ASSESSMENT

Formative Assessment: What We Don't Learn from "Just Answers"

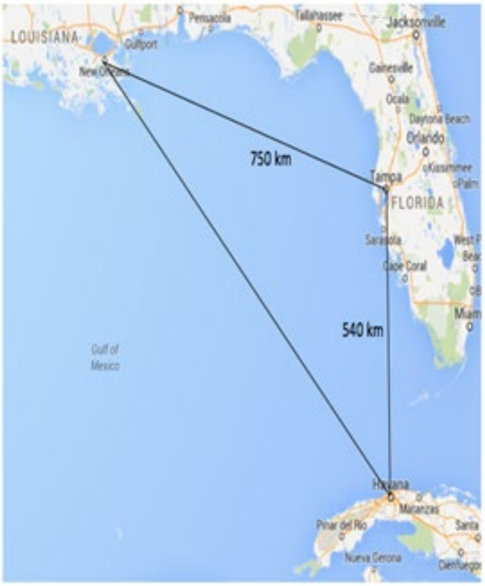
Shelbi K. Cole, Student Achievement Partners

High stakes summative assessments often use test questions that are proxies of the expectations outlined in standards as evidence of students' proficiency in mathematics. While these proxies can provide valuable information as part of a snapshot of student learning in the context of summative assessment, this value does not always neatly translate to classrooms in support of formative assessment. In 2006, the Council of Chief State School Officers (CCSSO) Formative Assessment for Students and Teachers State Collaborative on Assessment and Student Standards (FAST SCASS) defined formative assessment as "…a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievements of intended instructional outcomes." Defined this way, we should make a clear distinction between "testing" and the broader term, "assessment." We might even make the claim that there is no plural form of "formative assessment." The moment you hear the plural "formative assessments," it implies that the phrase is no longer defined by a process, but rather by more frequent testing. This article was written using the definition of formative assessment offered by the FAST SCASS, and will focus on a single mathematical modeling problem to highlight some important differences between formative assessment and other types of assessment.

Mathematical Modeling Problem and Student Responses

Consider the mathematical modeling problem shown in Figure 1. The problem was originally developed as part of an effort to illustrate the standards in the Smarter Balanced Assessment Consortium's[1] Item Specifications document, which now includes over one thousand examples of mathematics summative/interim assessment questions for grades 3–11. To determine whether to include the item as shown (i.e., with no explanation required), the problem was administered to eighty-five grades 4 and 5 students. The version administered to students required an explanation, but the version presented in Figure 1, intended for the summative assessment, did not require an explanation. The purpose of the small scale administration was to evaluate the information, or evidence of student learning, that might be lost by not asking students to write about the mathematics that they were using to provide a reasoned estimate in the problem. The alignment of the problem to the Common Core State Standards for Mathematics, also shown in Figure 1, illustrates that mathematical modeling problems often ask students to apply skills that they have developed over multiple years of learning. The ability to compare the lengths of three line segments is an expectation of the grade 1 standards, while the actual comparison given the specific measurements provided in the problem raises the problem to about grade 4. The problem was also given to grade 5 students to ensure that the 4th grade standards could be classified as securely held content for at least part of the sample.

1 Smarter Balanced Assessment Consortium (SBAC) is the counterpart to PARCC. For futher information, go to: http://www.smarterbalanced.org/smarter-balanced-assessments/



The distance between New Orleans and Tampa is about 750 kilometers. The distance between Tampa and Havana is about 540 kilometers. Estimate how far it is between New Orleans and Havana.

Alignment to CCSS-M:
MP 4. Model with mathematics.
MP 1. Make sense of problems and persevere in solving them.
MP 2. Reason abstractly and quantitatively.
1.MD.A.1 Order three objects by length; compare the lengths of two objects indirectly by using a third object.
4.MD.A. Solve problems involving measurement and conversion of measurements from a larger unit to a smaller unit.
4.NBT.A. Generalize place value understanding for multi-digit whole numbers.

Figure 1. Tampa to Havana problem and Common Core alignment.

First, let's examine five student responses as they would appear absent of any explanation of their thinking: 720, 800, 870, 1000, 1290. Table 1 provides some potential inferences that a teacher might make given just these answers to the problem.

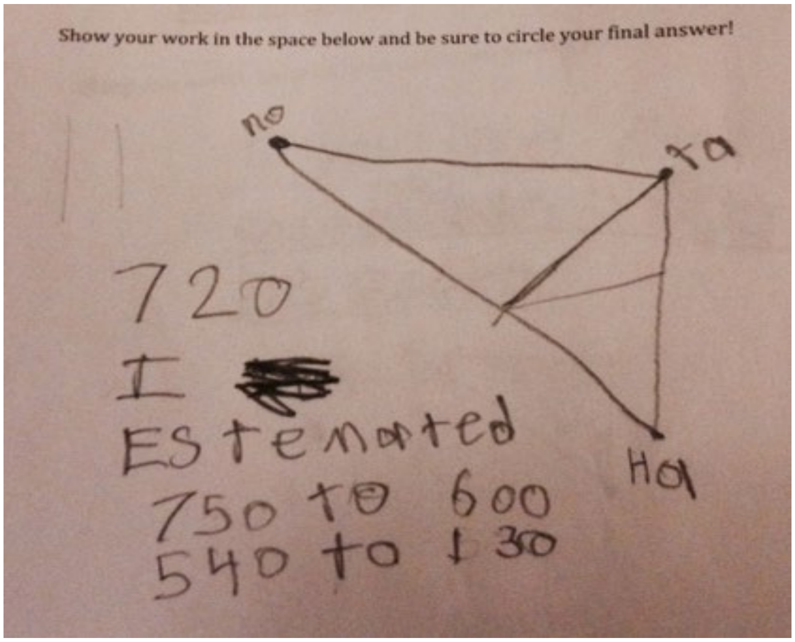
Table 1
Answers without Explanations and Potential Inferences

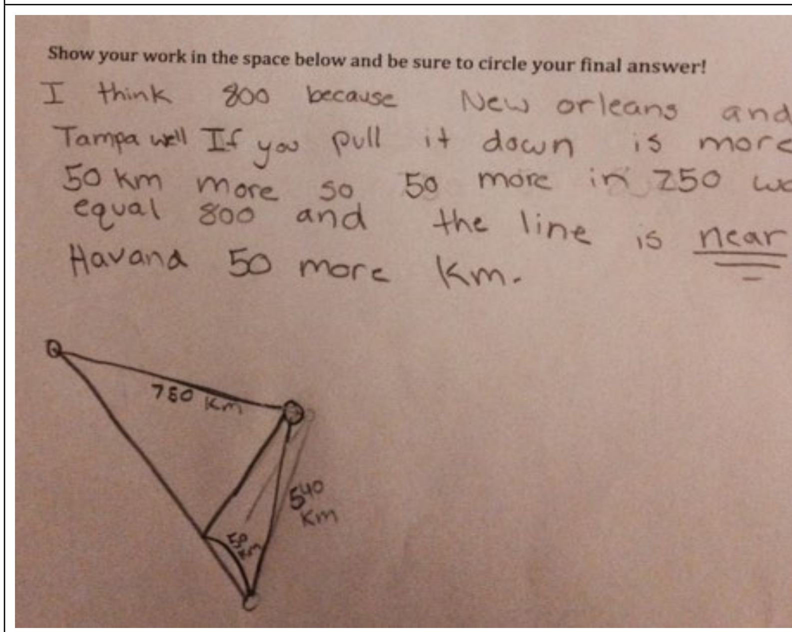
Answer	Potential Inferences Based on Answer
720	Student doesn't recognize that the unknown side length is longer than 750 km.
800	Student recognizes that the length is more than 750 km and less than 1290 km, but underestimates a bit.
870	Student is within a reasonable range for grades 4-5.
1000	Student is within a reasonable range for grades 4-5.
1290	Student seems to have just added the two numbers. He/she may have applied a methodology that works with other problems that "look" like this one, rather than making sense of the problem and given information in the context of the problem.

Now, let's examine the actual student work on which these responses are based. In each case, there is more that can be inferred about student understanding from the work students showed than could be gleaned from the answers alone.

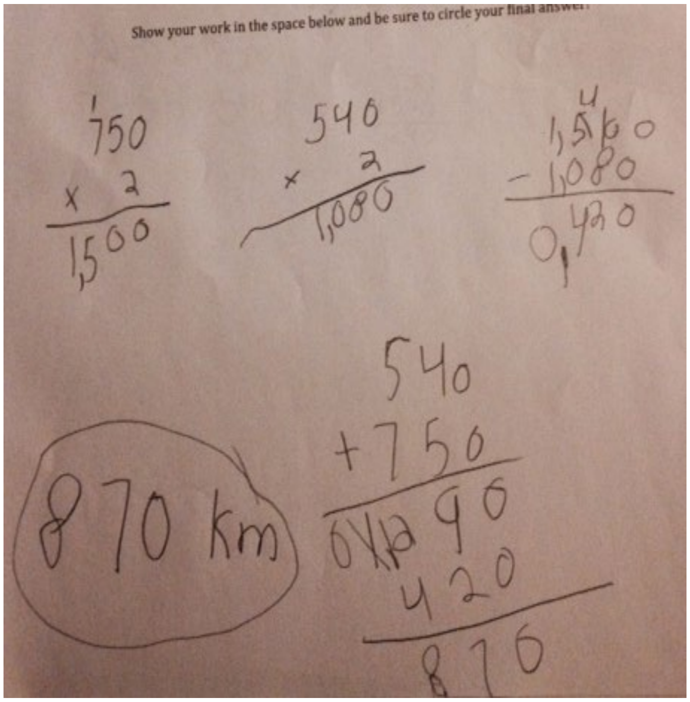
Table 2

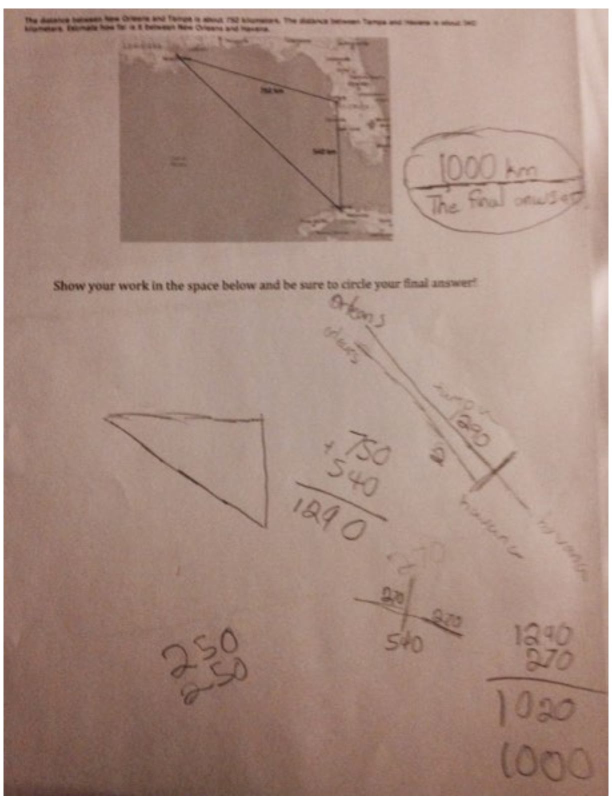
Student Work and Potential Teacher Inferences for Problem from Figure 1

<table>
<tr><td>Answer</td><td>Student Work, Teacher Inferences and Follow Up</td></tr>
<tr><td>

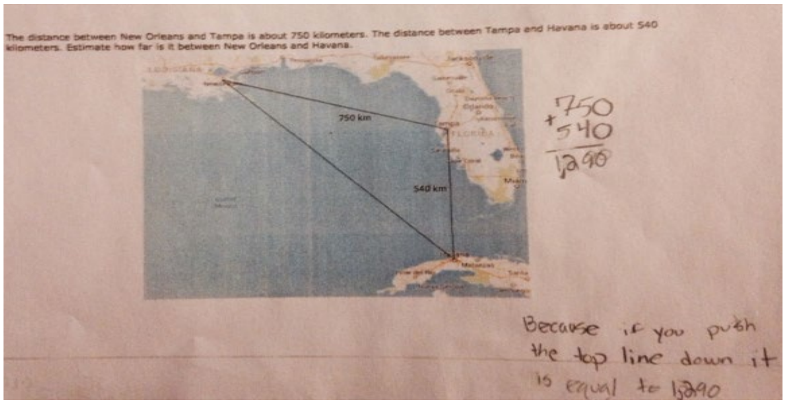
720</td><td>We see from the student's drawing that he dropped a perpendicular to partition the longest side. He then estimated the length of each segment of the partition, but provided estimates that were too low and then an incorrect sum of the segments. The student appears to have deployed a potentially successful (even sophisticated) strategy, but did not assess the reasonableness of his answer. A teacher may want to follow up with this student to get more information about the decisions he made while solving the problem.</td></tr>
<tr><td>

800</td><td>This student creates an isosceles triangle to replicate the 750 km length on the unknown side. She then provides an estimate of the remaining portion of the side length, which falls a bit short of being "close." This is exacerbated by the student's drawing not mirroring closely the original. This student appears to have a sophisticated understanding of the mathematics computations and strategies needed for this problem. The teacher may want to follow up to help the student find ways to get a more precise estimate of the 50 km.</td></tr>
</table>

<table>
<tr>
<td>

870</td>
<td>The mathematics seems to be an arbitrary collection of computations that lead to what could be considered a reasonable estimate of the unknown length. Does this student have a sense of what the "answer should be," but doesn't know how to articulate it or is something else going on here? The teacher could follow up with questions to find out more from the student about his thinking during each step.</td>
</tr>
<tr>
<td>

1000</td>
<td>This student sums 750 and 540, then appears to recognize that the length is less than that sum so divides the 540 in half, getting 270. He then subtracts the 270 from the "too big" 1290 and gets 1020. At the end, the student rounds this to 1000. After a mathematically sound process and reasonable estimate, the final step may represent some confusion between estimation and rounding.</td>
</tr>
</table>

 1290	Although this student does what many others did (simply add the two numbers), she adds a note at the bottom of her paper that suggests that there is a bit more going on here. Her note, "Because if you push the top line down it is equal to 1,290" suggests that she may be visualizing "straightening out" the shorter two segments into a single segment. The teacher might offer physical objects, such as spaghetti, that the student could use to represent the 750 km and 540 km lengths to help her model pushing the "top line down."

From Individual to Group Information

While we can learn a lot about individual students from their mathematical explanations, we can also use item level classification analyses based on groups of students to make inferences about changes to instruction and/or curriculum. If we take the same problem as shown above and examine the group results classified based on students' approaches to the problem, we begin to get a picture of general trends in the kinds of strategies students are using and a sense of some of the more global misconceptions that students have about the mathematics.

Table 3
Classification of Student Strategies at the Group Level

Strategy	Student Response Classification	Number of Students (n=85)
1	Subtracted to find distance	5
2	Subtracted to find distance after rounding both numbers	2
3	Added to find the distance (including responses with minor computation errors)	33
4	Added to find distance and then rounded sum	9

5	Rounded both numbers to the nearest hundred and then added to find distance	11
6	Estimated distance at greater than 750 and less than 1290 using valid mathematical strategies	14
7	Used other not valid computational strategies (e.g., multiplied 750 and 540)	8
8	Used a potentially viable strategy, but estimated less than 750	2
9	Estimated distance at greater than 750 and less than 1290 using mathematical strategies that do not appear valid	1

At least two of the classified strategies stand out as potential curricular/instructional issues. First, 33 out of 85 students simply added the two numbers shown in the problem. This implies that a large number of students seem to dive into the problem with a familiar strategy without making sense of the problem. These students may need more work on Standards for Mathematical Practice MP1: Make sense of problems and persevere in solving them. Students may be accustomed to solving perimeter problems where all of the side lengths are given, and may be inclined to use the strategy: add up the numbers you see. Unfortunately, that strategy does not work here and this problem has both an unknown perimeter and an unknown side length, which takes it from a more general problem solving question into the mathematical modeling category. Students need opportunities to grapple with problems where reasonable estimates are required as part of the problem solving process. In early grades, students should be presented with questions like, "What else do I need to know to solve this problem?" In a growing arc of sophistication with mathematical modeling and as students progress through mathematics, they should encounter more and more problems in which making assumptions, weeding out extraneous information, and retrieving information from external resources are expectations of the problems they are solving.

Another notable trend looking at the table is the number of students who thought that rounding was an essential component of the problem solving process. Strategies 2, 4, and 5 indicate that at least 22 students believed that rounding was an important component of this problem. In this case, teachers may want to evaluate whether "rounding" and "estimation" are being used synonymously either during instruction or in the school's curriculum resources, or whether students are only being presented with problems that ask them to round, and not being asked to think more broadly about the verb "estimate." This problem highlights an important use of the word "estimate," where it is asking students to provide an estimate of a distance within a reasonable range of the true distance from Havana to New Orleans. A reasonable expectation based on the grades 1 and 4 content standards in conjunction with the Standards for Mathematical Practice would be that students could recognize that the distance is greater than 750 km and less than 1290 km.

Conclusion

Ultimately, content experts representing the states in the Smarter Balanced Assessment Consortium decided not to include the problem in the Grades 3-5 Item Specifications. Since the grades 3-5 computer adaptive test does not currently include problems that require explanations (although the performance task section does), the content experts felt that too much information was lost from this particular problem without the student explanations, including some students who would have gotten the problem correct without the required mathematical understandings and other students who would have gotten the problem incorrect who seemed to have sophisticated mathematical modeling strategies.

Although this problem did not make the cut as a proxy for student performance within the summative assessment, it is useful in highlighting the difference between students giving just answers and students providing mathematical explanations of their thinking. This problem and student responses illustrate the need to continue to require students to demonstrate their mathematical thinking in writing, even when summative assessments use other proxies for mathematical thinking. The formative value of the work that students do provides meaningful information about individual student understanding of the mathematics, as well as group level information that may highlight the need to revisit and modify curriculum and/or instruction.

We often treat the terms "testing" and "assessment" as synonymous. It is important to recognize that assessment is a much broader term that encompasses all of the activities that educators use to understand student learning and processes that allow students to understand their own learning. Establishing a purpose first and foremost for the things that we ask students to do, can help us ensure that what they produce will lead to useful information that drives teaching and learning.

References

Council of Chief State School Officers. (2006). Distinguishing formative assessment from other educational assessment labels. Retrieved from http://www.ccsso.org/Documents/FASTLabels.pdf