

## Friday Focus: “Variety of Assessment Methods”

(Standard I, Indicator I-B-1)

Proficient: Designs and administers a variety of informal and formal methods and assessments, including common interim assessments, to measure each student’s learning, growth, and progress toward achieving state/local standards.

Exemplary: *Uses an integrated, comprehensive system* of informal and formal assessments, including common interim assessments, to measure student learning, growth, and progress toward achieving state/local standards. *Is able to model this element.*

Have reread the first line of the Exemplary descriptor several times, encountering the same problem over and over again. My problem is not with the *idea* of “an integrated, comprehensive system of formal and informal assessments, including...etc.” My problem arises when I try to define and describe what that looks like.

Let me start with the easy ones. Formal and informal assessments. Every teacher in this building can easily list a variety of formal and informal assessments that inform his or her instruction. The formal ones are district-expected, sometimes district-determined measures, that give us our measure, year to year. MCAS. WrAP. Stanford 10/OLSAT’s. One important characteristic of the “formal” assessment is that there is little or no wiggle room for altering the administration procedures. Once you do that, the resulting score cannot provide the perspective of comparison. The bell curve goes out the window. Arguably the most important score to come out of MCAS – the individual Student Growth Profile (SGP) – tells us to what extent the student made an expected amount of growth. The number may be heartening (90<sup>th</sup> percentile! Yay!) or discouraging (30<sup>th</sup> percentile. Yikes!) and both of these numbers can accompany the same Raw Score, Standardized Score and Performance Score. But none of these numbers can tell a teacher where the weakness is and, therefore, where to focus future instruction. (Fortunately, the MCAS does offer a breakdown by Standards. So that’s something.)

The Stanford 10/OLSAT – that seemingly endless series of subtest upon subtest – has one unique feature. The achievement tests are not expected to be reflective of the Standards at all – in fact – at this point it seems archaic to even use a test booklet filled with multiple choice test questions! The important part about this test is the AAC – the Aptitude Achievement Correlate – which tells us whether or not the students is achieving up to his or her “potential” according to what the aptitude scores would predict. Interesting to know<sup>1</sup> and it helps us to identify students who are underperforming so that we can offer them quantifiable proof that they need to apply themselves more. The OLSAT, administered in grades 2 and 5, is the only measurement we use

---

<sup>1</sup> If you believe there is such a thing as a measurable aptitude

that provides data on aptitude and, as such, is an important piece in any gifted and talented referral. The achievement subtests are a necessary part of the AAC equation, but they are redundant on their own because we use curriculum based measures of achievement for reading, math and science all the time.<sup>2</sup>

Our formative assessments (Words Way Inventory, Reading Street Comprehension, etc.) provide valuable information regarding individual and collective student performance and are more easily used to influence instruction than the summative. We use formative to know where to go next with instruction, we use summative to measure progress – with no promise of using the data to tell us anything else but “have we (and our curricula, planning, programming, materials selection) done a good job?” Not to say there can’t be overlap – increases in raw scores throughout the year (such as a midyear and End of the Year Words Their Way inventory) do measure progress as well as shape instructional planning. Rubrics can also serve both purposes if the same assessment tool is used over time.

What of all the time and effort devoted to Galileo? Galileo presents that elusive ideal blend of formative, summative, and criteria referenced. It is summative in that it goes outside our curriculum to tell us how well our students really know the math standards at their grade level. It is formative in that it can help teachers shape instruction around collective and individual strengths and weaknesses, and provides a red flag when students or groups of students have failed to master a standard when it seems other groups similarly situated have done so. It is criterion referenced in that – for those who are patient enough to pore through the resulting matrix – it will tell you exactly which students have mastered which standards. The on-line format is the likely test vehicle of the future. So, not perfect, but helpful in many ways I think.

So – teachers who administer these schoolwide formal and curriculum based assessments with adequate frequency and timeliness could most likely consider themselves at least proficient in this indicator. The whole assessment system is certainly *varied*.

But what constitutes **exemplary** performance in this area? Key words are *comprehensive, integrated*, and – simply enough – *use*.

Comprehensive, to me, means that the assessment(s) includes everything that should be included, if not in one sitting, then by the end of the marking period. It means that when trying to determine an instructional level in order to match appropriate materials, the student has the opportunity to demonstrate performance that may go beyond grade level, i.e. reaching the “ceiling.” Since the lesson unit contains essential concepts and questions, the assessment should also get at the students’ understanding of the essential concepts and questions, not just facts and vocabulary. Students’ spelling skills should be assessed in an authentic writing context, not in

---

<sup>2</sup> Know what’s nice about the Stanford Achievement tests? The Spelling subtest. Which used be consistently average and now is consistently above average. There, I said it.

isolation. These are just a few thoughts about what assessments should do in order for the practice to be considered comprehensive.

Integrated assessment? This idea is less clear to me and I welcome any and all discussion on what it means to others.<sup>3</sup> Some elegant performance assessments I've seen have combined several important components being assessed concurrently, such as writing in various genres, essential concepts, multi-media, projects. But now I'm wondering if, by "integrated," the authors of the teacher rubric mean that the teacher integrates the information from the various assessments into one useful learning profile. At special education team meetings, when reviewing evaluations, each evaluator will offer his or her report and conclusions one at a time, but the real important decision-making won't take place until all assessments have been reviewed so that the team members can integrate them – take all the pieces and put them together in order to understand the child's learning profile. We all know what that process looks and feels like, but how does an individual teacher integrate the variety of comprehensive assessments to make one meaningful profile of each student? I think I need to study the practices of teachers to better understand how this works. A few things that come to mind, however, is that the teacher who is integrating student assessments would have to have some kind of system for at least *keeping track* of all the assessments. Folders or portfolios, digital or paper. If asked, the exemplary teacher would be able to answer the question, "How does this assessment inform *this* [other] assessment?" or "How has your thinking about this student's learning changed over the course of several assessments?" or "What does this student's comparative performance across assessments tell you about his or her learning profile?" These are just a few thoughts – again – I would love to hear from you about how you integrate assessments – combine them to make meaningful conclusions about the student.

Finally, the word *use*. Not design, not administer. Assessments are meaningless unless they are used for some purpose, and since the majority of our assessments are curriculum based, they should be used to improve students' experience with the curriculum. Remediating some relative weakness with *one more* mini-lesson. Flexible grouping. Monitor progress after implementing some MTSS intervention. Understanding individual learning needs for conferencing. Designing pre-assessments so as to offer compacted instruction for students who already demonstrate understanding. Writing that Master Memo. A teacher who was exemplary in this standard, upon being asked "How do you use assessment in your instruction?" would respond, "How much time have you got?"

Finally, how to model this element? There is a place for assessment in Atlas Curriculum Mapping. Documents and ideas can be easily uploaded. Be a role model for your colleagues by completing required assessments in a timely manner and providing data to Sarah, who needs it to integrate and make meaning on behalf of our literacy program. If you have a method for organizing data, share it with colleagues. Share it with me! I could use it!

---

<sup>3</sup> That of course goes for all topics everywhere anytime. But you know what I mean.