# EL EDUCATIONAL LEADERSHIP

February 2018 | Volume **75** | Number **5**
**Measuring What Matters** Pages 14-20

# Three Key Questions on Measuring Learning

*Jay McTighe*

**To gauge different types of learning, we need a broader collection of measures, with a greater emphasis on authentic, performance-based projects.**

**BUY THIS ISSUE**

Share |

Educators, policy makers, parents, and others interested in improving the way we measure learning in today's schools need to examine three essential questions: 1) What really matters in a contemporary education? 2) How should we assess those things that matter? 3) How might our assessments enhance learning that matters, not just measure it?
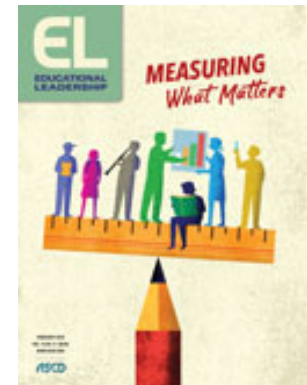
## What Matters in a Contemporary Education?

Any consideration of educational measurement must begin with the desired outcomes to be measured. In our work on the Understanding by Design® framework, the late Grant Wiggins and I described four key types of educational goals—knowledge, basic skills, conceptual understanding, and long-term transfer goals. All of them are essential to a successful education in the 21st century (Wiggins & McTighe, 2011). While these goals are interrelated, their distinctions are important because each type requires different approaches to both teaching and assessment. Let's look at each in turn.

### Knowledge

Knowledge goals specify what students should know—factual information (state capitals, multiplication tables), vocabulary terms, and basic concepts (climate, balance). The attainment of knowledge goals can be best gauged through objective test or quiz items and teacher questioning.

## *Basic Skills*

Skill goals state what students should be able to do. Every subject area contains basic skills (addition, handwriting, drawing, dribbling a basketball) that are essential to building competency and mastery. Teachers can assess student proficiency in a particular skill through direct observation of a performance or by examination of an end product that required use of the skill. Unlike with assessments of knowledge, for which there is usually a single, "correct" answer, skill performances can be best tracked along a continuum of proficiency levels from novice to expert.

## *Understanding*

Understanding goals refer to students' grasp of conceptual "big ideas." Such ideas are inherently abstract. They may be in the form of concepts (patriotism), principles (F=ma), themes (friendship), issues (government regulations), or processes (problem solving). Understanding in this context generally cannot be assessed through multiple-choice or fill-in-the-blank test items. Instead, students need to provide explanations, justify conclusions, and support answers with evidence. As with a doctoral dissertation, there is a need for the *defense*, not just the answer.

## *Long-Term Transfer*

Long-term transfer goals refer to students' capacity to apply what they've learned to a new situation or different context. Transfer goals are process oriented; they specify what we want students to *be able to do with their learning* in the long run when confronted by new opportunities and challenges. They tend to be reflected in the anchor standards or framework practices in official academic standards, but they are often transdisciplinary in nature (encompassing complex skills like critical thinking and collabora-tion, or developmental Habits of Mind such as persistence and self-regulation).

Transfer abilities—or qualities inherent in them—are also increasingly valued in the 21st century workplace, in a way that they haven't needed to be in the past (see fig. 1) (National Association of Colleges and Employers, 2016; Darling-Hammond & Adamson, 2013). Indeed, it's not too much to say that the future belongs to those who can apply their learning effectively in new situations.

## FIGURE 1. Attributes Employers Seek in Job Candidates

| Attribute | Percentage |
|---|---|
| Leadership skills | 80.1% |
| Ability to work in a team | 78.9% |
| Communication skills (written) | 70.2% |
| Problem-solving skills | 70.2% |
| Strong work ethic | 68.9% |
| Communication skills (verbal) | 67.2% |
| Initiative | 65.8% |
| Analytical/quantitative skills | 62.7% |
| Flexibility/adaptability | 60.9% |
| Technical skills | 59.6% |
| Interpersonal skills (relates well to others) | 58.4% |
| Computer skills | 55.3% |
| Detail oriented | 52.8% |
| Organizational ability | 48.4% |
| Friendly/outgoing personality | 35.4% |
| Strategic planning skills | 26.7% |
| Creativity | 33.3% |
| Tactfulness | 18.6% |
| Entrepreneurial skills/risk-taker | 18.6% |

Source: National Association of Colleges and Employers. (2016).
Job Outlook 2016.

Transfer abilities can best be measured through authentic, performance-based tasks, with well-developed rubrics for evaluation.

## How Should We Assess the Things That Matter?

Assessment is a process by which we make inferences about what students know, understand, and can do. To allow valid inferences to be drawn from the results, an assessment must align with, and provide an appropriate measure of, a given goal. Moreover, because all forms of assessment are susceptible to measurement error, our inferences are more dependable when we consider multiple sources of evidence.

Given that there are different types of learning goals, we need an associated variety of assessment types to gather valid evidence of learning. Think of assessment as analogous to photography. Like the results on a test, a picture can be informative; however, no *single* photo can provide a complete portrayal of a situation. To continue the analogy, what we need is a photo album of evidence on student learning, not a snapshot—a collection of multiple measures, appropriately aligned to different types of learning outcomes that matter.

This raises a vital question concerning the alignment between assessments and educational goals: Are we currently assessing everything that matters, or only those things that are easiest to test and grade? With respect to large-scale, standardized assessments, the answer is fairly obvious. For example, virtually all current standards in English language arts include listening and speaking skills, which are generally acknowledged as the foundations of literacy. Yet those skills are rarely, if ever, assessed on standardized tests. Similarly, most standardized tests have limited capacity to assess transfer goals, or related complex skills like scientific investigation, historical inquiry, research, argumentation, and creative thinking.

Now for a related question: If some outcomes that matter are slipping through the cracks of standardized testing, are we doing a better job of measuring all valued outcomes through classroom assessments? Studies of classroom assessments raise doubts (Frey & Schmidt, 2010). For one study, a district collected classroom assessments from all K–12 teachers in all subjects during a six-week period (Gibble, 2000). Of the total of 664 assessments collected, 20 percent were identified via a random sample for analysis by a committee of teachers and administrators. Here were two of their findings: 1) The majority of the assessments (75.5 percent) measured the lowest levels of cognition (levels I and II on Bloom's Taxonomy), and 2) assessment items were predominantly (80 percent) in multiple-choice, true-false, matching, or fill-in-the-blank formats.

Although this study is dated, its findings are consistent with patterns that I have observed more recently as "test prep" pressures result in classroom assessments that mimic the formats (generally selected- and brief-constructed-response) of state and national accountability tests (McTighe, 2017). Ask yourself: What would the results be if you replicated this study in your school or district today? Are any of your valued outcomes not being properly assessed?

Given the limitations of large-scale testing and the status of classroom assessments, what changes do we need to make to ensure that we are assessing outcomes that matter? What assessment photos do we need for a composite album of evidence of learning? Traditional types of assessments offer sufficient ways of measuring students' knowledge and basic skills. For example, we can use multiple-choice or fill-in-the-blank test items to gauge students' knowledge of historical or scientific facts. However, to properly assess conceptual understanding, long-term transfer, and other complex skills,

we need greater use of authentic, performance-based measures in which students are asked to: 1) *apply* their learning to a *new* situation, and 2) *explain* their thinking, show their reasoning, or justify their conclusion.

Authentic tasks are like the game in athletics. While the players have to possess knowledge (the rules) and specific skills (dribbling), playing the game also involves conceptual understanding (game strategies) and transfer (using skills and strategies to advantage in particular game situations). Assessing what matters must include assessing performance in a "game" in addition to tests of requisite knowledge and skills.

Another dimension of making sure we are assessing things that matter shifts the spotlight from educators to students. A photo album approach to assessment can enable learners to contribute their own personal "photos" as evidence of their accomplishments. They can be invited to propose ways of showing that they are meeting academic standards. Students and parents can be asked to contribute evidence of creativity, persistence, or community contributions accomplished outside of the school day. After all, isn't that what transfer means? Maintaining high standards does not require standardization of all measures.

Involving learners in creating the assessment portfolio builds students' capacity for self-assessment. The ability to honestly appraise one's performance against established criteria and performance standards is a life-long skill and a sign of intellectual maturity.

## How Might Assessments Serve Learning?

Which leads us to our third big question: How might assessments become more integral to learning, as opposed to just evaluating it? To say that assessment should serve learning is a nice slogan, but what exactly does it mean? Over the years, Wiggins and I conducted a workshop exercise to explore this question. We asked participants to think of a highly effective learning experience and then identify ways in which the associated assessments contributed to that learning. The responses to the exercise have always been remarkably similar across groups and provide a veritable blueprint to guide teachers' classroom assessment practices. Figure 2 offers a representative list of assessment characteristics that we found contribute to deep and effective learning (McTighe, 2013).

## FIGURE 2. Assessment Practices that Enhance Learning

**Assessments enhance learning when:**

- The "end" learning goals are known in advance, as are the assessments of them.

- The evaluative criteria (such as rubrics) are presented and explained at the beginning.

- Models of excellence, aligned to the criteria, are shown to provide a clear "picture" of desired performance.

- The assessment tasks are set in a realistic context and require application of skills.

- The assessments are challenging, yet attainable.

- The assessments are "open." In other words, there isn't a single correct answer or a single way of accomplishing the task.

- Students produce tangible products and/or performances to show evidence of their learning.

- There is an audience other than the teacher.

- The learners have some choice regarding how they will demonstrate their learning (such as via products and performances) and/or how they go about the task (process).

- There are opportunities to work with others (collaboration).

- The teacher functions like a coach or an advocate— a critical friend, not an adversary.

- Detailed feedback is provided along the way.

- Learners have opportunities to practice, refine, or "re-do" after feedback is given.

- Learners are encouraged to self-assess, reflect on their learning, and set future goals based on the assessment results.

*Source:* McTighe, J. (2013). *Core learning: Assessing what matters most.* Midvale, UT: School Improvement Network.

This list explains why I am an advocate for expanding the use of authentic tasks and projects in schools. Such assessments offer more than just another way to measure student achievement. Like the game in sports or the play in theater, authentic performances are motivating to learners. They give relevance and purpose to learning, and they underscore the need for practice.

Authentic tasks also influence teaching. Coaches recognize that their job is not to simply "cover" the playbook play-by-play and teach individual skills. They understand that knowledge and skills are in service of larger ends, and that their role is to prepare players for authentic transfer performance in the game. Performance-oriented teachers in all subjects understand this role as well.

## Measuring What We Value

How we answer the three questions posed in this article has significant consequences for educational measurement, instructional practices, and ultimately, student learning. The outcomes we choose to measure, as well as the methods of assessment we use, signal to students, parents, and others what matters. If we claim to value critical thinking, creative problem solving, oral communication, and the ability to work effectively in groups, then we need to teach and assess those outcomes. Otherwise, students will quickly get the message that these goals don't really count. If our assessments consist primarily of "single-correct answer" items, we validate rote learning and formulaic responses.

While there is still a place for traditional measures of knowledge and skills in an assessment photo album, greater attention must be given to gathering evidence of authentic student work through performance tasks and projects. By collecting authentic student work samples in digital portfolios, students can compile a literal "album" of growth and evidence of genuine achievements over their school career. As Wiggins (1998) opined many years ago, "Students should graduate with a resume of accomplishments, not just a transcript of 'seat time' and a GPA." It's time to act on this idea.

Of course, form follows function. The integration of multiple assessment measures, including expanded use of authentic, performance-based assessments, requires major shifts in school structures, including in grading practices and scheduling. Teachers will need time to collaborate on assessment task design, within and across subject areas (especially because many authentic tasks can be naturally interdisciplinary). To make the most of performance-based assessments, teachers will also need built-in opportunities to meet in professional learning communities to examine the student work resulting from those tasks (McTighe, 2008). These activities cannot be properly accomplished during a 30-minute team planning block or during an after-school meeting.

But we know the importance of identifying learning goals that matter, of determining ways to assess all of our valued outcomes, and of designing assessments in ways that can be integrated with learning. What's needed in most schools is the will and a systematic plan to do so.

## References

Darling-Hammond, L., & Adamson, F. (2013, June 6). The next-generation of assessments can—and must—be better. *ASCD Express*, *8*(18).

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal*, *5*(3), 107–117.

Gibble, J. (2000). *Report from the office of the supervisor of curriculum and instruction*. Mt. Joy, PA:

Donegal School District.

McTighe, J. (2008, May). Making the most of professional learning communities. *The Learning Principal*, *3*(8), 1, 4–7.

McTighe, J. (2013). *Core learning : Assessing what matters most*. Midvale, UT: School Improvement Network.

McTighe, J. (2017, March 13). Beware the test prep trap [blog post]. *Newsela Blog*. Retrieved from https://blog.newsela.com/blog/2017/7/25/u2tlzoasa0gur7ua7uij7ihtop6dkp

National Association of Colleges and Employers. (2016). *Job Outlook 2016*. Retrieved from www.naceweb.org/s11182015/employers-look-for-in-new-hires.aspx

Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

Wiggins, G., & McTighe, J. (2011). *The understanding by design guide to creating high quality units*. Alexandria, VA: ASCD.

Jay McTighe (www.jaymctighe.com) is an education writer and consultant. He is co-author of the Understanding by Design series with Grant Wiggins. Follow him on Twitter.

## KEYWORDS

Click on keywords to see similar products:
21st century leaning, assessment and grading, formative assessment, instructional strategies, active and project-based learning, student engagement and motivation, whole child: engaged, whole child: challenged, audience: administrators, audience: district-based-administrators, audience: higher-education, audience: new-principals, audience: new-teachers, audience: principals, audience: teacher-leaders, audience: teachers, audience: building-level-specialist, audience: instructional-coaches, audience: superintendents, audience: students, level: k-12

### Requesting Permission

- For **photocopy**, **electronic and online access**, and **republication requests**, go to the Copyright Clearance Center. Enter the periodical title within the "**Get Permission**" search

field.

- To **translate** this article, contact permissions@ascd.org