in California's community colleges under AB 705* [Policy report]. Policy Analysis for California Education.

- Hern, K., & Snell, M. (2010). *Exponential attrition and the promise of acceleration in developmental English and math*. Chabot College.
- Llosa, L., & Bunch, G. C. (2011). *What's in a test? ESL and English placement tests in California's community colleges and implications for US-educated language minority students*. Report prepared for the William and Flora Hewlett Foundation. <http://escholarship.org/>
- Mejia, M. C., Rodriguez, O., & Johnson, H. (2016). *Preparing students for success in California's community colleges*. Public Policy Institute of California. [https://www.ppic.org/wp-content/uploads/content/pubs/report/R\\_1116MMR.pdf](https://www.ppic.org/wp-content/uploads/content/pubs/report/R_1116MMR.pdf)

# The Research and Planning Group for California Community Colleges

As the representative organization for Institutional Research, Planning, and Effectiveness (IRPE) professionals in the California Community Colleges (CCC) system, The RP Group strengthens the ability of CCC to discover and undertake high-quality research, planning, and assessments that improve evidence-based decision-making, institutional effectiveness, and success for all students.

## Project Team

Craig Hayward

Mallory Newell

Terrence Willett

Loris Fagioli

Daisy Segovia

Kelley Karandjeff

[www.rpgroup.org](http://www.rpgroup.org)