

Will's LPTM (Language Proficiency Tracking Model)

An approach for a U.S. high school AP Spanish Language & Culture course

Table of Contents:

- [Introduction](#)
- [Expected Outcomes](#)
- [Starting with AP Spanish Language & Culture Assessments](#)
- [The Assessment Tasks](#)
- [How to Predict a Student's AP Exam Score](#)
- [Checking My Under-Qualification and Possible Bias](#)
- [Estimating ACTFL Proficiency Ratings](#)
- [Baseline Assessment Scores](#)
- [Data Analysis](#)
- [Setting Year-End Goals](#)
- [Personalized Feedback](#)
- [Strategic Curriculum / Plan](#)
- [Final Assessments](#)
- [Evidence of Student Growth](#)
- [Conclusions & Reflections](#)

Introduction: This file shows one possible method for tracking the Spanish language proficiency and growth of students based on a year-long AP Spanish Language & Culture high school course in the United States with students who have between beginning and intermediate levels of proficiency according to the ACTFL scale.

Expected Outcomes:

- We will fulfill each student's maximum potential for growth by the end of the school year.
- We will be able to measure the growth of each student, as well as the group, throughout the year.
- We will be able to predict what score each student is likely to score on the AP Spanish Language & Culture exam.
- We can estimate each student's *expected* Spanish language proficiency according to the ACTFL scale.

Starting with AP Spanish Language & Culture Assessments:

The first step to any language proficiency tracking model is to create or identify some sort of common assessment framework that you can use for pre- and post-learning evaluations. The content of these assessments shouldn't be the same, but the format and level of difficulty should be nearly identical so that way the student results on each of them are truly comparable.

I decided to use AP Spanish Language & Culture assessments that the College Board has made available to the public. I had been a part of this type of process once before and saw how my student-teacher's school district used self-made questionnaires and prompts coupled with ACTFL rubrics. I really liked this method because ACTFL is one of the most respected organizations related to foreign language education in the US. In this case, though, not only did I want students to improve their proficiencies, but I also had to use the course to prepare students for the AP exam, so I figured that using previous AP assessments would be the most efficient way to work towards both of these objectives at the same time.

The AP Spanish Language & Culture exam has two parts, or halves: One that involves reading and listening with multiple choice questions and a second that involves open-ended writing and speaking tasks. To access multiple choice exams that you can use for your pre- and post-learning assessments, you need to create an AP Classroom account. AP Classroom is an online platform of resources for teachers and students. At the time of this writing, you can access multiple choice practice exams from 2013, 2019, and 2020. To access free-response questions that you can use for your assessments, visit the page linked below, where you can find all of the free-response questions from previous AP Spanish Language & Culture exams dating back to 2012.

<https://apcentral.collegeboard.org/courses/ap-spanish-language-and-culture/exam/past-exam-questions>

The Assessment Tasks:

I see the AP Spanish Language & Culture exam as having six tasks:

- 1A: Reading + multiple choice
- 1B: Listening + multiple choice
- 2C: Write a formal email
- 2D: Write an argumentative essay
- 2E: Speak in a simulated conversation
- 2F: Present orally (speak) a cultural comparison

How to Predict a Student's AP Exam Score

Disclaimer: I am not a certified AP exam grader. But, below is a logic and research-based approach.

Students can earn scores that range from 1-5 on an AP exam. AP exam graders review each student's exam and assign one of these scores. Here is a table from the College Board's website that shows how they interpret these scores:

<u>AP Exam Score</u>	<u>Level of Qualification to Receive College Credit</u>	<u>College Course Grade Equivalent</u>
5	Extremely well qualified	A+ or A
4	Very well qualified	A-, B+, or B
3	Qualified	B-, C+, or C

2	Possibly qualified	—
1	No recommendation	—

Source: <https://apstudents.collegeboard.org/about-ap-scores/ap-score-scale-table>

This numerical score is based on a student's performance on all six tasks of the exam. Each half of the exam has a 50% weight on the student's final score. Within each of these halves are tasks that are meant to assess a student's four language skills: Reading, listening, writing, and speaking. The first exam part assesses reading and listening, while the second part assesses writing and speaking. You can also interpret the exam as giving each skill area performance a 25% weight on a student's final score. Here is a chart that helps show these weights:

<u>Part</u>	<u>Weight (by section)</u>	<u>Skill Area</u>	<u>Weight (by skill area)</u>	<u>Task</u>	<u>Weight by Task</u>
1. Multiple choice	50%	Reading	25%	1A: Reading + multiple choice	25%
		Listening	25%	1B: Reading + multiple choice	25%
2. Open-ended	50%	Writing	25%	2C: Write a formal email	12.5%
				2D: Write an argumentative essay	12.5%
		Speaking	25%	2E: Speak in a simulated conversation	12.5%
				2F: Speak a cultural comparison	12.5%

In order to estimate a student's expected AP exam score (1-5), you first need to evaluate their skill abilities based on their performances on the six exam tasks. Then, you need to use these individual scores to calculate an overall score.

To identify a student's score on the first part of the exam (reading and listening with multiple choice questions), you need to calculate the percentage of correct answers and then use a conversion scale to assign an AP score (1-5) based on the percentage score. On average, the AP Spanish Language & Culture scale looks like this:

<u>% Score</u>	<u>ACTFL rating</u>
0-32%	1
33-49%	2

50-66%	3
67-80%	4
81-100%	5

However, these conversion scales change each year based on the level of difficulty of each year's exam. Yes, the conversion scales are normally similar to the one above, but very often not exactly the same. Therefore, it is important to find and use the conversion scale for the specific multiple choice practice exam (2013, 2019, 2020) that you take from AP Classroom and use for your assessment. These conversion scales are also available in AP Classroom.

Next, you have to evaluate a student's free response submissions on part two of the assessment. The College Board provides rubrics that you can use to give students a score of 0-5 for each of these responses. Each number matches with a word that describes the quality of a response that would receive the numerical score:

- 5 = Strong
- 4 = Good
- 3 = Fair
- 2 = Weak
- 1 = Poor
- 0 = Unacceptable

And then, within each rubric, there are a variety of components, such as treatment of topic, vocabulary, grammar, comprehensibility, accuracy of information, organization, and more, that the evaluator considers when deciding what numerical score to give a student's response. You can view these rubrics on the AP Central portion of the College Board's website. Click this link [here](#).

Finally, once you have a score of 0-5 for each of a student's six tasks, you can calculate his/her *expected* overall score like this:

<u>Task</u>	<u>Score</u> (Enter 0-5 below)	<u>Weight</u> (Multiply the task score by this percentage.)	<u>Points Tally</u> (Add these numbers up to get the overall score below.)
Reading + multiple choice		x 0.25	
Listening + multiple choice		x 0.25	
Writing a formal email		x 0.125	
Writing an argumentative essay		x 0.125	
Speaking in a simulated conversation		x 0.125	
Speaking a cultural		x 0.125	

comparison			
Final Score (out of 5) →			

Checking my under-qualification and possible bias

It is important to acknowledge the possibility that I may make mistakes when evaluating my students' free response submissions, as I am not a certified AP reader. I also may be unaware of any biases or emotions that impact my evaluations of my students. Perhaps in some cases due to a desire to show them their growth as a motivational tactic, I give scores that are higher than they should be. Or, in other cases, perhaps I am too hard on them and give lower scores than they deserved for one reason or another. When I took the AP Summer Institute course in 2021 to learn how to teach the AP Spanish Language & Culture course, the greatest lack of clarity in our classroom involved the open-ended submissions. Our course leader implemented an activity during which we worked in small groups around the room to evaluate various free response submissions and there was never one time when all groups agreed on what score (0-5) to award to a submission. This event speaks to the fact that if one is not a certified AP Reader, he/she may likely not give the same score to a piece of student work that the AP would give.

Therefore, it is important for me to seek feedback about my evaluation methods and the accuracy of the scores that I award. In addition to taking the APSI course, I spoke with other AP Spanish teachers and linguists to find out how they use the rubrics, and I asked ChatGPT to evaluate my students' work using the College Board rubrics so that I could compare its evaluation of their work with my own.

These steps revealed that around 50% of the time, I was evaluating my students' work accurately, and that in the other 50% of instances, I was scoring the work slightly higher than I should have. For example, in these 50% of instances, if a students' work bordered on "high 3" to "low 4," I would tend to give the "low 4," while the more accurate score would be a "low 3."

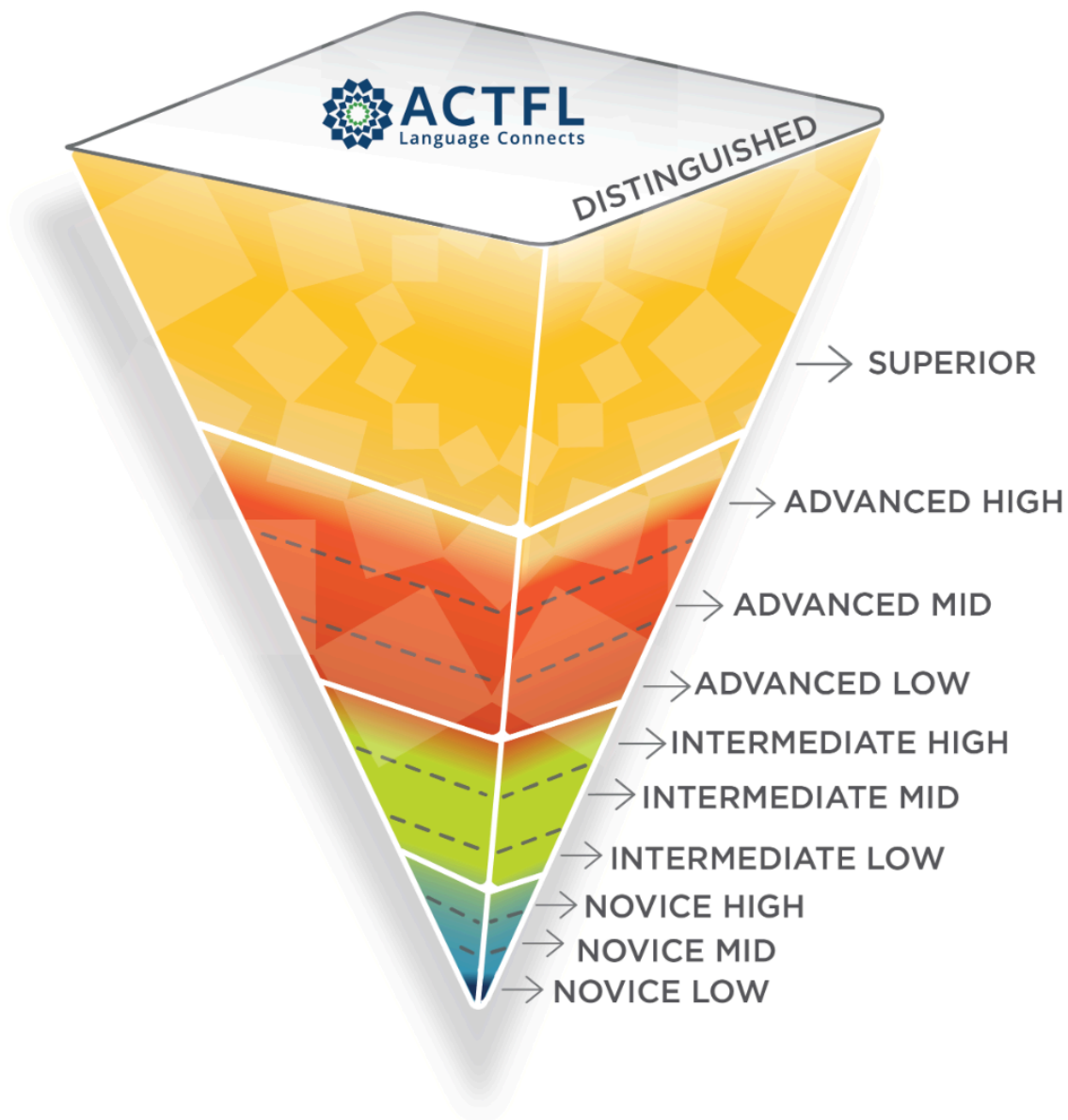
In order to make sure that I was scoring students as accurately as possible throughout this project, I reviewed and graded every piece of student work multiple times. Also, when I was stuck between two scores, unsure of which to give, I always gave the lower score.

Estimating ACTFL Proficiency Ratings

I also wanted to give my students an estimated proficiency ranking for their Spanish language abilities. When we talk about proficiency in a second language, the word that is most often thrown around is "fluent," which means "able to express oneself articulately." But, what if one doesn't have the confidence of a native speaker? What if your abilities fluctuate week-by-week or month-by-month based on your context? What if you can speak perfectly, but cannot read? Being "fluent or not" is way too narrow of a lens through which to view language proficiency. It would give my students a lot of clarity if they could describe their abilities by referencing the ratings of an official scale.

Around the world, there are different language proficiency scales that are used to help people rate their abilities in a language. One, for example, is CEFR (Common European Framework Reference) for Languages. You will hear most Europeans rate themselves according to this scale, which includes from low to high levels A1, A2, B1, B2, C1, and C2.

In the United States, we most commonly use the ACTFL scale. ACTFL stands for American Council for the Teaching of Foreign Languages. The picture below shows the levels of ACTFL's proficiency scale:



Here are to links to pages where you can read more about the ACTFL proficiency scale:

- In short:
 - <https://www.languagetesting.com/actfl-proficiency-scale>
 - <https://theglobalseal.com/actfl-language-proficiency-levels>
- In detail:
https://www.actfl.org/uploads/files/general/Resources-Publications/ACTFL_Proficiency_Guidelines_2024.pdf

In order to officially earn an ACTFL Spanish proficiency rating, there is no option other than to schedule and pay for a formal evaluation with a certified evaluator. However, there have been studies done and data published that show that one's score on the AP Spanish Language & Culture exam is a very reliable predictor of the same person's language proficiency according to the ACTFL scale.

Here is a chart that shows how one's AP exam score indicates his/her likely ACTFL rating:

<u>AP score</u>	<u>ACTFL rating</u>
1	Novice low-mid
2	Novice high
3	Intermediate low-mid
4	Intermediate high
5	Advanced low (at least)***

*****Note:** The AP exam ceiling is obviously lower than that of the ACTFL scale. In fact, GamutSpanish.com says that the level AP exam is somewhere between a CEFR B1 and B2 standard (the European scale's intermediate rankings). This means that a student who earned an AP 5 could be someone who "worked hard to barely get there," or a native speaker who passed easily and only didn't score higher because a higher score on the AP scale doesn't exist. If you earn an AP 5, it might be worth considering taking an official evaluation to see if your ACTFL rating could be any higher.

Sources:

- <https://www.actfl.org/uploads/files/general/SealofBiliteracy-ComparisonofScales-TalkingPointsv3.pdf>
- <https://www.doe.mass.edu/scholarships/biliteracy/skills-comparison.docx>

Baseline Assessment Scores

After doing my research and arriving at the conclusions mentioned above, I was ready to implement the project. First, here are the scores and estimated ratings of my AP Spanish Language & Culture students at the beginning of the school year.

Note: Instead of marking the individual task scores with AP "numbers" (0-5), I used percentages. This is one area where one could say that I am out of alignment with "AP grading methods." The AP rubrics are meant to be holistic, in the sense that when scoring a student's piece of work, you consider all the components and then based on your trained expertise, award a number that matches the student's overall body of work. I chose to score each component in the rubric separately before calculating an overall score. This method would allow me to give more specific feedback to students about their strengths and areas for improvement because I could reference specific components within the rubric. I also felt that reading percentages on a scale from 0-100 would offer greater clarity than scores from 0-5.

(baseline assessments continued...)

<u>Skill Area:</u>	<u>Reading</u>	<u>Reading + Listening</u>	<u>Writing</u>		<u>Speaking</u>		<u>Total Average</u>	<u>Predicted AP Score</u>	<u>Estimated ACTFL Rating</u>
<u>Test Section:</u>	<u>1A</u>	<u>1B</u>	<u>2C</u>	<u>2D</u>	<u>2E</u>	<u>2F</u>			
Student 1	10%	37%	40%	4%	37.5%	20%	25%	1	Novice low-mid
Student 2	43%	43%	32.5%	49%	40%	38%	41%	2	Novice high
Student 3		69%	60%	44%	100%	100%	N/a	N/a	N/a
Student 4	70%	77%	94%	62%	85%	88%	79%	4	Intermediate high
Student 5	63%	51%	57%	42%	20%	20%	42%	2	Novice high
Student 6	47%	45%	40%	64%	42.5%	56%	49%	2	Novice high
Student 7	47%	43%	57%	51%	45%	36%	47%	2	Novice high
Student 8	43%	40%	57%	47%	62.5%	52%	50%	2	Novice high
Student 9		23%	20%	27%			N/a	N/a	N/a
Class Average (test section):	46%	48%	51%	43.33%	54.06%	51.25%	48%	2	Novice high
Class Average (skill area):			47%		53%				

Data Analysis

These scores don't only tell us each student's "year-start level of Spanish," but they also provide insight into what specific skills and learning objectives each student needs to practice and work towards, as well as classwide trends that will tell the teacher how to plan and implement instruction throughout the year.

To see how I assembled and analyzed these baseline assessment scores in order to plan the course and work towards fulfilling each student's maximum potential for growth, you can open this file: [Will's LPTM: Baseline Assessments Data & Analysis](#).

Setting Year-End Goals

I set the following 13 goals based on the data:

1. Practice all four skill areas (reading, listening, writing, and speaking) strategically and more or less equally throughout the year.
2. Prepare for all six sections of the AP Spanish Language & Culture exam and give more or less equal attention to these sections throughout the year.
3. Practice all reading skill categories and learning objectives throughout the year. (There are four skill categories, eight skills, and 24 learning objectives specified in the AP curriculum.)
4. Give extra time and attention to reading creative pieces of writing, such as short stories, poems, letters, and journal entries, throughout the year. During these learning activities, make sure that there are questions that force students to analyze the literature.
5. Make culture an integral part of the course. The teacher needs to provide students with opportunities to reflect on their own cultures and learn about Spanish-speaking cultures in a variety of contexts on a near-everyday basis.
6. Practice listening to all five types of listening selections (informative report, conversation, interview, instructions, and presentation) throughout the year.
7. Practice all listening skill categories and learning objectives throughout the year. (There are four skill categories, eight skills, and 24 learning objectives specified in the AP curriculum.)
8. Learn and practice all the components of writing an effective formal email and argumentative essay.
9. The one student who scored 94% on his formal email baseline assessment should opt out of instructional modules that cover writing a formal email and instead use the time to practice the skills that, according to his scores, need more practice: Reading and listening with multiple choice questions.
10. Learn and practice how to use a variety of words, phrases, and verb tenses in order to alter the tone of their writing.
11. Learn about all the components of speaking effectively and practice speaking everyday.
12. The one student who scored 100% on both speaking tasks of the baseline assessment should opt out of speaking activities and instead practice a skill that his scores show he needs to develop more, which, in his case, is writing.
13. Learn and practice skills that will help students navigate moments in conversation when they don't know how to express themselves or make a mistake trying to do so: Circumlocution, self-correction, accepting and not dwelling on mistakes, and perseverance.

Personalized Feedback

After reviewing the baseline assessments, I assemble individualized feedback for each student in order to help make them aware of their individual strengths, areas for improvement, and goals for the year.

For the multiple choice sections, one idea is to have students review their own exams using the answer key provided in AP Classroom and make a list of the skills and learning objectives that their exam results are telling them they need to most focus on this school year. The AP Course Guide identifies four skill categories, eight skills, and 24 learning objectives to developing the interpretive mode of communication (reading and listening). In AP Classroom, along with each of the three available practice exams, you can access an answer key that not only provides the answers to each question, but also provides the specific skills and learning objectives associated with each question. As students review their work, they can make a list of the skills and learning objectives associated with their incorrectly answered questions in order to help themselves begin familiarizing themselves with their individual areas for growth.

Such a resource is not available for the free response submissions, so for these, I have to put more time and effort into providing each student with personalized feedback. I make sure that students understand the rubrics and then use them to show scores and provide specific pieces of advice about how to hone their strengths and make improvements based on their struggles. To view samples of the personalized feedback that I provide to students, click here: [Will's LPTM: Feedback Samples](#).

Strategic Curriculum / Plan

Implementing this type of project adds a lot more to think about during a school year for any teacher who has a workload that is already full, so it is important to plan a year of instruction that will help you stay on track, manage the balance within your instruction, and finish the year with maximized student growth. Therefore, I created and implemented a yearlong plan that checked the following boxes:

- Taught and provided practice opportunities for all six sections of the AP exam
- Provided a balanced mix of reading, listening, writing, and speaking opportunities
- Covered all six chapters of the AP Spanish Language & Culture curriculum
 - Families and communities
 - Technology, the individual, and society
 - Beauty and aesthetic
 - Contemporary life
 - Global challenges
 - Personal and public identities
- Touched on all 21 Spanish-speaking “places” (20 countries plus one territory) in the world in at least once context at least once

Click the following link to access an outline of my curriculum: [Will's LPTM: Yearlong Curriculum](#).

Final Assessments

After a year of hard work, we ended with these scores:

<u>Skill Area:</u>	<u>Reading</u>	<u>Listening</u>	<u>Writing</u>		<u>Speaking</u>		<u>Total Average</u>	<u>Predicted AP Score</u>	<u>Estimated ACTFL Rating</u>
<u>Test Section:</u>	<u>1A</u>	<u>1B</u>	<u>2C</u>	<u>2D</u>	<u>2E</u>	<u>2F</u>			
Student 1	60%	63%	80%	62%	60%	48%	62%	3	Intermediate low
Student 2	77%	74%	97%	87%	98%	76%	83%	4	Intermediate high
Student 3	67%	77%	97%	71%	100%	100%	82%	5	Advanced low (at least)
Student 4	87%	86%	100%	96%	98%	88%	91%	5	Advanced low (at least
Student 5	63%	43%	77%	76%	75%	38%	60%	3	Intermediate low
Student 6	70%	77%	87%	87%	100%	38%	76%	4	Intermediate high
Student 7	70%	80%	80%	84%	98%	60%	78%	4	Intermediate high
Student 8	67%	71%	100%	100%	85%	78%	80%	4	Intermediate high
Student 9	37%	57%	54%	56%		20%	40%	N/a	N/a
Class Average (test section):	66%	70%	86%	80%	89%	61%	72%	4	5.75
Class Average (skill area):			83%		75%				

Evidence of Student Growth

Student	Percentage Scores			AP Scores (1-5)			ACTFL Ratings		
	Year-Start	Growth	Year-End	Year-Start	Growth	Year-End	Year-Start	Growth	Year-End
Student 1	25%	37%	62%	1	2	3	Novice low	3 levels	Intermediate low
Student 2	41%	42%	83%	2	2	4	Novice high	3 levels	Intermediate high
Student 3			82%			5			Advanced low (at least)
Student 4	79%	12%	91%	4	1	5	Intermediate high	1 level	Advanced low (at least)
Student 5	42%	18%	60%	2	1	3	Novice high	1 level	Intermediate low
Student 6	49%	27%	76%	2	2	4	Novice high	3 levels	Intermediate high
Student 7	47%	31%	78%	2	2	4	Novice high	3 level	Intermediate high
Student 8	50%	30%	80%	2	2	4	Novice high	3 levels	Intermediate high
Student 9			40%						
Total (Class Averages)	48%	24%	72%	2	2	4	Novice high	2.5 levels	Intermediate mid-high

Conclusions & Reflections

- **Yes, the project was a success.** On average, students in the class grew their Spanish language proficiency by 24%, (2) AP test points, and (2.5) ACTFL proficiency levels.
- **Surprise #1:** I am surprised that even at the end of the year out of the four skill areas reading is the lowest on average in the class.
- **Surprise #2:** I am also surprised that the simulated conversation scores (2E) at the end of the year were so high (89% average).
- **Idea for Improvement #1: Introduce more authentic texts and increase the level of reading and listening challenges at lower course levels.** Before implementing this project, I would have said with confidence that out of all four language skill areas, the speaking scores would be the lowest in the class. Naturally, I just figure that speaking a language is the hardest part, or at least the last piece of the puzzle to fall into place for someone learning the language. But, the “surprises” mentioned above seem to suggest otherwise. I suspect that maybe speaking “comes last” because of the psychological challenges that come with it—Not only needing skills and knowledge, but also having to get past the nerves and anxiety of performing something new in live-time in front of other people who may be better than you, for example. The data actually shows that at a high, authentically immersive level in Spanish, reading and listening are harder to do well than writing and speaking. It actually makes a lot of sense, too. Speaking is a more “controllable skill” because you can “operate within your means.” In other words, you can circumlocute and use the words that you know in order to be comprehensible and communicate basically, but still effectively. Reading and listening, however, are “uncontrollable” because in authentic, real-world situations, there are so many variables that may present themselves to you whether you want it or not—For example, in Spanish, the wide variety of dialects, vocabulary, and accents. The high dialectal variation, or dialect continuum, in Spanish could be a reason that in my classroom, proficiency in the interpretive modes/skills was the hardest to achieve.
- **Idea for improvement #2: Show students how to identify and log the skill areas and learning objectives related to their incorrect multiple choice answers of the baseline assessment** as I described in the “Personalized Feedback” section. (I checked this information myself, but I did not show and talk about it to students.)
- **Idea for improvement #3: Return baseline assessment feedback to students immediately after taking the assessment.** It took me so long to grade some of these assessments that I was only able to return them bit-by-bit at relevant points throughout the year. I sometimes feared that at later points in the year, some of the feedback wasn’t as relevant as it could have been because the students at that point were so far removed from their performance.