BUSTING MYTHS ABOUT PROFESSIONAL JUDGMENT AND SUBJECTIVITY IN ASSESSMENT

Removing the human element from testing doesn't make it more accurate.

By Ken O'Connor & Matt Townsley

When Ken was an international field hockey umpire, players would often exclaim that someone had violated a rule. "Why aren't you making this call, ump?" He had to tell these well-intentioned players that the rule they were so sure about wasn't in the rule book at all. These myth-busting conversations became a frequent occurrence during his umpiring career.

We have a similar issues in classroom assessment. Teachers believe, and act upon, myths about their *professional judgment* and *subjectivity* to the detriment of quality classroom assessment.

KEN O'CONNOR (kenoc@aol.com) is an independent consultant who specializes in issues related to the communication of student achievement, especially grading and reporting. He is the author, most recently, of *A Repair Kit for Grading: 15 Fixes for Broken Grades* (3rd ed., First Educational Resources, 2023) and *How to Grade for Learning: K-12* (4th ed., Corwin, 2018). **MATT TOWNSLEY** (matt.townsley@gmail.com) is an associate professor of educational leadership at the University of Northern Iowa in Cedar Falls. He is the co-author of *A Parents' Guide to Grading and Reporting: Being Clear About What Matters* (Rowman & Littlefield, 2023) and author of *Using Grading to Support Learning* (Routledge, 2022).

AT A GLANCE

- Decisions about assessment are often built on myths about teacher professional judgment and subjectivity that prioritize standardized assessment over classroom assessment.
- Allowing teachers to exercise their professional judgment does not mean allowing them to do whatever they want. Instead, they must work within certain guidelines.
- Teachers should be encouraged to review a variety of assessments over time to assess student learning.
- The goal of assessment should not be to eliminate subjectivity, but to create consistency in how teachers assess students.
- Standardized assessments do not provide a more accurate measure of student abilities. In fact, a focus on classroom assessments can solve some of the problems with standardized assessments.

Understanding where these myths come from and how to dispel them starts with a clear understanding of the purpose of classroom assessment.

Understanding classroom assessment

In a widely cited *Kappan* article, Dylan Wiliam and Paul Black (1998) defined assessment as "all those activities undertaken by teachers — and by their students in assessing themselves — that provide information to be used as feedback to modify teaching and learning activities" (p. 140). Indeed, the primary purpose of classroom assessment is to gather evidence to improve student learning — every student's learning. Yet, in the era of data-driven accountability and publisher-provided classroom assessment materials, this purpose can be fraught with misunderstandings and erroneous implementation.

Too often, educators overemphasize externally imposed mathematical formulas to determine classroom assessment cut scores (e.g., "80% is considered 'proficient' on this test for placing students into intervention"). They also depend on long-standing assessment and grading policies (e.g., "90% of the points and better is an A on this test") because they believe these approaches are more objective than their personal appraisal of the students they teach daily.

We believe classroom teachers can lean into their professional judgment to make *better* decisions about student achievement, but it is not always an easy case to make. The thought of teachers' using their professional judgment may elicit feelings of uncertainty ("Is he qualified to do that?" or "It's too subjective") and even a sense of misbehavior ("Is she permitted to do that?") from outsiders.

In our experiences supporting thousands of classroom teachers, we have observed a variety of classroom assessment practices based on myths related to professional judgment and subjectivity. Many of these myths are easily debunked when we look at current scholarly research and best practices.

Professional judgment

Lawyers, medical providers, airline pilots, and other professionals rely on their judgment every day to make informed decisions in their workplace. For example, physicians in the U.S. follow clinical practice guidelines established by professional organizations such as the American Academy of Family Physicians. However, following these guidelines to put their knowledge into practice requires them to use their professional judgment (Mugerauer, 2020).

Teachers must do the same as they review student evidence of learning and plan next steps. Indeed, the Ontario Ministry of Education (2010) contends that for teachers, "professional judgements are at the heart of effective assessment, evaluation, and reporting of student achievement" (p. 8). Yet two common myths have created an atmosphere of suspicion around the idea of assessments that rely on teachers' professional judgment.

Myth 1. Using their professional judgment means every teacher can do whatever they want

One reaction to the suggestion that teachers should use their professional judgment is an assumption that they will throw the rule book out the window, and classroom autonomy will prevail. Teacher A may believe they can assess and grade students in their own way, even if it is significantly different from the way Teacher B does it. In this scenario, using professional judgment permits teachers to make decisions without any shared parameters. However, this is a false understanding of professional judgment.

Damian Cooper (2011) defines professional judgment as "decisions made by educators, in light of experience, and with reference to shared public standards and established polices and guidelines" (p. 13). Adhering to shared public standards and established policies and guidelines improves consistency within and across classrooms, which is necessary for quality classroom assessment. However, over a century of research suggests that teachers have implemented grading practices inconsistently across classrooms (Brookhart et al., 2016). This illustrates the need for shared policies and guidelines to set the parameters within which teachers exercise their judgment. Such systemwide polices support teachers in implementing research-informed assessment practices (Guskey, 2020; O'Connor, 2018).

Classroom assessment and grading practices will be enhanced when teachers follow certain established policies that allow them to exercise some judgment. For example, instead of requiring teachers to use a percentage-based scale, a school could ask them to use four distinct performance descriptors (e.g., beginning, developing, nearly proficient, and proficient). When using a percentage-based scale, teachers

must be able to differentiate among 101 levels of learning, for example, between a student that has an 85% from a student that has an 86% on Pythagorean's Theorem, a nearly impossible task. In contrast, when using fewer decision points, the higher the likelihood of consistent teacher judgements for students on the level of their understanding of Pythagorean's Theorem.

Myth 2. Classroom assessment should be protected from teachers' professional judgment

Because teachers often have been inconsistent in their professional judgment, state, provincial, and district policies have sought ways to insulate classroom assessment from these inconsistent judgments. Usually, this means requiring teachers to use programs and textbook materials that rely on multiple choice, fill in the blank, and other constructed-response assessments.

Unfortunately, these scripted and often standardized materials produce a goal-method mismatch. For example, a seventh-grade learning goal from the Common Core English Language Arts standards requires students to "Demonstrate two or more central ideas in a text and analyze their development over the course of the text; provide an objective summary of the text" (CCSS.ELA-LITERACY.RI.7.2). A multiple-choice item can efficiently ask students to demonstrate two or more central ideas after reading a short passage. However, analyzing the *development of these ideas over the course of a text* and *providing an effective summary of the text* likely require an open-ended response. A standardized commercial assessment may feel efficient and valid, but the information obtained may not accurately represent the full extent of what the student knows in relation to the learning goal.

To bust this myth, schools should require teachers to obtain assessment information through a variety of means, including performance assessments, observations, discussions, questioning, teacher/student conferences, projects, portfolios, tests, essays, and self-assessments. When teachers use a wide range of assessments, they make it more likely that the assessment method will capture students' progress toward the goal (Cooper, 2022).

Another way to bust this myth is for teachers to collect evidence of student achievement *over time* from different sources. By observing students, conversing with them, and reading their written explanations over time, teachers can increase the reliability of their evaluations (Alonzo, 2019). This is most needed when summarizing learning at the end of a unit of study or reporting period. Looking at students' achievement over time will enable teachers to capture how much progress students have made. This way, a low score at the beginning of a unit will not pull down the overall grade. Sadly, summative assessment in the U.S. has often relied on standardized instruments with limited room for teacher judgment (Brookhart, 2016). Instead, determining a report card grade should involve teachers' professional judgment and interpretation of evidence. It should reflect the student's most consistent level of achievement, with special consideration given to more recent evidence.

Subjectivity

In addition to discussion surrounding the merits of teachers' professional judgment, critics and reformers see subjectivity in grading as a problem to be solved (Bowers, 2011; Feldman, 2023; Link & Guskey, 2022). For example, Joseph Kunnath (2017) noted that because teachers' philosophical beliefs often overtly incorporate subjectivity when grading evidence of student learning, schools should consider adopting alternative grading practices. Beliefs such as prioritizing student effort may result in a teacher awarding more points to one student than another student, despite both students answering the same questions correctly on an assignment.

At first blush, it seems like assessment and grading would include opportunities to eradicate subjectivity. For example, the number of correct spelling words or the percentage of multiplication facts correct are objective measures. However, even such seemingly objective measures include some subjective elements (e.g., the difficulty of spelling words and multiplication facts or the role of speed in students' accuracy). Yet certain myths about the dangers of subjectivity and superiority of objectivity persist.

Myth 3. Assessment and grading can be more or less objective or subjective

Educators often believe that subjectivity is bad and that teachers should strive for objectivity when assessing students. What they often fail to recognize that *all* human judgment is subjective, as Joshua Eyler (2024) reminds us:

> Grades rely on criteria that are almost always subjective insofar as they are developed by an individual instructor based on an assignment created by that same teacher who brings to the process her or his own values, beliefs about error and feedback, perspectives on the role of the teacher in helping students learn, and a host of other factors. (pp. 25-26)

In a January 2000 post in an ASCD chat, Grant Wiggins pointed out that the question is not whether assessment and grading practices *are* subjective, but whether they are credible and defensible. Advanced Placement and International Baccalaureate programs are credible and defensible because they go to great lengths to train their examiners to make *consistent* judgments. But their assessments are scored by humans; therefore, they are subjective.

However, this human element in assessment is not something to fear. In fact, as Ruth Sutton (1991) reminds us, it is an inevitable part of classroom assessment:

> It is worth noting, right from the start, that assessment is a human process, conducted by and with human beings, and subject inevitably to the frailties of human judgment. However crisp and objective we might try to make it, and however neatly quantifiable may be our "results", assessment is closer to an art than a science. It is, after all, an exercise in human communication. (p. 2)

The goal should not be the elimination of subjectivity but the development of consistency. We can achieve this by giving teachers frequent opportunities to engage in moderation, the shared evaluation of student assessment evidence. Jurisdictions such as New Zealand, Scotland, and Queensland, Australia, have made moderation a required practice for teachers and provide the time and resources for teachers to participate effectively.

Myth 4. The results of classroom assessments are subjective; therefore, standardized assessments are better

Those who perceive classroom assessments and letter grades to be subjective often devalue the results from classroom assessments and overemphasize the results from standardized state, provincial, and district assessments.

Despite the assumption that standardized assessments are more accurate measures of student learning, there are many problems with them:

- They provide a snapshot, not the photo album provided by classroom assessment evidence.
- They may not assess the taught curriculum.
- They are insensitive to the endless variety of students' conditions on the day of the assessment, such as mental and physical health, or their anxiety levels.

Classroom assessments can solve some of these problems:

- Teachers can create electronic portfolios of student evidence of learning to make the photo album of assessment evidence available to parents (and students).
- Teachers can collaboratively plan assessments to ensure they measure the taught curriculum.
- Teachers can make appropriate accommodations because they know and care about their students and their situations.

In many jurisdictions, teachers have improved assessment by moving away from *norm-referenced* assessments that rank students and using *criterion-referenced* assessments that teachers develop together to show how students are achieving certain established, shared standards (Ontario Ministry of Education, 2010, p. 30). This makes the assessment process more transparent to the public and helps bust the myth that standardized assessments are more accurate.

Assessments that prioritize learning

The *Growing Success* document from the Ontario Ministry of Education (2010) stresses the purpose of assessment: to grow learners:

> The use of assessment for the purpose of improving learning and helping students become independent learners requires a culture in which student and teacher learn together in a collaborative relationship, each playing an active role in setting learning goals, developing success criteria, giving and receiving feedback, monitoring progress, and adjusting learning strategies. The teacher acts as a "lead learner," providing support while gradually releasing more and more responsibility to the student, as the student develops the knowledge and skills needed to become an independent learner. (p. 30)

Dismantling myths about professional judgment and subjectivity in assessment begins with a deeper understanding of the true purpose of classroom assessment. By embracing shared standards and implementing varied assessment approaches, educators can enhance the reliability of their judgments and provide meaningful, authentic insights into student learning. This approach involves creating supportive policies and professional learning opportunities that empower teachers to make well-grounded decisions in their classrooms. Ultimately, this myth-busting approach will lead to assessment practices that prioritize learning and uphold the professionalism of teaching. ■

References

Alonzo, A.C. (2019). Defining trustworthiness for teachers' multiple uses of classroom assessment results. In S.M. Brookhart & J.H. McMillan (Eds.), *Classroom assessment and educational measurement* (pp. 120–145). Routledge.

Bowers, A.J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation, 17* (3), 141-159.

Brookhart, S. (2016). The use of teacher judgement for summative assessment in the USA. In V. Klenowski (Ed.), *International Teacher Judgement Practices*. Routledge.

Brookhart, S.M., Guskey, T.R., Bowers, A.J., McMillan, J.H., Smith, J.K., Smith, L.F., Stevens, M.T., & Welsh, M.E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research, 86* (4), 803-848.

Cooper, D. (2011). *Redefining fair*. Solution Tree.

Cooper, D. (2022) *Rebooting assessment: A practical guide for balancing conversations, performances and products*. Solution Tree.

Eyler, J. (2024). *Failing our future: How grades harm students, and what we can do about it*. John Hopkins University Press.

Feldman, J. (2023). *Grading for equity: What it is, why it matters, and how it can transform schools* (2nd ed.). Corwin.

Guskey, T.R. (2020). *Get set, go! Creating successful grading and reporting systems*. Solution Tree.

Kunnath, J. (2017). Teacher grading decisions: Influences, rationale, and practices. *American Secondary Education, 45* (3), 68-88.

Link., L.J. & Guskey, T.R. (2022). Is standards-based grading effective? *Theory into Practice, 61* (4), 406-417.

Mugerauer, R. (2020). Professional judgement in clinical practice (part 2): Knowledge into practice. *Journal of Evaluation in Clinical Practice, 27* (3), 603-611.

O'Connor, K. (2018). *How to grade for learning: Linking grades to standards* (4th ed.). Corwin.

Ontario Ministry of Education (2010). *Growing success*.

Sutton, R. (1992). *Assessment: A framework for teachers*. Routledge.

Wiliam, D. & Black, D. (1998). Inside the black box: Raising standards through classroom assessment. *Kappan, 80* (2), 139-144.