Locally Developed/Managed Performance Assessment
ESL English Language Essay Sample

<u>Standardization and Administration</u>

An ESL English Language Essay Sample assessment was developed to guide student placement recommendations into the college's ESL curriculum sequence. The curriculum comprises a sequence of 5 ESL levels with each level corresponding to one ESL course. The final course in the sequence is equivalent to a transfer-level English composition course.

The ESL Essay Sample consists of 10 different writing prompts, which had previously been developed by ESL and English faculty. Initially, the prompts were used to elicit end-of-term essays for students in ESL and English courses. The ESL Essay Sample is scored using the scoring rubric presented in Table 1. The scoring rubric comprises 4 rubric criteria which are considered holistically to determine a final score on a scale from 0 and 6 points. The scoring rubric was developed by two ESL faculty members, who also conducted the scorer training for the remaining ESL faculty. The scorer training was conducted in a two-hour training session that provided instruction in the interpretation and application of the scoring rubric and provided an opportunity for practice scoring. Each student essay is scored by one ESL faculty member. The final essay score is then used to make a recommendation for placement into one of the five ESL course levels. Cut scores for placement recommendations were determined in a separate cut score study conducted by the ELS faculty. The cut score study is described in a later section of this report.

The ESL Essay Sample is administered to all ESL students prior to course enrollment. Students are allowed to select one prompt and have two hours to compose an essay addressing the writing prompt. Students are not allowed to use a dictionary or thesaurus.

Table 1. ESL Essay Sample Score Rubric

Criteria	Score 5-6	Score 3-4	Score 1-2	Score 0
1.Response to prompt	Exhibits an insightful response to the text; effectively addresses all tasks	Exhibits an adequate response to the text; addresses most aspects of the task	May not respond adequately to the text; may ignore some aspects of the task	Demonstrates a failure to comprehend the tasks at hand
2.Organization, development, and support	Is well organized and substantially developed with effective examples and evidence	Is unevenly organized and generally developed with some effective examples and evidence	May lack coherent structure and effective examples	Lacks any structure or development; may be inappropriately brief
3.Style (diction and syntax)	Makes sophisticated syntactic choices and uses precise diction	May lack syntactic variety and exhibit inexact diction	Often lacks precise word choice and syntactic variety	Lacks control of syntax and vocabulary

4.Writing conventions	Usually employs correct grammar, punctuation, and spelling	Contains a few grammar errors, but generally observes conventions	Contains errors that interfere with meaning	Contains numerous grammatical errors that interfere with meaning

Accommodations

Accommodations are granted on the basis of a student's individual college accommodation plan. Approved accommodations include distraction-free rooms, Braille or large print, extended time, wheelchair-accessible testing stations, screen-reading technology, provision of scribes, and permission to bring and take medication.

Content Validation

A content validation study was conducted in which five members of the ESL faculty were asked to rate on a five-point scale the congruence of each rubric element to specific course prerequisites. The five point scale had the following scale points:

1 = *no match* between the scoring rubric and the course objectives/skills,
2 = *little match* between the scoring rubric and the course objectives/skills,
3 = *moderate match* between the scoring rubric and the course objectives/skills,
4 = *good match* between the scoring rubric and the course objectives/skills,
5 = *strong match* between the scoring rubric and the course objectives/skills.

Table 2 shows the mean faculty ratings of the congruence between the scoring rubric criteria at each score point and the specific course prerequisites/skills necessary for each of the five levels of the ESL sequence. An average rating across the four rubric criteria is also provided for each score point, the largest of which is highlighted in grey. For lower ESL course levels, the faculty found on average a stronger degree of congruence (i.e., highest average rating) between rubric criteria at lower score points. For higher ESL course levels, the faculty members found higher degrees of congruence with the rubric criteria of higher score points.

Table 2: Faculty Mean Rating of congruence between scoring rubric criteria and course prerequisites/skills

Scoring Point	Rubric Criteria	Congruent with ESL 111	Congruent with ESL 123	Congruent with ESL 133	Congruent with ESL 141	Congruent with ESL 151
Score 0	Criterion 1	5.0	1.8	1.0	1.0	1.0
	Criterion 2	5.0	2.4	1.6	1.0	1.0
	Criterion 3	5.0	2.8	2.0	1.0	1.0
	Criterion 4	5.0	3.0	1.8	1.0	1.0
	Average	**5.0**	**2.5**	**1.6**	**1.0**	**1.0**
Score 1-2	Criterion 1	4.0	5.0	2.4	2.0	1.4
	Criterion 2	4.2	5.0	2.8	3.0	2.4
	Criterion 3	4.2	5.0	3.4	2.2	1.8
	Criterion 4	5.0	5.0	2.2	1.4	1.4
	Average	**4.35**	**5.0**	**2.7**	**2.15**	**1.75**
Score 3-4	Criterion 1	1.8	3.4	4.0	5.0	5.0
	Criterion 2	2.6	3.6	4.0	4.0	4.2
	Criterion 3	3.8	4.0	4.0	3.8	3.0
	Criterion 4	1.8	3.2	3.8	3.0	2.4
	Average	**2.5**	**3.55**	**3.95**	**3.95**	**3.65**
Score 5-6	Criterion 1	1.0	1.6	3.2	5.0	5.0
	Criterion 2	1.0	2.2	3.8	4.2	4.4
	Criterion 3	1.0	1.2	3.2	4.0	4.0
	Criterion 4	1.8	1.8	3.6	4.8	5.0
	Average	**1.2**	**1.7**	**3.45**	**4.5**	**4.6**

<u>Criterion Validation</u>

A faculty survey was conducted in the spring semester of 2021 to collect faculty judgments of the appropriateness of student placement into the ESL courses. ESL instructors were asked to rate how well each student in their ESL course was prepared for the level of work in the course using the following rating scale:

1 = the student is under-prepared,
2 = the student is adequately prepared,
3 = the student is over-prepared.

Table 3 shows faculty judgments of placement accuracy by ESL level. Overall, the ESL instructors judged that 80% of their students were placed accurately (i.e., received a rating of 2).

Table 3. Faculty Judgments of Student Placement Accuracy

Course	Curriculum Level	Agreement Rate	Number of Students
ESL 111	1	76%	114
ESL 123	2	79%	119
ESL 133	3	80%	136
ESL 141	4	79%	128
ESL 151	5	84%	125

A parallel survey was conducted to collect student judgments of the appropriateness of their placement into the ESL course sequence. Students from 25 sections (i.e., 5 sections from each level of the ESL curriculum) were asked to answer the question "Which of the following is most true of your placement?" by selecting from the following three answer choices:

1 = This course is too easy for me.
2 = This course is the right fit for me.
3 = This course is too difficult for me.

Table 4 shows student judgments of placement accuracy by ESL level. Overall, between 75% and 88% of students agreed with their course placement (i.e., selected answer choice 2). On average, the agreement rate for students is 80%, indicating the students agree with their placement into the ESL sequence.

Table 4. Student Judgments of Accuracy of ESL course placement

Course	Curriculum Level	Agreement Rate	Number of Students
ESL 111	1	75%	105
ESL 123	2	75%	109
ESL 133	3	79%	124
ESL 141	4	84%	115
ESL 151	5	88%	118

An additional student survey will be conducted in the spring semester of 2023. The student survey will be administered together with the ESL essay sample, and will include all students taking the ESL Essay Sample in that semester. In the survey, students will consider a list of English proficiency skills organized from low to high, and select the proficiency level that best represents them. Figure 2 below shows the survey students will complete. Student responses will then be scored based on the level they select, with the lowest level receiving a score of 1 and the highest level receiving a score of 4. Scores from the survey will then be correlated with students' ESL Essay Sample score to yield information on the strength of the relationship between student self-assessment and performance on the ESL Essay Sample.

Figure 2. Student Survey Form

Student Self-Assessment Survey
Consider the English proficiency skills listed in each box. Select the box that best describes your English skills using the sentence "Today I believe I can …"   Choose ONLY 1 box.

Today I believe I can … _____	☐ Write 5 to 7 page essays in academic English with little or no help ☐ Read college level texts in English, including a 300-page novel or nonfiction with little or no dictionary help. ☐ Fully understand a college lecture in English on academic topics such as Biology, History, or Sociology.
Today I believe I can … _____	☐ Write 2 to 3 page essays in academic English with some help. ☐ Read short college- level texts in English, including a 200-page novel or nonfiction with dictionary help. ☐ Understand most of a college lecture in English on academic topics such as Biology, History, and Sociology
Today I believe I can … _____	☐ Write a group of sentences in English with help. ☐ Read a group of sentences in English, but sometimes I do not know all of the words. ☐ Understand a slow-paced conversation in English on a familiar topic or in a practical everyday situations such as shopping or the weather.
Today I believe I can … _____	☐ Write some words and a couple of sentences. I know my English ABCs, numbers, and some words, but I need a lot of help. ☐ Read and understand some familiar words in simple sentences. ☐ Understand some simple questions, directions, or greetings.

<u>Consequential Validity</u>

Starting in the fall semester of 2022, data will be collected from each academic term over the next three years to determine rates of ESL course completion after initial placement. Data will also be collected from students who did not enroll in the initially recommended ESL course.

Data will be analyzed to provide the following information:

- Percentage of students who did and did not enroll in recommended ESL course
  - Overall
  - Disaggregated by course level
- Percentage of students per initial recommended ESL course level, who completed ESL curriculum sequence within three years
  - overall
  - disaggregated by gender
  - disaggregated by cultural/linguistic group

The college plans to disaggregate data for the following cultural/linguistic groups:

- Middle Eastern
- East Asian
- African
- Spanish
- Filipino
- Eastern European

A test bias survey was conducted during the spring 2022 semester with a panel of 6 ESL faculty members and a separate panel of 12 ESL students. The demographic composition of each panel is described in tables 5 and 6 below.

Table 5. ESL Faculty Panel Demographic Composition

		ESL Faculty Panel Member					
		1	2	3	4	5	6
Gender	Male	X					X
	Female		X	X	X	X	
Ethnicity	Asian	X					
	Black	-	-	-	-	-	-
	White		X	X		X	
	Hispanic				X		X
	Filipino	-	-	-	-	-	-

Table 6. ESL Student Panel Demographic Composition

		ESL Student Panel Member											
		1	2	3	4	5	6	7	8	9	10	11	12
Gender	Male	X		X	X			X		X			
	Female		X			X	X		X		X	X	X
Age	<20	X	X	X			X	X		X			
	21-30					X			X		X		
	>30				X							X	X
Cultural/ Linguistic Group	East Asian					X					X		
	African	X											
	Mid.-Eastern		X				X	X		X		X	X
	Hispanic			X					X				
	Filipino				X								

The panel members examined each of the 10 essay prompts for evidence of bias for gender, age, and ethnicity. The faculty and student panels convened separately to examine each of the 7 essay prompts for evidence of bias. Panelists were led through a training that explained the purpose of the meeting, the notion of test bias, and the rating task. Panelists were then asked to consider each essay prompt individually and to evaluate whether the prompt was in topic or language unfair or distasteful with respect to any particular group of gender, age, or ethnicity. Panelists who deemed a prompt unfair or distasteful were also instructed to add explanations of their rating decision. After the rating process, panelists discussed any prompts for which bias had been identified by a panelist. After the discussion, panelists were given the chance to revise any ratings. Responses from each panel member are summarized in tables 7 and 8 below. An "N" in the tables indicates that the panel member did not identify any potential bias for at least four of the seven essay prompts. A "Y" would indicate that a panel member did identify potential bias for three or more prompts. None of the faculty or student panel

members identified bias for three or more prompts. The survey therefore found no evidence of test bias.

Table 7. ESL Faculty Panel Member Responses

	ESL Faculty Panel Member					
	1	2	3	4	5	6
Evidence of Bias for Gender	N	N	N	N	N	N
Evidence of Bias for Age	N	N	N	N	N	N
Evidence of Bias for Ethnicity	N	N	N	N	N	N

Table 8. ESL Student Panel Member Responses

	ESL Student Panel Member											
	1	2	3	4	5	6	7	8	9	10	11	12
Evidence of Bias for Gender	N	N	N	N	N	N	N	N	N	N	N	N
Evidence of Bias for Age	N	N	N	N	N	N	N	N	N	N	N	N
Evidence of Bias for Ethnicity	N	N	N	N	N	N	N	N	N	N	N	N

<u>Disproportionate Impact Evidence</u>

Placement data were collected from a sample of 2,127 ESL students who were placed into one of the five ESL courses during the past two academic years on the basis of the ESL Essay Sample score. Placement rates were computed for six cultural/linguistic groups, all representative of the ESL student population in this college. To determine disproportionate impact, the EEOC 80% criterion was applied.

Placement into the ESL curriculum is broken up into the five ESL credit levels. The lowest credit level represents students placing into ESL 111. The second credit level represents those students who placed into ESL 123. The third credit level represents those students who placed into ESL 133. The fourth and fifth levels represent ESL 141 and ESL 151, respectively.

Tables 9 through 13 show placement rates into each of the five difference ESL course levels by cultural/linguistic group. The tables are ordered from highest/transfer-level ESL course (Table 9 for ESL 151) to lowest-level ESL course (Table 13 for ESL 111). Each table presents two disproportionate impact indicators: percentage point gap and proportionality index. Values in red, bold font indicate that the number of students from the particular student group that places into the ESL course is below expectations, thus suggesting disproportionate impact. For the percentage point gap, those values must be negative and outside the margin of error calculated for each group's sample size. For the proportionality index, those values must be less than 0.85.

Tables 9 shows that for placement into the highest level ESL course ESL 151 disproportionate impact was only observed for Middle Easter students, and only on the basis of the proportionality index. The percentage point gap indicator did not suggest disproportionate impact for this group. For the remaining lower-level courses, disproportionate impact was indicated by either one or both indicators for East Asian, Spanish-speaking, and Eastern European student groups. no disproportionately impacted based on the percentage point gap indicator. The proportionality index also suggests that placement of Middle Eastern students into ESL 141 was also disproportionately impacted.

The college will monitor placement rates of Middle Eastern students into the highest level ESL course. It is also possible that the high overall proportion of Middle Eastern students in our ESL student population affected results of the disproportionate impact analyses.

Table 9. Disproportionate Impact Indicators for ESL 151 Course Placement by Cultural-Linguistic Group

Cul./Ling. Group	ESL 151 Placement		Percentage Point Gap	Proportionality Index
	N	%		
Mid.-Eastern	60	35.7%	17.4%	**0.50**
East Asian	82	48.2%	43.2%	3.11
Spanish	20	11.8%	4%	1.41
Eastern European	7	4.3%	-3.8%	1.01
Total	169	8.0%		

Table 10. Disproportionate Impact Indicators for ESL 141 Course Placement by Cultural-Linguistic Group

Cul./Ling. Group	ESL 141 Placement		Percentage Point Gap	Proportionality Index
	N	%		
Mid.-Eastern	227	58.5%	31.4%	**0.81**
East Asian	115	29.8%	**14.5%**	1.92
Spanish	28	7.4%	**-11%**	0.87
Eastern European	17	4.3%	**-14.1%**	1.02
Total	387	18.4%		

Table 11. Disproportionate Impact Indicators for ESL 133 Course Placement by Cultural-Linguistic Group

Cul./Ling. Group	ESL 133 Placement		Percentage Point Gap	Proportionality Index
	N	%		
Mid.-Eastern	378	74.6%	52.9%	1.04
East Asian	63	12.4%	**-12.6%**	**0.80**
Spanish	46	9.2%	**-15%**	1.09
Eastern European	20	3.9%	**-20.3%**	0.92
Total	507	24.1%		

Table 12. Disproportionate Impact Indicators for ESL 123 Course Placement by Cultural-Linguistic Group

Cul./Ling. Group	ESL 123 Placement		Percentage Point Gap	Proportionality Index
	N	%		
Mid.-Eastern	354	76.2%	57.5%	1.06
East Asian	53	11.4%	**-11.7%**	**0.74**
Spanish	32	6.9%	**-15%**	**0.82**
Eastern European	25	5.4%	**-16.4%**	1.29
Total	464	22.1%		

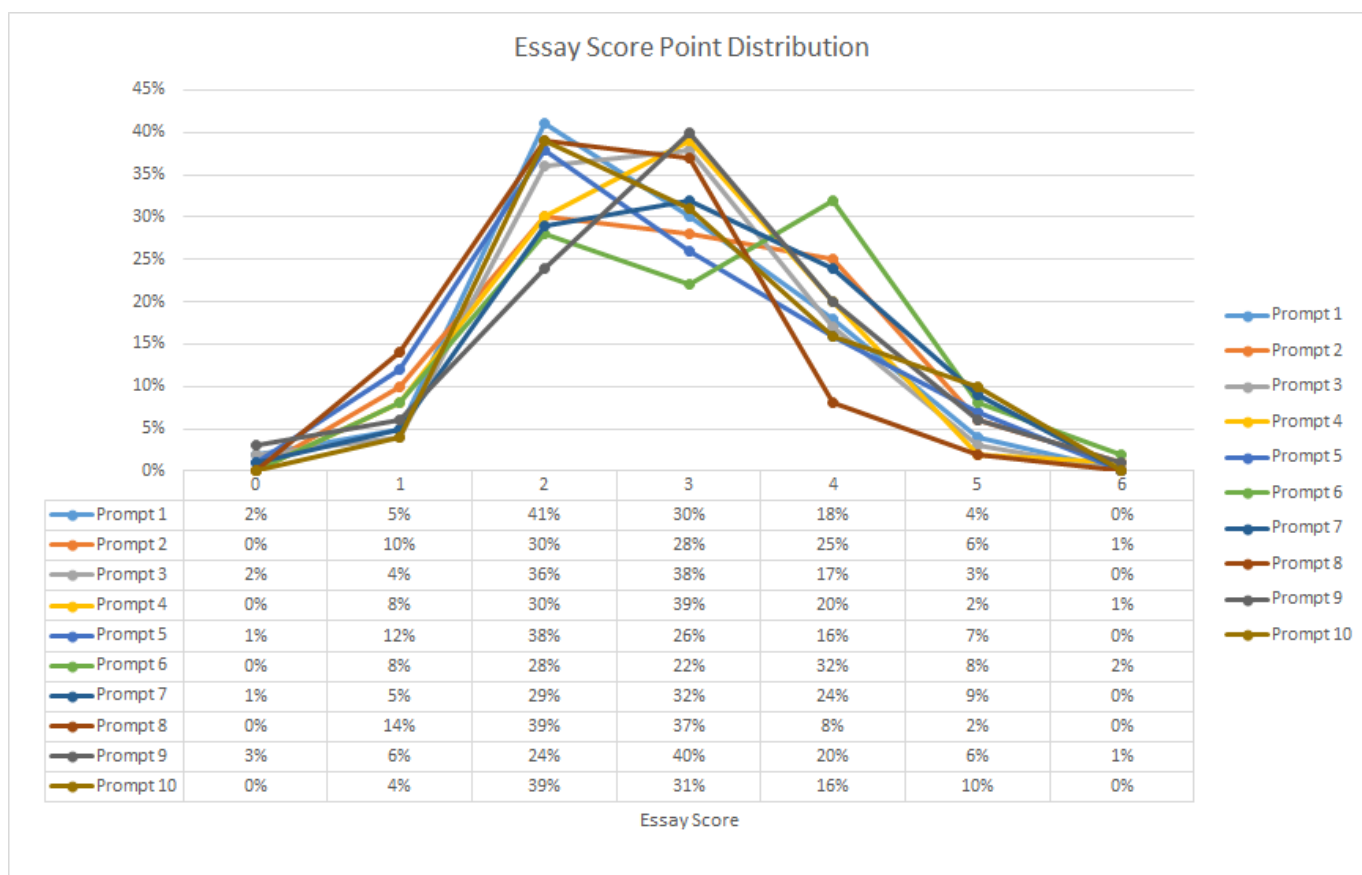Table 13. Disproportionate Impact Indicators for ESL 111 Course Placement by Cultural-Linguistic Group

Cul./Ling. Group	ESL 111 Placement		Percentage Point Gap	Proportionality Index
	N	%		
Mid.-Eastern	493	85.6%	71.5%	1.19
East Asian	13	2.3%	**-29.4%**	**0.15**
Spanish	50	8.7%	**-19%**	1.03
Eastern European	20	3.5%	**-24.1%**	**0.82**
Total	576	27.4%		

Reliability Evidence

Interrater reliability was computed for a sample of 225 student essays that had been written between July 1, 2020 and December 1, 2021 by ESL students who were subsequently placed into the ESL curriculum. Each essay was scored by two ESL instructors. A total of 5 ESL instructors participated in the scoring of these essays. On a 7-point scale (scored from 0 to 6), 99.7% of the scores agreed within one point. There was complete agreement on 78.2% of the essays, and disagreement by one point on 21.5% of the essays. No essays elicited disagreement by two or more points. Any disagreements were resolved by a third reader, who made the final scoring decision. The interrater reliability coefficient computed for this sample was $r = 0.95$, indicating high reliability of essay scores.

The equivalency of essay prompts was examined by graphing the distributions of essay scores for each prompt. The graphs are shown below in Figure 3. The considerable overlap of the distributions shows that the set of prompts result in approximately equivalent essays.

Figure 3. Distribution of Essay Scores by Prompt

Essay Score Point Distribution

	0	1	2	3	4	5	6
Prompt 1	2%	5%	41%	30%	18%	4%	0%
Prompt 2	0%	10%	30%	28%	25%	6%	1%
Prompt 3	2%	4%	36%	38%	17%	3%	0%
Prompt 4	0%	8%	30%	39%	20%	2%	1%
Prompt 5	1%	12%	38%	26%	16%	7%	0%
Prompt 6	0%	8%	28%	22%	32%	8%	2%
Prompt 7	1%	5%	29%	32%	24%	9%	0%
Prompt 8	0%	14%	39%	37%	8%	2%	0%
Prompt 9	3%	6%	24%	40%	20%	6%	1%
Prompt 10	0%	4%	39%	31%	16%	10%	0%

Essay Score

ESL English Language Essay Sample

Cut Score Setting

A cut score study was conducted in the fall of 2020 to set cut scores for the ESL Essay Sample in order to facilitate appropriate ESL course placement decisions.

A panel of 5 ESL faculty members was convened. All 5 members were experienced ESL instructors who had been teaching at least two or more of the courses in the ESL course sequence, with a combined teaching experience of 32 years. The 5 panelists were also participants in the test bias survey of the ESL Essay Sample. The demographic composition of the panel is shown in Table 5 above.

The cut score study was facilitated by another member of the ESL faculty who had been involved in the development and training of the scoring rubric. The facilitator began with an overview of the scoring rubric and a walk-through of prerequisite skills for each ESL course level. The facilitator then asked panelists to discuss the skills of a student who is minimally qualified for each course level. Panelists then created a list of minimally qualifying skills and abilities for each course level. The facilitator then presented a sample of 10 scored student essays representing all score points except '0'. Without knowledge of each essay's score, panelists were instructed to individually consider each essay and place it into the most appropriate course level using the list of minimally qualifying course skills and abilities as basis for their decision. After panelists made individual placement decisions, the facilitator informed the panel of each essay's score. Armed with this feedback, panelists were then asked to discuss and decide as a group the best course placement for each student. The resulting score ranges for each course level are presented in Table 10.

Table 10.

ESL Course	ESL Essay Sample Score Range
Level 1	0-1
Level 2	2
Level 3	3
Level 4	4
Level 5	5-6