Fishing for the Right Assessment Language

Myron Dueck

With clear objectives and cooperative assessments, educators can help students set the hook on their own learning.

Over the past year or so, my son has dragged me into the world of fly-fishing. Now, don't be fooled for one second by the romantic portrayal of this pastime as depicted in the Brad Pitt film *A River Runs Through It*. Fly-fishing is a frustrating pursuit, rife with complexity. The underlying challenge of the sport is that the lure is very light, unlike in most other styles of fishing, and it will not pull the line from the reel when hurled into the water. Instead, fly-fishing uses a delicate combination of a heavy line and water tension to "load" the rod and get the fly out to where the fish are waiting.

Confused yet? I was.

But that's exactly the point I'm trying to make: Fly-fishing has its own language. It wasn't until I took lessons, and learned this language, that the sport started to make sense. This is true for many hobbies—from model trains to quilting. There are a flurry of terms and techniques you need to know to understand the sport and feel empowered in conversations with your fellow hobbyists.

It's also true for assessment in schools. For students to feel empowered in their learning, they must understand the language, purpose, and goals of assessment. They need to understand what they're supposed to learn, and more important, be able to determine whether they actually learned it.

For too much of my teaching career, I was fixated on what I wanted to teach, losing sight of the real goal—what my students were *learning*. Dylan Wiliam (2017) aptly describes assessment as "the bridge between teaching and learning" (p. 52)—since, clearly, not all that is taught is learned. Additionally, assessment is derived from the Latin word *assidere*, which means "to sit beside" (Stefanakis, 2002). Both these ideas suggest we might want to approach assessment not as something done *to* students, but rather *with* them.

We can start by shifting our language to welcome students into the process.

Rewrite Standards for Clarity

From my experience, when teachers are asked, "What's a learning standard?", not everyone has a clear idea. Students are often even more mystified. EdGlossary.org defines *learning standards* as "concise written descriptions of what students are expected to know and be able to do at a specific stage of their education."

Standards, therefore, are essentially a combination of a *verb* (which gives me some direction of what to *do*) and a *noun* (which tells me what I need to *know* or use). Check out the sampling of Next Generation Science Standards and Common Core State Standards below. For our purposes, I've italicized the verbs and underlined the basic nouns.

> *Recognize* the distinguishing features of a sentence. (CCSS ELA 1)

> *Apply* scientific principles to *design* a method for monitoring and minimizing a human impact on the environment. (NGSS MS-ESS3-3)

> *Analyze* how an author's choices concerning how to structure a text, order events within it, and manipulate time create such effects as mystery, tension, or surprise. (CCSS RL 9-10.5)

Before we talk more about standards, let's look closely at the importance of a good verb. While boarding a recent flight, I observed a baggage handler yelling at the passenger ahead of me. Over the roar of nearby jet engines, I could hear her hollering, "Baggage tag! Baggage tag!" The confused passenger stood paralyzed, just looking at her, not clear on what to do. On that noisy tarmac, the helpless passenger was simply in need of a verb. I leaned over and shouted, "Rip it off!" He tore off the baggage tag and immediately appeared relieved.

After delivering a workshop to member teachers of the Alberta Assessment Consortium and working alongside its executive director, Sherry Bennett, I've come to appreciate how incorporating a verb into the standards-based criteria of our rubrics greatly helps students understand what to do. To do a quick audit, I encourage you to find your favorite rubric.

Seriously, go get one. I'll wait. …

OK, now that you've got one, look at the criteria—the items typically found in the left column. If it is an English language arts rubric, there's a good chance the criteria are listed as nouns such as *topic sentence* or *list of sources*. The problem is that these are *things*, and just like "baggage tag," the user is often unsure what to do with these *things*. When I performed this little activity at the American School of Honduras, a 4th grade teacher named Tim Van Ness took one look at his class's rubrics and said, "I never noticed my criteria were all nouns! I can easily fix that!" Before long, Tim had adjusted his rubrics from *things* such as "topic sentence" to "*Write* a topic sentence that introduces the theme of your paragraph."

Science teacher Deborah Allen, from the Roseville School District in California, sent me a fantastic one-page unit plan she uses with her 8th grade students, in which she's not only using verbs to help clarify, but also breaking the components down into pieces easily understood by students.

There are many elements that make Allen's plan easy to understand: a compelling question (How ethical is genetic manipulation?), a phenomenon to investigate (dog breeds), and a clear breakdown of knowledge, reasoning, skill, and product targets, all defined with verbs (What I need to know, What I can do with what I know, What I can demonstrate, and What I can make to show my learning). The assessment plan includes vocabulary words that students should know, statements of learning goals they should master, and a description of the presentation and written response they are expected to do at the end of the unit.

Stop Speaking "Percentages"

As I was breaking down standards for my own students, and looking closely at the verbs driving the learning, I realized that a shift in language might be required. If I was going to assess and report on the extent to which my students would *model, analyze, interpret*, and *evaluate*, some of my traditional evaluative terminology seemed lacking. Perhaps most glaringly, these higher-order verbs, and the assignments and activities that accompanied them, were bringing into question the grading vernacular I had spoken for a long, long time: percentages language.

I'll be the first to admit it: I've used percentage-based grading and reporting for many years, as it seemed to make sense and offered the feeling of precision. But my 21-year relationship with percentages is largely coming to an end. (Thankfully we are parting on good terms, and will likely still see each other outside of the workplace from time to time.)

Earlier in my career, when I was fixated on content, percentages were a good fit for my grading and reporting. If I asked my students to *list* the planets and their characteristics, or *memorize* the 50 states, I could easily and reliably attach a percentage score. But what's being asked of our students today has drifted away from content acquisition and more toward what students can do with that content. Our standards are now driven increasingly by verbs such as *analyze, develop*, and *assess*. Our students are challenged to *create, modify*, and *determine*. I'm not sure about other teachers out there, but if I'm asked to assess how two different students *develop* a model to *generate* data (NGSS MS-ETS 1-4), do I really want to split hairs between Jennifer's model being a 94 percent and Susan's a 95 percent? Can I even tell the difference?

Many people will argue, "But we use percentages all the time!" or "Percentages are part of the 'real world!'" Perhaps that's true for the chance of precipitation or likelihood of survival, but really, how *useful* is this information? Will the difference between a 76 percent and a 77 percent chance of rain alter your picnic plans? In *The Empire Strikes Back*, when C-3PO irritatingly reports Master Luke's chances of survival as 725 to 1, or well under 1 percent, I'm not sure his fellow rebels would've felt much better about 300 to 1, or 100 to 1. Luke was in grave danger by any of these incredibly precise measures.

If you're still married to percentages, try using them more often in the real world and see how it goes. If after a day of snow skiing, you proudly proclaim, "I'm a 74 percent skier!", no one will want to hang out with you. Tell the carpet cleaners that you think the cleanliness of your shag rug is down to 13.4 percent from when it was new, and they'll probably look at you very strangely. The next time you're considering calling in sick, tell your principal you're feeling 42 percent healthy and wait for her reaction. Maybe just try, "I'm not feeling very well," and see if you get a better response.

Thomas Guskey covers this conundrum wonderfully in his article "The Case Against Percentage Grades" (2013). He argues that when using percentages, we suffer from the "illusion of precision" (p. 3). We might triumphantly determine that a student arrived at a 97.3 percent in physics, but would we not quietly admit that if that same student did those same assignments again, our grading discrepancy could just as easily result in a 97.6 or 96.7 percent? Furthermore, Guskey argues that the more categories, or levels, we have to use to report on learning, the greater the chances are that we will misclassify the results of any given example of student work, and thereby erode our reliability.

Lastly, when considering the feedback we provide students, we should not confuse precision with utility. Percentage-based terminology may appear precise, but it's not useful as feedback. Feedback is only useful if it allows me to understand where I'm at and provides me with a direction of what I might do next. Hearing that I got an 83 percent accomplishes neither of these things.

Zones of Misclassification

Imagine you are asked what kind of day you're having, and there are two categories: a good day or a bad day. Normally this would be a quick and easy decision, but perhaps you're unsure of which classification to use to describe your day, as some things have gone well, and others not so much. In trying to decide, your mixed day positions you somewhere between the upper part of bad and the lower part of good.

As you are pressed to decide, you find yourself waffling between the two options. Welcome to *the zone of potential misclassification*—the region between two grading categories where you might make a determination, only to realize later that perhaps you should've chosen another.

The more categories we add in our grading, the more zones of potential misclassification we also add. If we add three more categories to our good day/bad day scale, for example—great day, good day, OK day, bad day, and terrible day—we might gain a little more precision, but we also add three more zones of potential misclassification. Imagine the potential for error when deciding how to classify something into the 100 zones present in a percentage system!

Guskey and Brookhart (2019) argue that the fewer categories in our measuring system, the less chance we have to face zones of potential misclassification. In my province, British Columbia, many school districts have been piloting a new reporting language consisting of only four categories—emerging, developing, proficient, and extending—and only three zones of misclassification (see Figure 1). As a parent with two students in this school system, I'm encouraged about this kind of language, as it is not only more reliable, but it's stated in strength-based, student-friendly language that my own son and daughter tend to understand.

Figure 1. Four-Category Assessment System

■ Zone of misclassification
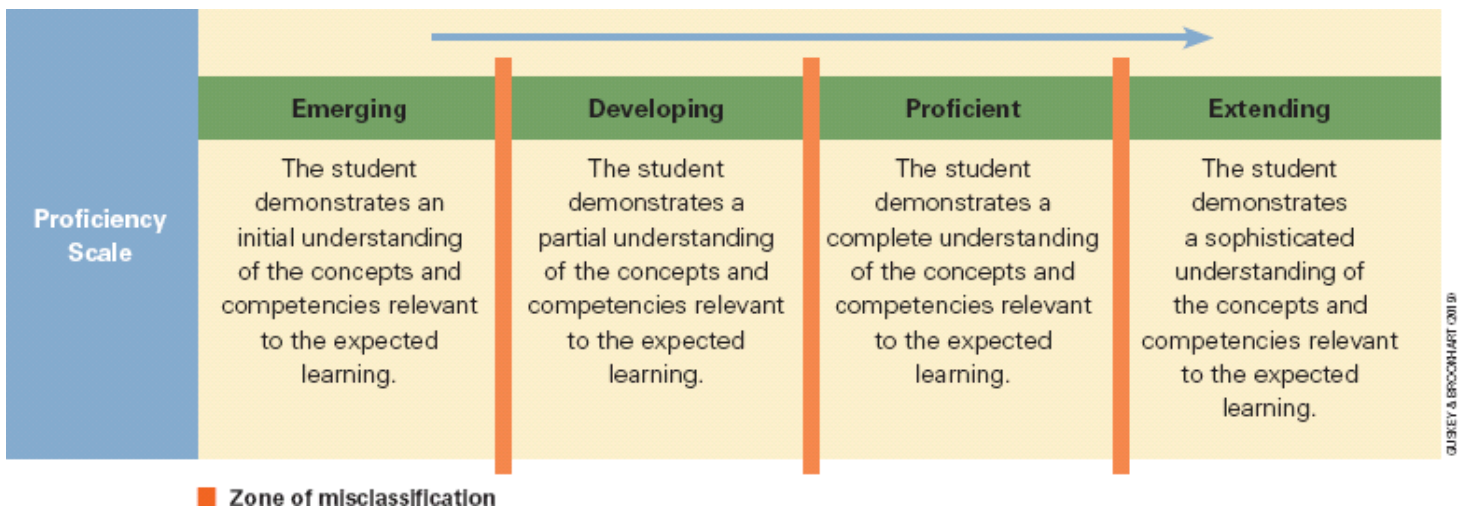
Conversation-Based Grading

Thinking back to my fly-casting lessons, I recall being completely frustrated by my struggles to fling that little fly using the bending of the rod and weight of the line. It didn't help being told by the instructor, "Wow, your son is doing really well. He'll soon be on to more complicated casting." Grrrrrrrr.

Despite my frustration, what I can fondly reflect upon now was the conversational nature of the assessment method. I was never told I was a 48 percent or a 4 out of 10. Rather, I was given feedback such as, "That's really close, but you're missing one little thing—let me show you," or, later on, "That's good, you're nearly there, just try this." I think we can learn something from this type of assessment—the same version likely used by your grandma teaching you to make a pie crust, or by the person who taught you to ride a bike. A descriptive nature to the assessment, along with clear directions on next steps, results in a considerably more useful and empowering experience than simply doling out a percentage score.

A few years ago, I considered reframing how I graded my students and reported their learning, and thankfully I had a few colleagues in my high school ready to explore this change also. We examined John Hattie's research and studied his list of 150 influences on student achievement. To our surprise, *student self-reported grades* was number one (2012). Around this same time, I was a part-time administrator in my high school, and I had many opportunities to chat with students. When asked, "How are you doing in your classes?", far too often the students shrugged their shoulders or commented, "I don't know."

Our team came up with two goals: to increase student self-reporting in our classes and to decide on a language to facilitate this conversation. We wanted our students to speak about their learning similarly to how they talk in everyday life, so we adopted a six-level proficiency scale that had common descriptors for each level. These six levels were used to give feedback on just about any example of student work—from an essay to defining a set of terms on a test. We listened for how students were describing the categories to each other and included this as "student lingo," which was useful because sometimes students have a more effective and efficient way of conveying a message to other students. Our language went through many changes and adaptations and evolved over time into the framework seen in Figure 2.

Figure 2. Six-Level Proficiency Scale

		Teacher Speak	Student Lingo
Expert	6	Near perfect demonstration of understanding/skill; high confidence; mastery of learning standard	"You could teach this!"

☐	5	Strong demonstration of understanding/skill; high confidence; slight error involved	"Almost perfect, just one little error!"
Apprentice ▪	4	Good demonstration of understanding/skill; confidence evident; a few errors	"Good understanding, with just a few errors."
	3	Satisfactory demonstration of understanding/basic skills; key concepts are lacking; errors common	"You are on the right track, but understanding is lacking on a key concept."
Novice ●	2	Minimal understanding of key concepts and rudimentary demonstration of basic skills; many errors	"You have achieved the bare minimum to pass the learning outcome."
	1	Inadequate understanding of key concepts and little to no demonstration of basic skills; errors throughout	"You have not met the minimal proficiency on the learning outcome or the skill."

Dave Van Bergeyk, a calculus teacher in Salmon Arm, British Columbia, worked with our team and showed us how he negotiated the final grades for all his students through individual conversations at the end of the course. After this enlightening meeting, we began to ask students to look for evidence from their own work and to present a case for how they thought they were doing. *Conversation-based grading* was all about us sitting beside the students and letting them talk about their learning. As we further developed this system, we developed a mental model to guide the conversation. The model had a circle in the center with the question, *What is my grade?*, with every type of student assignment that impacts that grade (projects, tests, quizzes, journals, etc.) on the outside pointing to the center. The model reminded students that learning is represented in many ways and that they have multiple chances and avenues to demonstrate understanding.

The results were incredible as we welcomed students to talk about their own learning, and to report on themselves. When I asked one of my students how she felt about reporting on her own learning, she boldly stated, "I love it. No one knows me like I know me!"

Empowered Assessment

I'll never forget the first successful fly-fishing excursion my son and I had. We launched our boat on a remote lake in British Columbia and caught many trout, releasing them after a few photo opportunities. I now realize that my empowerment was directly related to my understanding and comfort around the language of our new hobby. I'm convinced that student empowerment in the realm of assessment and learning also hinges largely on the language we choose. As you break down learning goals, consider fewer categories and welcome a more conversational nature to learning. With those tools, assessment can truly become something we do *with* students rather than *to* them.

REFLECT & DISCUSS

➼ Examine one of your lessons or unit plans. Can you see a way to rewrite it to be clearer about its learning objectives?

➼ Where might you find ways to include students more in the process of evaluating their own work?

➤ What challenges do you see to implementing conversation-based grading? How might one of those challenges be overcome?

References

Guskey, T. (2013). The case against percentage grades. *Educational Leadership, 71*(1), 68–72.

Guskey, T., & Brookhart, S. (2019). Are grades reliable? Lessons from a century of research.*Education Update, 61*(5).

Hattie, J. (2012). *Visible learning for teachers*. Routledge: New York.

Stefanakis, E. (2002). *Multiple intelligences and portfolios*. Portsmouth, NH: Heinemann.

Wiliam, D. (2017). *Embedded formative assessment* (2nd ed.). Bloomington, IN: Solution Tree.

Myron Dueck works with educators around the world on topics related to grading, assessment, and reporting. He is district vice principal for assessment and innovation in SD67, British Columbia, Canada, and the author of *Grading Smarter Not Harder* (ASCD, 2014). He is working on his second book with ASCD focusing on student-empowered assessment. Follow him on Twitter at @myrondueck.