Ambitious Teaching and Equitable Assessment

A Vision for Prioritizing Learning, Not Testing

By Lorrie A. Shepard

Claims that testing will serve equity very often aren't true. That's my conclusion from 50 years of studying the impact of high-stakes standardized assessments. Yes, I should have retired years ago. But I feel compelled to call attention to how seemingly well-intentioned efforts to increase student achievement actually diminish student learning—and, more importantly, to offer an alternative vision of assessment so integrated with instruction that it actually furthers learning.

Here's the arc of the last 50 years, at warp speed: minimum competency tests in the '70s; basic-skills tests in the '80s; "tests worth teaching to" in the '90s; high-frequency, high-stakes tests in the '00s; and added layers of commercial interim tests in the '10s.* After testing ourselves into a maniacal focus on reading and math, there's now a growing effort to tack on other variables, like social-emotional development, as if that could solve the horrific imbalance of accountability testing over all else.

What has all this testing accomplished? Very little. We've known since the 1980s that standardized testing in basic skills, when there are any consequences attached, results in test score inflation and curriculum distortion.[1] And we've known since 2011 that the high-stakes testing required by No Child Left Behind increased achievement only minimally. One methodologically strong study found an increase of 0.10 of a standard deviation in fourth-grade math,[2] while a research synthesis found an average increase of 0.08, with gains mainly in elementary math.[3] These findings translate to roughly 3–4 percentile points.

What has all this testing cost? Far too much. Each year, testing consumes weeks of instructional time, pulls millions of dollars away from student services and enrichment, and demoralizes budding learners across our country.

It's long past time to reckon with how the accountability testing strategy has failed, accepting that we cannot incentivize our way to equity and excellence, and to redirect our efforts to assessments that support learning. It's time to value teachers, strengthen local curricula, build on the knowledge students bring to class, foster caring classroom environments, and focus on assessments that enable next steps for instruction.

Lorrie A. Shepard, distinguished professor and dean emerita of the School of Education at the University of Colorado Boulder, has been researching assessment and working with educators to enhance student learning for 50 years. She has served as president of the National Council on Measurement in Education, the National Academy of Education, and the American Educational Research Association.

*For a more complete explanation of the dominant testing ideas of the last 50 years, please see this presentation I gave on April 21, 2021: go.aft.org/afg.

ILLUSTRATIONS BY ANGELA HSIEH

Because of the pandemic, this school year will be different from the past, even in places where in-person schooling was maintained throughout. Many educators and members of the public now more fully understand that relationships matter for learning. Debilitating inequities in educational resources and learning opportunities, known yet effectively ignored by policymakers for decades, are now undeniably exposed. And for the first time, federal monies are available on a scale sufficient to enact meaningful changes.

Many of my fellow educators and researchers have long known[4] that genuine opportunities for deep learning require two basic things: rigorous, authentic learning goals and instructional supports that ensure a sense of safety and belonging. The fact that these remain aspirational goals, rather than widespread practices, tells us that the impediments are real and enduring. Inequities are systemic in American society. Learning opportunities are not equal before or during schooling, and testing often intensifies the marginalization of students with low scores because they are given remediation instead of enrichment. So the students who face the most challenges outside of school (due to everything from residential segregation to inadequate access to healthcare[†])—specifically, Black, Latinx, and Indigenous students, nonnative speakers of English, and students with special needs—tend to be given the least inside of school. Therefore, any effort to invest in creating more equitable learning opportunities requires recognizing built-in racist and classist causes of inequity and working consciously to remove these systemic injustices.

It's long past time to reckon with how the accountability testing strategy has failed.

Developing equitable assessment practices requires starting over, with up-to-date research on teaching and learning. My goal in this article is to provide a new vision of assessment integrated with instruction for the sole purpose of supporting learning—not ranking students, teachers, or schools. I explain how *ambitious teaching practices*, framed by sociocultural theory, are essentially one and the same as *equitable assessment practices*.[5] I begin with a summary of the outmoded beliefs about learning and motivation underlying our current accountability systems. In the concluding section, I address what teachers need from district leaders and higher-level policymakers.

Outdated Beliefs About Learning and Motivation Underlie Assessment Mandates

Countless studies have shown the curriculum-narrowing effects of accountability pressures. In schools worried about raising test scores in reading and mathematics, science and social studies are driven out of the curriculum along with art, music, and PE. Worse still, testing pressure can undermine learning even in reading and mathematics because low-scoring students often receive repetitive drills, using decontextualized worksheets and other formats that closely resemble multiple-choice test items.

I don't believe that anyone is truly in favor of children receiving such a dry, uninspiring education—so how did we get to this point? To better understand our current situation, let's look back a few decades.

The cognitive researchers who helped politicians launch the first wave of standards-based reforms in the 1990s had some good ideas. Importantly, they pointed to the evidence that thinking and reasoning abilities are developed (not genetically fixed).[6] They sought to make rich and challenging curricula available to all students, rather than an elite few, hence the slogans "all students can learn" and "world-class standards." They already had evidence from the 1980s showing the harmful effects of teaching to basic-skills tests, so they called for performance assessments aligned with ambitious new standards. These would be "tests worth teaching to," with students writing essays, conducting chemistry experiments, and engaging in other

demonstrations of their current competencies. The researchers emphasized that their aspirations for a "thinking curriculum" were unprecedented and would require substantial "capacity building" and resources to help teachers teach in profoundly different ways.

Unfortunately, the idea of capacity building was replaced almost immediately by a competing theory of change based on incentives that used test scores to mete out rewards and punishments for educators.[7] The name "standards-based reform" had been hijacked. Under the new theory of action, it was assumed that with sufficient motivation (from accountability pressure), teachers and other school personnel would find the means to improve instruction and that improvement would show up in students' test scores. What research over the next decade showed, however, was that many administrators and educators did not understand the instructional changes that were needed or lacked the capacity to make them happen in a sustained, impactful way.

[†]For an in-depth discussion of the connections between race, opportunity, and well-being, see "Healing a Poisoned World" in the Fall 2020 issue of *AFT Health Care*: aft.org/hc/fall2020/washington.

No Child Left Behind (NCLB), enacted in 2002, dramatically increased both the stakes and the amount of testing, from milestone testing in grades 4, 8, and 12 to every-grade testing from 3 to 8, plus high school testing. Because of the amount of testing required, the elaborate performance assessments that had appeared briefly in the '90s were too expensive and were replaced by mostly multiple-choice tests. In addition, because of draconian NCLB performance expectations—all students proficient by 2014—districts began purchasing commercial interim tests to get ready for state tests. Just as "standards-based reform" was hijacked, so was "formative assessment." Machine-scored, multiple-choice tests are called "formative assessments," but they are nothing like the curriculum-based, ongoing, interactive processes documented in the literature on formative assessment.[8]

Thus, we now have a multilayered testing system that is limited in its ability to document progress toward deep learning goals, much less cultivate deeper learning. State tests must be curriculum-neutral to allow for local control, interim tests purchased by districts have to be generic enough to sell to national markets, and costs preclude portfolios or performance tasks. Although external tests could be useful once-per-year barometers of programmatic trends (if they did not have performance-distorting stakes attached), they are sold as if they have instructional meaning for individual students. Worse, frequent test-score reports give students the wrong idea about the purpose of learning. Feedback about how many additional points are needed to reach proficiency does not help students improve. In fact, research on motivation shows compellingly that data walls and other types of normative comparisons are harmful to learning. Initiatives for culturally responsive and sustaining pedagogy,* for example, cannot help if students experience public shaming for their low scores. Simply put, test-driven schooling is antithetical to what research on learning tells us about effective teaching and productive learning environments.

Frequent test-score reports give students the wrong idea about the purpose of learning.

Current Research Supports Integrating Ambitious Teaching with Equitable Assessment Practices

Creating truly equitable and excellent educational opportunities means ensuring that each child has access to rigorous curricular resources and is supported to participate fully in instructional activities that enable deep learning. This "ambitious teaching" centers on each student's engagement and participation; it requires paying explicit attention to who students are as they enter the classroom, including their prior learning experiences (inside and outside formal educational settings), their family- and community-based funds of knowledge, and their races, ethnicities, gender identities, social classes, and other aspects that influence their identities as learners.

Ambitious teaching practices are consistent with asset-based pedagogies, culturally responsive and sustaining teaching, and learning research in literacy (including biliteracy and bidialectism†), mathematics, and science. The foundation for all of this work is sociocultural learning theory,‡ which is the state-of-the-art model for understanding how learning happens and why context, culture, and sense of belonging are an integral part of learning.[9] Sociocultural learning theory builds on important lessons from cognitive research (in laboratory and classroom settings) about sensemaking, prior knowledge, and metacognition; it also attends to the ways that social and cultural contexts shape development, identity, and new learning. Importantly, it explains why motivational aspects of learning—feelings of self-efficacy, belonging, and purpose—are completely entwined with cognitive development.[10] Because learning is seen as transforming one's ability to *participate* in a community of practice, learning involves developing communication skills and gaining experience with tools for thinking along with an increasing sense of competence and ability to contribute.

Sociocultural learning theory, thus, creates an imperative to deeply know each student—academically, emotionally, socially, and culturally—and to offer a supportive classroom environment where students feel safe to talk together about their thinking and reasoning. That's why ambitious teaching is only possible when equitable assessment is fully integrated into instructional practice. Unlike our existing testing regime, equitable assessment is almost entirely formative—but not the *so-called formative* of today's widely used benchmark assessments (which are mainly another form of test prep). True formative assessment takes many forms, from peer conversations and sharing out of group work to classroom quizzes and exit tickets, but a core feature is that it is grounded in the classroom curriculum and makes visible useful

*For details on culturally sustaining practices that reinforce the ambitious teaching I describe in this article, see "Liberatory Education: Integrating the Science of Learning and Culturally Responsive Practice" in the Summer 2021 issue of *American Educator*: aft.org/ae/summer2021/hammond.

†To learn about bidialectism, or speaking more than one variety of English, see "Teaching Reading to African American Children" in the Summer 2021 issue of *American Educator*: aft.org/ae/summer2021/washington_seidenberg.

‡Although various lists of effective teaching practices have been derived without a theoretical foundation, theory is important for continuing to improve; it aims to explain why certain practices work, helps us think about improvements when initial efforts fall short, and provides a model of how all the pieces fit together.

information for guiding day-to-day instruction. Often, students do not know they are being assessed—they are simply sharing their thoughts and participating in activities as a normal part of the learning process.

Assessment integrated in ambitious teaching is equitable in several important ways. First and foremost, it positions students as capable learners and offers helpful information about what next, rather than a sometimes overwhelming list of all the things not known. Because students are assessed on the specific knowledge and skills they have been taught, questions and expectations are more recognizable and relevant as compared with curriculum-general state assessments. In addition, because the teacher is engaging with the student, the results are more meaningful; problems like bad days, issues at home, or simply misunderstanding a question do not skew the teacher's understanding of the student's progress.

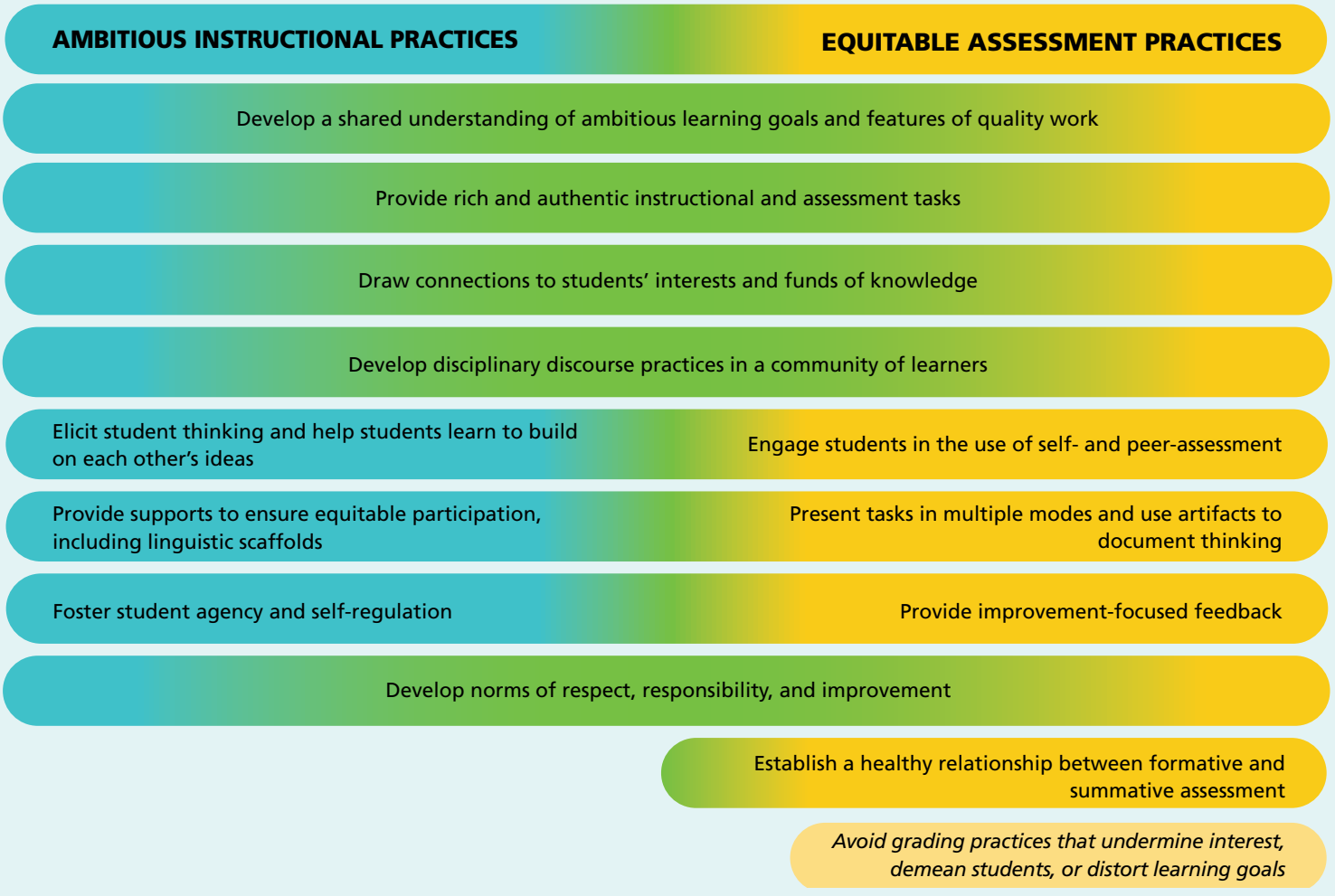
When well-integrated, equitable assessment is embedded in and enables ambitious teaching. The summary of shared and unique practices shown below follows from a set of "Classroom Assessment Principles"[11] that my colleagues and I developed in collaboration with our district and state assessment partners. We also invited and received extensive feedback from an array of participants who attended a national conference on classroom assessment. There was only one criticism to which we were unresponsive: the complaint that our principles for assessment "looked mostly like high-quality instruction." Yes, that's exactly our intention.

In the remainder of this article, I walk through each item in the figure. Although these practices would be daunting and likely incoherent if attempted piecemeal, they are highly interconnected and mutually supportive when viewed from a sociocultural perspective.

Develop a shared understanding of ambitious learning goals and features of quality work. Learning goals direct effort and shape thinking. Goals help to explain context and purpose and create a vision for what mature or expert practice looks like. To serve equity, goals must be challenging for all students (instead

Using Sociocultural Theory to Integrate Equitable Assessment with Ambitious Instructional Practices

Sociocultural theory is the foundation from which both ambitious instructional practices and equitable assessment practices derive.



of reserving ambitious goals only for some students and not others). Equity also requires that challenging goals be accessible and meaningful, which means they are not carved in stone and handed down from on high. Rather, they should be negotiated and connected based on students' interests and experiences outside of school.

True formative assessment is grounded in the classroom curriculum and makes visible information for guiding instruction.

For example, *Ambitious Science Teaching*[12] describes how *anchoring events or phenomena* can create a more particularized and compelling context for considering more general curriculum topics.* Teachers and eighth-graders in Washington state took up the problem of killer whale populations declining in Puget Sound as they learned about ecosystems. And kindergartners used their 5-year-old words like "sticky," "push," and "pull down to the ground"(for friction, force, and gravity) to talk about whether somebody little could bump somebody big off the end of a playground slide.

Involving students in shaping goals and in monitoring their own progress develops self-regulation capabilities as well as deeper understanding of success criteria. It is well recognized in the formative assessment literature that coming to understand the features of quality work—what it means to be a good writer, a good student of history, and so forth—is an integral part of developing subject matter expertise.

Provide rich and authentic instructional and assessment tasks. It follows that ambitious goals require instructional activities and assessment tasks that fully represent or embody those goals. If a goal is for students to be able to develop and evaluate historical claims and arguments, then instructional activities must involve this kind of experience, including reading across texts, examining primary documents, presenting and critiquing arguments, and the like. Formative assessment can occur as part of learning activities, with both planned-for and in-the-moment questions designed to elicit student thinking. To further the activity, some questions can be asked of the group, but individual questions are also needed to check for understanding, possibly as an exit ticket. Reporting back and showing students how their responses have helped shape next steps can enhance trust and demonstrate a joint commitment to learning (in contrast to more typical testing strategies that feel like catching and punishing students for what they don't yet know).

In a partnership of researchers and practitioners working to improve middle school mathematics instruction, the collaborators noted the importance of open-ended, high-cognitive-demand tasks.[13] Problems that can be solved in multiple ways and that expect students to engage in mathematical reasoning help students develop an understanding about why procedures work and which procedures are appropriate to use in particular situations. And when showing their reasoning is part of the learning activity, students give teachers an abundance of information to determine next steps instructionally. Similarly, *Ambitious Science Teaching* emphasizes the importance of authentic tasks, for both instruction and assessment, that simulate the kinds of intellectual work that is called for in real-world contexts—this often means deciding what the problem is that needs solving as well as transferring learning of a concept from one specific project or problem to another. Tasks with lower cognitive demands can also be informative, but these should be selected carefully, as something like scaffolds, to help teachers see students' partial understandings when they are not yet able to complete more ambitious transfer tasks.

Draw connections to students' interests and funds of knowledge. Most teachers are aware of the importance of eliciting and building on students' *prior knowledge*. But too often they've been told to probe for an inert list of prerequisite school skills. More up-to-date research acknowledges the profound ways that cultural patterns affect all aspects of learning and development. This makes experiences from home and community highly relevant to school learning. The term "funds of knowledge"[14] is becoming widely used to recognize the robust, accumulated wisdom developed in families and communities about daily concerns like cooking, budgets, first aid, and automobile repair and about core cultural values regarding morals and ethics. This knowledge, always there but sometimes disregarded in school, can be explicitly engaged as a resource for teaching. Attending to students' lived experiences furthers learning in several important ways. It shows respect and helps to counter negative positioning of students from communities that have long been marginalized. Drawing connections and providing scaffolds from everyday knowledge to academic knowledge also support intellectual development while contributing emotionally to a student's feeling of belonging.

A similar concept emerged from research on teachers who were recognized by parents and principals as successful with African American students.[15] The practices they had in common, now known as *culturally relevant pedagogy*, affirm students' cultural identities while at the same time challenging and helping

*While this discussion draws from the *Ambitious Science Teaching* book, there is also a companion website with excellent, free resources: ambitiousscienceteaching.org.

them to succeed academically. Culturally relevant pedagogy also entails engaging students in recognizing and critiquing social inequities. There is no simple formula or list of strategies; this approach centers on teachers forming strong relationships with students and being committed to their success.

In the original research,[16] the successful teachers did not use any "language of lacking," such as attributions to single-parent homes or poverty. They fought against the idea of one-right-answer thinking and demanded that students work to high intellectual levels. Teachers' assessment strategies focused on allowing multiple ways to demonstrate learning (which is discussed again later in this article). Assessment insights were also gleaned from all of the connections teachers made between school and community. Connecting rap lyrics to poetry is the most obvious; other examples included examinations of zoning laws that allowed liquor stores near schools with predominantly Black students but not near those with mainly white students and critical analyses of social studies textbooks under consideration by a state evaluation panel.

Develop disciplinary discourse practices in a community of learners. Students learn more from talk-based instructional practices,[17] primarily because explaining your reasoning to someone else is usually more challenging than passively listening without any meaningful check for understanding. In addition, developing language and inquiry skills specific to each discipline, such as mathematics, history, and science, helps students develop a much deeper understanding of that discipline. Including disciplinary practices as learning goals along with big ideas in content domains† addresses equity by ensuring that students from all backgrounds have opportunities to develop problem-solving, reasoning, and communication skills that enable participation.

Disciplinary practices, such as posing questions, analyzing and interpreting data, modeling, and argumentation, can be thought of as tools for thinking. Talk-based instructional strategies and documentation (e.g., poster presentations), as students engage with these disciplinary practices, provide opportunities for assessment and feedback without the need for separate quizzes and tests.

Elicit student thinking and help students learn to build on each other's ideas. By now, the integration of these various ambitious instructional and equitable assessment practices should be apparent. An overarching idea is that engaging students in challenging intellectual work requires emotional support, respecting who students are, and specific academic scaffolds, enabling next steps for learning. A critical aspect of assessment that's embedded in instruction is that students' thinking has to be visible—at least to the teacher and the student, and often to classmates. This means intentionally creating a productive classroom learning environment in which students are not afraid of offering inaccurate or incomplete thoughts.

In the research partnership on middle school mathematics noted earlier, the participants recognized the importance of whole-class and small-group discussions to further conceptual understanding. Making this happen requires high-quality tasks, otherwise there's nothing to question or discuss. In addition, it is crucial that classroom *norms* be established about the value of learning from mistakes, what counts as an acceptable explanation for the academic discipline, and the importance of giving a rationale for why particular steps were taken. The researchers emphasized the critical role of the teacher in "pressing students to elaborate their reasoning and to make connections between their peers' solutions and key mathematical ideas."[18]

Engage students in the use of self- and peer-assessment. Self- and peer-assessment are two of the many formative assessment practices that have been shown to increase student learning.[19] Self-assessment was initially thought of decades ago as a specific strategy to help students develop more explicit understandings of success criteria and to get better at applying these criteria to improve their own work. Subsequent theoretical framings and studies have helped us understand the power of self-assessment to enhance metacognition and executive function and also, from a sociocultural perspective, to enhance self-regulation and student agency.

Showing students how their responses helped shape next steps can enhance trust and demonstrate a joint commitment to learning.

Peer-assessment shares many of the same benefits for learning as self-assessment.[20] Peer-assessment provides a vehicle for student talk and more explicit attention to the features of quality work. Getting better over time at learning from and being able to critique the work of classmates contributes directly to students' use of disciplinary practices, such as making claims from evidence or being explicit in the use of definitions.

None of these claims about the benefits of self- and peer-assessment are true, however, if students aren't taught how to engage in this kind of feedback meaningfully or if these

†This is one positive feature of some recently developed standards, including the Common Core State Standards and the Next Generation Science Standards.

strategies are imposed only as an additional bureaucratic requirement. This is not about checking right/wrong answers or assigning grades. Rather, self- or peer-assessment practices should be taken up with explicit attention to the ways that they are expected to further self-monitoring and new learning. Because an essential purpose of formative assessment is to provide feedback about how to improve, there must be subsequent opportunities to revise. Too often, students complete assignments "to be done," rather than seeing learning intentions and how new learning will be used and built upon in subsequent classroom work. Insights from self- and peer-assessment need to have a place in the rhythms of classroom activities.

Provide supports to ensure equitable participation, including linguistic scaffolds. Talking about thinking is one of the most powerful tools we have to ensure that all students have the opportunity to engage fully with rigorous learning goals. But hoped-for critical-thinking talk doesn't happen automatically. *Ambitious Science Teaching* explains the process by which participation in talk leads to higher-order learning. It also provides advice about how to set goals for classroom conversations, how to plan high-cognitive-demand questions, and how to use a repertoire of "talk moves" to be responsive to student ideas. For example, "pressing" is a way to ask students for more: "Can you give an example?" "What evidence supports that idea?" "Sounds like you have the beginning and the end of an explanation [repeat students' partial explanation]; can you say what happens in the middle?" *Ambitious Science Teaching* also offers scaffolds for student talk, which are especially important for students who have not participated in such conversations previously. Scaffolds include teacher modeling and coaching, sentence starters or sample questions to launch group work, and ways to simplify complex tasks without doing the thinking for students.

Engaging students in challenging intellectual work requires emotional support, respecting who students are, and specific academic scaffolds.

Scaffolds to support participation are especially important for English learners. Most teachers are aware of the importance of making sure that students are comfortable with the meanings of academic words like *compare*, *contrast*, *hypothesize*, *cause*, and *effect*; but like everything else, real understanding is more likely in discussions about specific applications of such words. Thus, it is beneficial to make these clarifications part of class discussions as new work is being launched. Later, during discussions when students are being asked to explain their group's model or problem solution, teachers may need to invent specific scaffolds for students who are not yet fluent in English. This might mean pointing to the relevant portion of a graphical aid, letting students ask a classmate for help, or coming back at the end to invite an English learner to say "I agree with..." and having permission to repeat what another student has said or posted. Again, these moves are done in the spirit of collaborative learning, with an appreciation shared by the class that our thinking is almost always ahead of our ability to put our thoughts into complete verbal explanations.

Present tasks in multiple modes and use artifacts to document thinking. In addition to talk-based instructional practices that elicit and build on student thinking, presenting tasks in multiple modes and allowing students multiple ways to demonstrate their learning can serve equity goals and affirm a positive learning culture.* In addition, representing learning in multiple ways can deepen students' conceptual understanding by drawing connections and offering more than one way to think about a new idea.

The teachers that parents and principals had identified as exemplary teachers of African American students held multifaceted conceptions of assessment and engaged students in work reflecting multiple forms of excellence.[21] In one example, a teacher helped her students choose the standards by which they would be evaluated and what evidence or work products they wanted to use as proof of mastery of specific concepts and skills. Another teacher emphasized questioning as a recurring pattern in classroom interactions, asking "Why are we doing this problem?" This invited students to interpret tasks and respond in ways that played to their particular strengths—it also created greater access to the content and the classroom discourse. As students' various answers and approaches were shared across the class, much more robust understandings developed about how targeted knowledge and skills were to be explained and used.

Foster student agency and self-regulation. Fostering student agency and developing self-regulation capabilities are broad, overlapping categories of practices that sum up several of the specific strategies and intentions addressed above and below. Self-regulation, which emerged from cognitive theory, and student agency, which emerged from sociocultural theory, are closely overlapping constructs having to do with both cognitive and affective aspects of learning. They are about developing the

*While multiple modes are beneficial, this should not be confused with learning styles. For details, see "Ask the Cognitive Scientist: Does Tailoring Instruction to 'Learning Styles' Help Students Learn?" in the Summer 2018 issue of *American Educator*: aft.org/ae/summer2018/willingham.

awareness, self-confidence, and skills to take responsibility for one's own learning—and they are critical for motivation.

In summarizing the vast research on motivation, the recent milestone report from the National Academies of Sciences, Engineering, and Medicine, *How People Learn II,* concluded that, "Motivation to learn is fostered for learners of all ages when they perceive the school or learning environment is a place where they 'belong' and when the environment promotes their sense of agency and purpose."[22] The report also summed up what educators can do as follows.[23]

> Educators may support learners' motivation by attending to their engagement, persistence, and performance by
> - helping them to set desired learning goals and appropriately challenging goals for performance;
> - creating learning experiences that they value;
> - supporting their sense of control and autonomy;
> - developing their sense of competency by helping them to recognize, monitor, and strategize about their learning progress; and
> - creating an emotionally supportive and nonthreatening learning environment where learners feel safe and valued.

With this understanding of motivation—which, notably, does not include external rewards like scores, stickers, or pizza parties—we can see how it is that equitable assessment bolsters motivation. Self-assessment in particular is intended to help students develop their agency and self-regulation. This practice can be especially fruitful when students are given feedback about the quality of their self-assessments. Indeed, all of these ambitious teaching and equitable assessment practices are intended to work together in support of student agency and self-regulation.

Provide improvement-focused feedback. Assessments that result in normative comparisons—posting Jacob's score as "Below Basic" or telling Keisha that she scored at the 55th percentile, for example—undermine learning. This conclusion comes from many hundreds of studies[24] showing that students who receive this kind of feedback do worse, on average, on subsequent measures of achievement than students in control groups that received no feedback. This type of feedback, where students are told how their performance compares to other students, is also called ego-focused or person-focused feedback. In contrast, task-focused feedback that shows students something about how to improve has a positive effect on learning.

How People Learn II provides this summary of the relevant research.[25]

> Feedback is most effective when it is
> - focused on the task and learning targets; that is, detailed and narrative, not evaluative and graded;
> - delivered in a way that is supportive and aligned with the learner's progress;
> - delivered at a time when the learner can benefit from it; and
> - delivered to a receptive learner who has the self-efficacy needed to respond.

In reflecting on how all the equitable practices in the figure on page 31 fit together, notice that they all attend to the identity and feelings of students as members of a learning community. Equitable assessment is not about offering false praise or lowering expectations. Rather, engaging students with specific information about how to improve their work conveys respect (because of the teacher's confidence that the student is able to do this higher-level work), and it invites students to take greater ownership and thereby have a greater sense of control. I have said this many times before, but it is worth repeating: feedback that helps students think about how to improve their work requires *substantive insights* and is, therefore, more often qualitative (e.g., written comments or a discussion) rather than quantitative (e.g., a score).

Task-focused feedback that shows students something about how to improve has a positive effect on learning.

Develop norms of respect, responsibility, and improvement. The practice of developing classroom norms in support of mutual respect, personal responsibility, and a shared focus on learning is fundamental to the idea of a learning culture. Creating an environment where students feel safe to ask questions, are willing to share partially formed explanations, and are able to offer critiques of each other's reasoning without meanness or injured feelings requires explicit negotiation and scaffolds (as is true when working toward any meaningful learning goals). Explicit work to jointly establish such norms is imperative if students have not become accustomed to such expectations in other classrooms or in prior years of schooling.

Ambitious Science Teaching suggests ways that norms can be co-constructed through role playing, for example, or by asking students directly what kinds of comments might keep them from participating and then discussing as a class how disrespectful

comments might be avoided. Examples of the kinds of norms that might be negotiated include students taking responsibility for their own learning by seeking clarification when they don't understand something and not interrupting or talking over classmates when they have the floor. For these norms to be felt and lived by, they must also be enforced, as when a teacher needs to step in and say, "That's a put-down rather than a fair critique of an idea; how can you rephrase what you said?"[26] A good indicator of when norms are embedded in the learning culture is when students remind each other to stay on topic or to let someone else have a turn.

Establish a healthy relationship between formative and summative assessment. Ideally, in an affirming learning culture, students are excited to engage with new learning because of the intrinsic appeal of the goals and tasks—and because it feels good to contribute to the efforts of the group. Formative assessment practices such as making posters or using Google Docs to report out group questions or initial models for discussion should be seen as helping the class learn together. For more formal assessment practices, such as feedback or peer- and self-assessment, it is important that there be clear conceptual linkages to culminating summative assessments. That way, formative assessment can be seen as supporting improvement toward the same criteria and goals that will be called upon by summative assessments.

Current summative assessment strategies, like regularly giving students formal tests and posting the scores in electronic gradebook management systems, may be handy for parents to check on progress, but they are antithetical to what we know from learning research. First, they involve reducing substantive insights about what students know to mere points, which cannot tell the story about student interest or effort or how to improve. Point systems can foster extrinsic rather than intrinsic motivation to learn (and can reduce intrinsic motivation, as discussed next). And, oddly, even when recorded weekly, each data entry is typically treated as if learning were finished rather than a step in a learning progression; it fails to note substantive improvements over time, much less guidance about how to improve further.

Avoid grading practices that undermine interest, demean students, or distort learning goals. Much of the research on grading tends to involve surveys of current practices rather than examining how grading practices affect learning. One comprehensive summary of these survey studies noted that teachers give credit for "enabling factors"[27] (e.g., effort, ability, improvement, work habits, attention, and participation) in addition to mastery of learning goals. Better evidence about how grades affect learning comes from the motivation literature and from formative assessment research (much of which is summarized earlier). In short, grading best supports learning (and a learning orientation, which is crucial for future learning) when grades reflect mastery of the specific learning goals toward which instruction and feedback have been aimed and when grades are *not* used to try to motivate or control student behavior.

Even with my decades of assessment experience, I find it difficult to convince fellow educators to give up grades as motivators.

Even with my decades of assessment experience, I find it difficult to convince fellow educators to give up grades as motivators. After all, extrinsic rewards do work, and students are more likely to turn in assignments and turn off their phones if you make these things "count" toward their grades. What was most convincing to me were the studies of extrinsic versus intrinsic rewards,[28] which show that students (even kindergartners) have less interest in learning activities after these activities had been "reinforced" with extrinsic rewards. Another important nuance related to grading is that many teachers want a way to credit nonachievement factors like effort because of their kindness and caring for students. An alternative practice that is kind and keeps the focus on learning is to allow students to submit new evidence of learning in place of earlier assignments. This makes sense especially when teachers and students together are aware of how a later assignment subsumes knowledge and skills required for an earlier assignment. As it stands now, many teachers feel they are bound by the bureaucracy of prior gradebook points. But ultimately, grades should be about the substantive learning that has been demonstrated—not whether it was demonstrated in week 14 or week 17.

Conclusion: What Districts Can Do to Support Ambitious Instruction and Equitable Assessment

Sociocultural theory helps us understand why personal relationships—between student and teacher and among students—are critical for academic achievement. It helps us see how to create affirming classroom environments by attending to students' identities and sense of belonging and, at the same time, how to ensure rigor by explicitly structuring and scaffolding students' participation skills to cultivate higher levels of thinking. Equitable assessment practices are themselves ambitious teaching practices and are entwined with additional instructional practices that together lead to high levels of engagement and deeper learning.

Some teachers already exemplify ambitious teaching and equitable assessment. After all, the science and math partnerships and the teachers succeeding with African American students described here are all grounded in observing teachers who had been identified as highly effective. Many teachers have implemented some but not all of these ideas. An especially familiar pattern is when teachers are able to implement one set of ambitious practices, such as discourse-based instructional practices, but then maintain more traditional, multiple-choice quizzes and tests instead of show-your-thinking, authentic assessment tasks. For other teachers, both ambitious teaching and equitable assessment practices are new; this is particularly likely in contexts where districts have not had the resources to invest in professional development or curricular materials connected to topics and problems of particular interest to local communities.

Learning to teach or improve one's practice in the ways described in this article can be daunting. Support for teacher learning is just as important as an equity-focused vision for student learning. In our "Classroom Assessment Principles,"* my colleagues and I identified five recommendations as to what school and district leaders could do to support equitable assessment practices in classrooms.

1. Implement coherent curricular activity systems that integrate curriculum, instruction, and assessment based on well-founded theories of learning.
2. Build collaborations between assessment and curriculum department staff to inform the design and implementation of coherent curricular activity systems in schools.
3. Provide professional development and coaching structures (e.g., time, supports for educator collaboration) that help to coordinate all of the different new things that teachers are being asked to learn, including learning and motivation theories, asset-based pedagogy, disciplinary practices as part of content standards, and classroom assessment principles.
4. Develop or adopt district-level assessments that embody the full range of desired learning goals.
5. Establish grading policies in support of grading practices aimed at creating clear success criteria, while avoiding the use of grades as motivators.

The single most important idea here is that district leaders must understand the research base informing these recommendations and themselves hold a coherent vision of how equity-focused assessment practices fit within commitments to asset-based pedagogies, rigorous subject-matter standards, and culturally responsive and sustaining teaching. The theories of learning (whether implicit or explicit) that govern district-level decisions matter: unhappily, districts are sometimes the cause of impediments to best practice. This happens when district leadership applies intense pressure to raise scores on accountability tests at the expense of other considerations, when districts invest in multiple-choice "formative" test products instead of substantively rich curricular and assessment resources, and when the rules for data management systems emphasize quantizing information rather than substantively describing progress. It also happens when districts create incoherence. Many districts, with the best of intentions, launch multiple worthwhile initiatives, but they do not coordinate among those initiatives, leaving educators struggling to make connections and to find time to squeeze in each mandated activity. Even under the current, highly counterproductive federal and state testing regimes, districts can and must do better.

As we emerge from the pandemic and take stock of our values, I hope we will fundamentally rethink how we approach teaching, assessment, learning, and youth development. The vision offered in this article would be one way to conceptualize how new monies and all the many reform ideas—about rigorous content standards; diversity, equity, and inclusion initiatives; social-emotional learning; culturally sustaining pedagogy; and equitable assessment practices—could be coherently aligned and mutually supportive. If not this particular vision, then the important thing is that a coherent plan be devised that is grounded in what research on learning has taught us about equity. □

I hope we will fundamentally rethink how we approach teaching, assessment, learning, and youth development.

*This detailed resource, which also addresses what state leaders and teacher educators should do, is available for free at go.aft.org/yui.

Endnotes

1. US Congress, Office of Technology Assessment, *Testing in American Schools: Asking the Right Questions* (Washington, DC: US Government Printing Office, 1992).

2. T. Dee and B. Jacob, "The Impact of No Child Left Behind on Student Achievement," *Journal of Policy Analysis and Management* 30, no. 3 (2011): 418–46.

3. National Research Council, *Incentives and Test-Based Accountability in Education* (Washington, DC: National Academies Press, 2011).

4. L. A. Shepard, "The Role of Assessment in a Learning Culture," *Educational Researcher* 29, no. 7 (2000): 4–14.

5. L. A. Shepard, "Classroom Assessment to Support Teaching and Learning," *Annals of the American Academy of Political and Social Science* 683 (2019): 183–200.

6. L. B. Resnick and D. P. Resnick, "Assessing the Thinking Curriculum: New Tools for Education Reform," in *Changing Assessments: Alternative Views of Aptitude, Achievement, and Instruction*, ed. B. R. Gifford and M. C. O'Connor (Boston: Kluwer Academic, 1992), 37–75.

(Continued on page 48)

All Children Thriving

(Continued from page 26)

15. Rose, *The End of Average*.

16. L. Vygotsky and M. Cole, *Mind in Society: The Development of Higher Psychological Processes* (Cambridge, MA: Harvard University Press, 1978).

17. Rose, *The End of Average*.

18. C. Dweck, *Mindset: The New Psychology of Success* (New York: Ballantine, 2016).

19. McEwan, "The Brain on Stress"; and G. Slavich and S. Cole, "The Emerging Field of Human Social Genomics," *Clinical Psychological Science* 1 (2013): 331–48.

20. C. Raeff, *Exploring the Dynamics of Human Development: An Integrative Approach* (New York: Oxford University Press, 2016).

21. D. Osher et al., *Thriving, Robust Equity, and Transformative Learning and Development: A More Powerful Conceptualization of the Contributors to Youth Success* (Washington, DC: American Institutes for Research and Forum for Youth Investment, 2020); and M. Tucker, *Race in America 2020* (Washington, DC: National Center on Education and the Economy, 2020), ncee.org/2020/06/race-in-america-2020.

22. Lerner, *Liberty*.

23. M. Spencer, "Phenomenological Variant of Ecological Systems Theory (PVEST): A Human Development Synthesis Applicable to Diverse Individuals and Groups," in *Handbook of Child Psychology: Theoretical Models of Human Development*, vol. 1, 6th ed., ed. W. Damon and R. Lerner (Hoboken, NJ: Wiley, 2006), 829–94.

24. D. Moore, *The Developing Genome: An Introduction to Behavioral Epigenetics* (New York: Oxford University Press, 2015).

25. M. Pigliucci and G. Mueller, "Elements of an Extended Evolutionary Synthesis," in *Evolution—The Extended Synthesis*, ed. M. Pigliucci and G. Mueller (Cambridge, MA: MIT Press, 2010), 3–17.

26. Slavich and Cole, "The Emerging Field."

27. T. Misteli, "The Cell Biology of Genomes: Bringing the Double Helix to Life," *Cell* 152 (2013): 1209–12.

28. D. Hebb, "Drives and the C. N. S. (Conceptual Nervous System)," *Psychological Review* 62, no. 4 (1955): 243–54.

29. K. Fischer and T. Bidell, "Dynamic Development of Action and Thought," in *Handbook of Child Psychology: Theoretical Models of Human Development*, vol. 1, 6th ed., ed. W. Damon and R. Lerner (Hoboken, NJ: Wiley, 2006), 313–99.

30. C. Lee, "Soaring Above the Clouds, Delving the Ocean's Depths: Understanding the Ecologies of Human Learning and the Challenge for Education Science," *Educational Researcher* 39, no. 9 (2010): 643–55.

31. Vygotsky and Cole, *Mind in Society*.

32. I. Kendi, *How to Be an Antiracist* (New York: Penguin Random House, 2019), 103.

33. N. Nasir et al., *Handbook of Cultural Foundations of Learning* (New York: Routledge, 2020).

34. P. Cantor et al., *Whole-Child Development, Learning, and Thriving: A Dynamic Systems Approach* (Cambridge, MA: Cambridge University Press, 2021), 21–22.

35. Fischer and Bidell, "Dynamic Development."

36. This description is adapted from L. Darling-Hammond et al., *Design Principles for Schools: Putting the Science of Learning and Development into Action* (Palo Alto, CA, and Washington, DC: Learning Policy Institute and Turnaround for Children, 2021).

37. This description of East Palo Alto Academy is adapted from L. Darling-Hammond et al., *Be the Change: Reinventing School for Student Success* (New York: Teachers College Press, 2016).

38. J. Fortson et al., *Impact and Implementation Findings from an Experimental Evaluation of Playworks: Effects on School Climate, Academic Learning, Student Social Skills and Behavior* (Princeton, NJ: Mathematica Policy Research, 2013).

39. B. Bloom, "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring," *Educational Researcher* 13, no. 6 (1984): 4–16.

Ambitious Teaching

(Continued from page 37)

7. National Research Council, *Testing, Teaching, and Learning: A Guide for States and School Districts* (Washington, DC: National Academies Press, 1999).

8. L. A. Shepard, "Formative Assessment: Caveat Emptor," in *The Future of Assessment: Shaping Teaching and Learning*, ed. C. A. Dwyer (New York: Lawrence Earlbaum, 2008), 279–303.

9. National Academies of Sciences, Engineering, and Medicine, *How People Learn II: Learners, Contexts, and Cultures* (Washington, DC: National Academies Press, 2018).

10. L. A. Shepard, W. R. Penuel, and J. W. Pellegrino, "Using Learning and Motivation Theories to Coherently Link Formative Assessment, Grading Practices, and Large-Scale Assessment," *Educational Measurement: Issues and Practice* 37, no. 1 (2018): 21–34.

11. L. A. Shepard et al., *Classroom Assessment Principles to Support Teaching and Learning* (Boulder, CO: Center for Assessment, Design, Research and Evaluation, University of Colorado Boulder, 2020).

12. M. Windschitl, J. Thompson, and M. Braaten, *Ambitious Science Teaching* (Cambridge, MA: Harvard Education Press, 2018).

13. P. Cobb et al., *Systems for Instructional Improvement: Creating Coherence from the Classroom to the District Office* (Cambridge, MA: Harvard Education Press, 2018).

14. L. Moll et al., "Funds of Knowledge for Teaching: Using A Qualitative Approach to Connect Homes and Classrooms," *Theory into Practice* 31, no. 2 (1992): 132–41.

15. G. Ladson-Billings, "Toward a Theory of Culturally Relevant Pedagogy," *American Educational Research Journal* 32, no. 3 (1995): 465–91.

16. Ladson-Billings, "Toward a Theory."

17. S. Michaels, C. O'Connor, and L. B. Resnick, "Deliberative Discourse Idealized and Realized: Accountable Talk in the Classroom and in Civic Life," *Studies in Philosophy and Education* 27 (2008): 284–97.

18. Cobb et al., *Systems for Instructional Improvement*, 49.

19. H. Andrade and A. Valtcheva, "Promoting Learning and Achievement Through Self-Assessment," *Theory into Practice* 48, no. 1 (2009): 12–19.

20. P. Black et al., "Working Inside the Black Box: Assessment for Learning in the Classroom," *Phi Delta Kappan* 86, no. 1 (2004): 8–21.

21. Ladson-Billings, "Toward a Theory."

22. National Research Council, *How People Learn II*, 133.

23. National Research Council, *How People Learn II*, 133.

24. A. N. Kluger and A. DeNisi, "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory," *Psychological Bulletin* 119, no. 2 (1996): 254–84.

25. National Research Council, *How People Learn II*, 155.

26. Windschitl, Thompson, and Braaten, *Ambitious Science Teaching*, 66–67.

27. J. H. McMillan, "Secondary Teachers' Classroom Assessment and Grading Practices," *Educational Measurement: Issues and Practice* 20, no. 1 (2001): 20–32.

28. E. L. Deci, R. Koestner, and R. M. Ryan, "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation," *Psychological Bulletin* 125 (1999): 627–68.