A Teacher’s Guide to Creating Quality Assessments for a Bilingual Population

By: Leigh Sanna
# Table of Contents

Common Core State Standards.......................................................................................... 3-4
Culturally Sensitive Assessments.................................................................................. 5-7
  Writing Directions........................................................................................................ 5
  Unbiased Language....................................................................................................... 5
  Aligning to the Common Core..................................................................................... 6
Creating Multiple Choice Questions................................................................................ 8-11
  Aligning to the Common Core..................................................................................... 8
Writing Essay Prompts..................................................................................................... 12-14
  Aligning to the Common Core..................................................................................... 12
Flexibility in Administration............................................................................................ 15
Checking-in After the Test............................................................................................... 16
Common Core State Standards

What are the Common Core State Standards?

The Common Core State Standards (CCSS) were created to establish a complete and ordered list of standards required for each grade level in mathematics and literacy. These new standards “include rigorous content and application of knowledge through higher order skills” (Marrongelle et al, 2013, p.202). The purpose of these standards is to better prepare students for higher education. Currently, these standards have been created both for Mathematics and English Language Arts.

These standards have been adopted by forty-five states. The only states yet to use such standards in their education systems are: Alaska, Minnesota, Texas, Nebraska, and Virginia. Although the remaining states have all officially decided to accept the Common Core, the dates for which full implementation will take place vary (Common Core State Standards Initiative, 2012).

Where can I find the standards?

http://www.corestandards.org/the-standards

- Mathematics: http://www.corestandards.org/Math
- English Language Arts: http://www.corestandards.org/ELA-Literacy

How do these standards affect my assessment?

State tests are being developed to reflect these new standards, and teachers must be prepared to mimic the higher level questions that will be asked on their own classroom assessments. This alignment with the state tests will help teachers better predict student performance not only on these tests, but also in the higher grade levels and in college.

How can I align my assessments to the core?

- Find the appropriate grade level and content strand of the Common Core:
  http://www.corestandards.org/the-standards
- Make sure each test item addresses a specific standard
- Ensure that the verbs on your test match the advanced level of thinking and processing required by the standard
- Use verbs that are also associated with Bloom’s Taxonomy
- Ask for help! Speak to coaches, Common Core representatives and specialists, and reference the Common Core website to ensure alignment.
The first example asks the student to classify the objects into two different categories: big and small, which is a perfect match for what the standard requires.

The second example does not address this specific standard of classification. It is asking the students to compare the two shapes based on size and to identify the larger shape, which would better match Kindergarten standard K.MD.A.2: **Directly compare two objects with a measurable attribute in common, to see which object has “more of”/“less of” the attribute.**

If test items do not assess the standard at hand, they will not provide students with accurate nor useful data.
Culturally Sensitive Assessments

What are culturally sensitive assessments?
Culturally sensitive assessments acknowledge and accept student differences, providing participants with an opportunity to see their culture reflected in test questions so that students of one culture or another are not more likely to perform higher on any given assessment due to its format or language choice.

Writing Directions
Clear test directions are essential in setting up students for success. If directions are confusing or overwhelming, students may not be prepared to answer the questions. Directions should be short and direct. Directions and prompts must include a verb that tells students exactly what is expected of them. If the directions are vague, the responses will be varied, and leaving room for student-made assumptions. If the directions are precise, the responses will be more streamlined, making the work of both the teacher and student focused on the same topic and representation.

Unbiased Language
Diverse student populations must not suffer from assessments written with the dominant culture in mind. To ensure that there is equity in assessments, it is necessary that teachers take care with the language used in directions and test items. It is the job of the teacher to consider his or her student population: their socioeconomic status, their race and culture, and their age and interests. Using these characteristics, teachers can create tests with more familiar language, eliminating unnecessary obstacles and ensuring a focus on content.

<table>
<thead>
<tr>
<th>Do</th>
<th>Do NOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Use proper nouns that the students are familiar with:</td>
<td>✓ Use high level English/Spanish terminology in a content test <em>unless</em> it is content specific vocabulary</td>
</tr>
<tr>
<td>▪ local stores</td>
<td>✓ Use foreign proper nouns that the students have never heard</td>
</tr>
<tr>
<td>▪ names of students in the classroom</td>
<td>✓ Use culture specific terms, unless they have been explained</td>
</tr>
<tr>
<td>▪ age and culture appropriate TV show characters &amp; celebrities</td>
<td>✓ Assume that students will understand words that you have not taught</td>
</tr>
<tr>
<td>✓ Use content specific vocabulary that has already been introduced</td>
<td>✓ Create unfamiliar contexts for reading selections, essay prompts, or other test items</td>
</tr>
<tr>
<td>✓ Eliminate excess wording</td>
<td></td>
</tr>
<tr>
<td>✓ Consider students’ daily lives and experiences in creating reading sections or other test items</td>
<td></td>
</tr>
</tbody>
</table>
In addition to writing directions and test items that are culturally responsive, teachers must be able to verify that all items relate directly to the Common Core State Standards. To do so, teachers must be able to use the verbs from Common Core State Standards and the verbs from Bloom’s Taxonomy. All students, regardless of race, ethnicity, or learning style, must be judged based on these high level standards. In trying to make testing language familiar and simple in order to avoid confusion, teachers must maintain the rigor that is required by the Common Core.

**Examples:**

Target standard to be assessed for Grade 2: 2.OA.A.1 Use addition and subtraction within 100 to solve one- and two-step word problems involving situations of adding to, taking from, putting together, taking apart, and comparing, with unknowns in all positions, e.g., by using drawings and equations with a symbol for the unknown number to represent the unknown.

- **Correct Example:**
  
  There are 27 students in 2nd grade and 31 students in 3rd grade. How many students are there all together?
  
  Show your work.

- **Wrong Example:**
  
  Monique and Xian went fishing at the lagoon. Monique caught 27 perch with her fishing rod and Xian caught 31 perch using his net. How many perch did they catch all together?
  
  Show your work.
Both of these examples accurately assess the Grade 2 standard mentioned above. The prompts require the same amount of mathematical rigor, however, they are unequal in the amount of English language arts rigor required. The first example uses familiar language to all students and provides a situation that all students can easily relate to and imagine. The simple language and relatable topic help keep the prompt’s focus on the math.

Although the numbers are identical in the second example, there are many factors that may cause unnecessary confusion to a diverse group of students: unfamiliar names (unless they are names of students in the classroom), unfamiliar vocabulary (lagoon, perch, rod) and unfamiliar situations (fishing). Although some aspects of this item may be relatable for some of the students in a bilingual/diverse classroom, they will not be familiar to all students, unless previously introduced by the teacher. Considering that all students share a very similar school experience, it is a safe idea to create questions referring to such situations and using school vocabulary: classroom, lunch room, field trips, school supplies, etc.
Creating Multiple-Choice Questions

What makes a high quality multiple-choice question?

Multiple-choice questions require two parts: a stem and answer alternatives. The language and vocabulary should be content and grade-level appropriate. There must be only one correct answer. Multiple-choice questions require a great deal of planning to ensure that they are written in a way that is truly testing if the students have mastered the content standard being assessed.

These questions should be designed in such a way that students who have a mastery of the content will perform well; they should not allow “savvy test takers” without mastery to perform well.

Each multiple-choice question should address a specific Common Core State Standard. Begin with the standard being assessed and build the questions from there. The questions should range in difficulty, beginning with lower level verbs from Bloom’s Taxonomy and moving upwards to help demonstrate a more in depth understanding of each standard. If possible, use the verb explicitly used in the corresponding CCSS.

How can I check that my multiple-choice questions are of high quality?

☐ My stem has one essential question/focus: addressing a specific Common Core State Standard
☐ The language is clear and concise
☐ My stem includes a verb which helps present the problem or question
☐ There is no irrelevant information in the stem
☐ Key words (and all negative words) in the stem, such as best, but, not, and except are in bold or italics
☐ I do not use any double negatives
☐ I have avoided using words such as never or all, which tend to stand out as obvious distractors
☐ I have avoided the choices “all of the above” and “none of the above” as they have been proven to reduce effectiveness of test items
☐ The stem is complete and includes all words that would otherwise have to be written in each alternative. The stem includes more information so that the alternatives can be as short as possible.

☐ There is one correct choice and there are three to four tempting distractors; if I could not create any more plausible distractors, I have only included the high quality options, meaning not all questions will have the same number of choices

☐ Each alternative is approximately the same length

☐ There are no words from the stem in the alternatives that would give away the answer

☐ The alternatives are other plausible options, perhaps likely misunderstandings or mistakes of the students

☐ The alternatives are all alike in terms of content

☐ Grammar links between the stem and alternatives do not lead the student to the correct option

☐ There is random placement of the answers throughout the test (no pattern exists in the answers A, B, C, D, E)

  ○ *To help prevent unconscious patterns in answer placement, randomize the choices using your computer program.

☐ Throughout the tests, the questions increase in difficulty and the verbs move along Bloom's Taxonomy

(Burton, Sudweeks, Merrill & Wood, 1991; Jacobs, 2004; Rudner & Schafer, 2002; Salend, 2011; Torres, Lopes, Babo, & Azevedo, 2011)
The first example has a complete stem that has 1 focus: to find the subject-verb agreement. There are four alternatives, each of which has a similar structure and length. There is no irrelevant information, and all of the choices are plausible, since they are all variations of the verb “to be”.

The second example contains two alternatives that stand out for negative reasons. Choice “b” is much longer than the other options, which can draw attention to it as the correct answer. Choice “d” draws immediate attention because it is the only alternative that does not use the verb “to be” and it may immediately stand out as a distractor.
The primary purpose of the stem of a plant is to...
- a. take in water and nutrients
- b. provide support for the plant
- c. make food for the plant
- d. none of the above

OR

The primary purpose of the stem of a plant is to...
- a. take in water and nutrients
- b. provide support for the plant
- c. make food for the plant
- d. produce fruit containing seeds

**EXAMPLES:**

Target standard for Grade 4, NYS Standard 4, 3.1b
The Living Environment Each plant has different structures that serve different functions in growth, survival, and reproduction.

The poor example uses “none of the above” which is not often used as a plausible distractor (except in the case of mathematical calculations). Students generally recognize this option immediately and discount it as an option. To make a stronger test item, it is best to choose a more plausible distractor or to eliminate the option completely. The good example shows both of these stronger formats. It is unnecessary for each test item to have four distractors, but if the teacher can find a fourth distractor that is both content related and tempting, it can also be included.
Writing Essay Prompts

What makes a high-quality essay prompt?
The best essay prompts for a bilingual population are clear and complete and do all of the following:
1. Give the student a specific role as the writer; are they a friend, a student, a disgruntled protestor, an endangered animal, or a scientist?
2. Indicate the student’s audience; is the student writing to a friend, a teacher, the president, or a community?
3. Provide a format that the student should use when writing: a letter, a school report, a speech, a science experiment, etc.
4. Give a clear topic: what is the purpose for writing this piece, and what information needs to be included?

What is the difference between a content-focused essay and a language-focused essay?
A content essay is used on tests that measure content: mathematics, science, and social studies. A teacher must first consider if an essay is the best format for assessing a specific content standard. If the teacher is measuring a student’s ability to write a scientific experiment, or explain his or her steps in solving a math problem, or describing the differences between urban and rural communities, the answer may be yes. In writing an essay prompt to assess content, one must be sure that neither the directions nor the literacy abilities of the student affect the student’s ability to show his or her understanding of the content. In some cases, therefore, it may be inappropriate to ask specific students to write content-focused essays. As addressed in the “Flexibility in Administration” section of this manual, other formats may more accurately show student knowledge.

Language-focused essays intend to measure a student’s writing abilities in a specific language. These essays may focus on organization of thoughts and ideas, vocabulary use, grammar and mechanics, and writing style. These types of essays will inherently favor native speakers of that language in ease of creation of thoughts and ideas in that language. However, providing a choice in writing topic is one simple way to help reduce bias and provide all students with an opportunity to demonstrate what they are able to do in writing. In this way, students will be able to find topics to which they are able to relate, about which they can form ideas and opinions. Providing this opportunity will help allow students to perform their best. (Writing and grading, 2012)

Essay questions should only be used to assess standards that require more than basic knowledge of facts. These prompts must address a specific standard. If they represent more than one standard, the parts of the prompt must be clearly separated and defined so that students know exactly what is being asked of them.
How can I check that my essay prompts are of high quality?

☐ The role of the writer is identified
☐ The audience is defined
☐ The required format is clearly presented
☐ The topic is made clear and includes all points that I expect the student to address
☐ I have provided a reasonable task that students should be able to complete within the allotted time
☐ A time and length limit are established to help guide students
☐ The prompts are separated into all of the components that the students must address in their answers
☐ The prompt requires students to reach to more than one level of Bloom’s Taxonomy
☐ The content prompt provides a detailed outline of what the essay should look like and contain so as not to base the score on format or writing style

(Jacobs, 2004; McMillan, 2007; Reiner, Bothell, Sudweeks, & Wood, 2002; Salend, 2011; Writing and grading, 2012)

Note: This research recommends special care with essay questions when working with a bilingual/diverse population. Although content prompts should not provide choice, literary prompts can have more flexibility. If the goal of the teacher is to assess a student’s use of language and organization, there is no proven harm in having flexibility in the writing topic. The teacher must use his or her judgment and unique student profiles to help create appropriate essay topics.

For example, in a city school district, it would be meaningless to ask students to write about their experience on the farm, unless they had previously been given this background knowledge. For Hispanic students, it is unnecessary to prompt them to write about an American football game if the purpose is purely literary; instead, ask students to write about their experience as a spectator at a sport of their choice. The guidelines for the rest of the prompt can be specific and detailed, but allowing the choice in sport suddenly levels the playing field for students who may never have been to a football game.
**EXAMPLES:**

Target standard for Grade 6 Writing: *W.6.1 Write arguments to support claims with clear reasons and relevant evidence.*

Write a speech that explains why you would make a good class president.

---

Leave yourself at least 25 minutes for this section of the test.

You are running for class president. Write a speech to give to your classmates that explains why you are a good candidate. State your argument and provide supporting evidence. Be sure to include

- Why you are running
- What qualities of yours make you a good candidate
- How you have already proven yourself as a good candidate

The first example is incomplete; it does not provide students with their audience, or with a clear topic. The prompt is vague and will therefore elicit an immense assortment of responses. The second response includes all four of the essential prompt pieces: role, audience, format, and topic. The topic is clearly outlined in a manageable format that allows students to check off each piece as they attend to that task.
Flexibility in Test Administration

What does it mean to be flexible when administering an assessment?

When dealing with a bilingual population, there is bound to be great diversity amongst the students. Knowing this, teachers must also be aware that despite all attempts to create a fair, unbiased, culturally sensitive, well-crafted assessment, there may still arise issues in test administration. These administration concerns require teacher attention and expertise to both notice and address any hurdles to student performance as they surface.

What kinds of accommodations can I make without compromising the integrity of the assessment?

An accommodation given to a student of a different culture, language, learning style, etc. should help level the playing field for that student. This means that when given to the child, the adjustment will help raise his or her score. It also means that if these same accommodations were given to students who did not need the modifications, they should not help raise their scores.

Examples of research-based accommodations:
• Reading directions orally to the student
• Providing choice in small ways (as described in the essay section)
• Translating test directions or content test questions
• Providing test-specific glossaries (that may include translations of specific words, or definitions of non-content terms)

The most important goal in assessing a diverse student body is to:
create and provide opportunities for all students to demonstrate their knowledge and understanding of the same high standards, through attention to student differences and alignment to the Common Core State Standards.
Checking-in After the Test

**What are some signs that my test (or parts of my test) needs improvement?**

- Certain items were consistently answered incorrectly (possibly by a specific gender or cultural group)
- Certain items were answered correctly by ALL students (even the lowest performing students)

Finding trends in performance on test items helps teachers to examine their assessment based on their own student data. If teachers find that there is a correlation between students’ language, culture, and diversity with their performance on test questions, it is time to examine those specific items for readability, cultural sensitivity, language use, and question format.

**What are some signs that my test items are high quality and culturally sensitive to my students?**

- The class distribution was as expected (per other assessments, observations, etc)
- Students scores did *not* correlate with their culture, race, or background
- Student scores on content tests did *not* correlate with the students’ dominant language

If a teacher finds that the scores were as expected and that there is no correlation between student cultural diversity and their ability to correctly answer test items, it is time to celebrate success!

🌟 Test performance should consistently reflect a student’s mastery (or lack thereof) of specific standards in content and language arts; likewise, performance should *not* reflect his or her inclusion in any specific cultural or socioeconomic group.