CASE STUDY

Developing Educational Assessments to Serve Learners

Susan M. Brookhart

Developing Educational Assessments to Serve Learners

Susan M. Brookhart

Abstract

In this chapter I examine aspects of educational assessment that could or should change or evolve to serve the needs of learners more effectively in the future. I approach this task in three main sections: by considering (1) what learners need; (2) how assessment can help meet these needs (i.e., examining how assessment meets learners' needs now and how that could be enhanced by new developments in educational tools and processes); and (3) how validity and validation concerns might develop accordingly. A final section summarizes these points in relation to three of the seven Principles for Assessment in the Service of Learning and makes recommendations to inform the development of assessment tools and processes that better support learners in their learning.

Children arrive at school as young—sometimes very young—people who have already accomplished a great deal of learning. Much of that learning was developmental and learned in the context of a family or other care group: learning to walk, learning to speak, learning to whom to turn for food and emotional support. Children add an additional learning context to their lives when they come to school. Cognitive outcomes become more salient in school (Resnick, 1987). As students age, learning becomes less physical and visual, and more mediated by language, but students still need to be healthy, safe, engaged, supported, and challenged (Slade & Griffith, 2013). They still need a learning community. Shepard et al. (2018) argued that sociocultural theory acknowledges this need by positing that self-regulation, self-efficacy, sense of belonging, and identity are interwoven with cognitive development. School learning is situated in classroom learning communities, somewhat similar to the way preschool learning was situated in a family or care group.

The primary assessment from a learning perspective is classroom formative assessment; other assessments for other purposes (e.g., accountability assessment) need to be coherent with and connected to formative assessment (Shepard et al., 2018). Formative assessment is assessment that occurs during learning, providing information to students and teachers that can move the learning forward. Formative assessment involves both a process and some sort of instrument, tool, or method (Bennett, 2011). The formative assessment process used in a classroom learning community will best meet students' learning needs if it uses the formative learning cycle (Where am I going?—Where am I now?—Where to next?) so that students can activate cognitive, affective, and behavioral regulation strategies to move them toward their learning goal (Andrade & Brookhart, 2019; Andrade et al., 2021), developing their evaluative judgment (Panadero et al., 2019).

To adopt a sociocultural stance and recognize that learning is done by whole persons does not negate all prior theories—for example, it is still useful to study and understand students' cognitive structures (Shepard et al., 2018; Wiliam, Fisher, & Frey, 2024). What a sociocultural framing does is allow us to focus on students' self- and co-regulation of learning as primary aspects of student learning needs and on features of formative assessment and feedback as primary ways in which assessment meets those needs (Bailey & Heritage, 2018). These aspects are detailed in the next two sections.

What Do Learners Need in Order to Learn?

Learners need a supportive classroom learning culture (sometimes called the classroom learning environment; Ames, 1992). They also need clear learning goals and success criteria, shared with them in such a way that they can activate self- and co-regulatory processes to actively move themselves closer to the learning goal. In classrooms, this is often called the formative learning cycle (Brookhart & DePascale, in press; Brookhart & McTighe, 2017). Finally, they need high-quality feedback from teachers, self, peers, and sometimes from computers or other learning materials and, importantly, the opportunity to use that feedback. Effective feedback and its use are part of the formative learning cycle, but feedback is so important that it warrants its own discussion, which appears in a subsequent section of this chapter.

Supportive Classroom Learning Culture

Learners need a classroom culture that views mistakes as opportunities to learn and encourages learning together. The idea that classroom culture can contribute to students' motivation to learn has been around a long time and can be traced to theories about students' learning goal orientations (Ames, 1992). How assessment is used is an important element in determining whether students perceive their classroom as a learning culture or an evaluative culture (Ames, 1992; Crooks, 1988).

Students need to think of themselves as competent learners who belong in a community of others who are also active learners. An important feature of a learning community is how students view being wrong and the role of productive struggle in learning (McMillan, 2018). Leighton et al. (2013) proposed one way to create a safe learning environment is to explain to students how making and then understanding errors has value for learning. This explanation helps students expect to make errors and makes it safe for them to discuss their errors as they learn complex content. While some learning can be accomplished individually, for example memorizing math facts, deeper and more meaningful learning is best accomplished in a community where students can work with peers, including more and less experienced peers, on learning tasks (Laal & Ghodsi, 2012).

To benefit from instruction, learners need appropriate background knowledge and experience. They need to know their own cultural background knowledge and experience is honored. Giving students voice and choice in assessment is an important way to bring students' cultural background into instruction and assessment (Ladson-Billings, 2014; Taylor & Nolen, 2022). Ladson-Billings (1995) used the term "culturally relevant pedagogy" to describe teaching that builds on connections between teachers and students' families, communities, and daily lives. More recently, others have added to the insights that students need school—and assessment—to reflect their own cultures and funds of identity (Esteban-Guitart & Moll, 2014; Randall, 2021). This is an important part of creating a classroom culture of learning where students feel safe, welcomed, and supported in their learning.

The principle of active student involvement in learning in a classroom learning culture, and the underlying beliefs that students construct their own understandings, can pose a difficulty for teachers who are used to thinking "What am I going to teach?" instead of "What will students try to learn?" Studies of formative assessment in both pre-service and in-service teacher education have found that at both levels, programs in which teachers developed effective formative assessment practices were those that helped teachers shift their beliefs about student learning to realize that the assessment information needed to inform students' regulation of learning more than teachers' lesson plans (Brookhart, 2017).

Clear Learning Goals and Success Criteria

Given a supportive classroom learning culture, the first thing students need is to understand what learning goal they are pursuing (Chen et al., 2017; Heritage & Wylie, 2020). Having a goal is what makes the difference between student compliance—students simply doing what the teacher asks—and students' regulation of their learning. Regulation of learning requires having a learning goal (Zimmerman & Schunk, 2011). The goal needs to be specific enough that students have a clear sense of what they are trying to learn. This is typically accomplished by sharing expectations or criteria for what counts as evidence of learning (sometimes called success criteria), by sharing concrete examples of different levels of work with students, and by discussions and activities around the criteria and examples (Chen et al., 2017; James et al., 2006; Heritage & Wylie, 2020). During learning, students apply the learning goal and success criteria to engage in the self-regulation of learning (Moss, 2022).

The Formative Learning Cycle

Once a learning goal is set, learners need to set their sights on it, aim for it, and activate and sustain cognitions, affects, and behaviors in pursuit of the goal; in other words, they need to muster thoughts, motivation, and effort to regulate their learning (Zimmerman & Schunk, 2011). This metaphor of "aiming" is the image behind the term "learning target." It is not the goal, but rather the student aiming, that comprises regulation of learning. This regulation may be broader than self-regulation and include assessment information from teachers, peers, technology, and other learning materials (Andrade et al., 2021); such regulation is known as co-regulated learning and involves the whole classroom learning culture.

As a whole, this process is sometimes called the formative learning cycle (Brookhart & DePascale, in press; Brookhart & McTighe, 2017). The formative learning cycle is a practical instantiation of the process of the regulation of learning, inspired by three conditions of formative assessment originally proposed by Sadler (1989) and expanded into a three-question model of feedback by Hattie and Timperley (2007). In the formative learning cycle, students set a goal (Where am I going?); gather feedback on formative practice work from multiple sources to compare where they are and where they need to be (Where am I now?); and consider suggestions for next steps (Where to next?). Feedback targeted to a learning goal can lead to students increasing effort and motivation, seeking additional information, changing learning strategies, and restructuring cognitions (Hattie et al., 2021). Formative assessment is especially effective for learning when students initiate self-assessment (Lee et al., 2020), in other words, when they take ownership of the formative assessment cycle.

To regulate their learning, learners need to be engaged and active, paying attention, exerting appropriate effort, and employing metacognitive skills (Andrade et al., 2021). Giving students voice and choice supports these efforts (Taylor & Nolen, 2022). Hattie and Clarke (2019) described a feedback culture—what I have been calling in this chapter a classroom culture that supports learning—as one based on the formative learning cycle and characterized by students who have the "skill, will, and thrill" (pp. 12–13) to use feedback to move from surface learning to deep understanding, where they can relate a concept to other ideas and apply it in other contexts. That is, students need learning skill, for example knowing how and when to focus their work and thinking and in what direction, to move their work and understanding closer to the criteria. They need the will or disposition to exert the effort needed to do this, based on the belief that this work will help them learn or make them smarter. They need the thrill or motivation to reach the success criteria, in other words, they need to be truly aiming toward the learning goal which has become their own and not just the teacher's.

Feedback and the Opportunity to Use It

Learners need descriptive, non-evaluative, ungraded feedback and opportunities to use the feedback to approach the learning goal (Brookhart, 2018; Hattie, 2009; Hattie & Clarke, 2019), differentiated according to the learner's proficiency level (Stobart, 2018), and meaningful to the learner (Taylor & Nolen, 2022). Several recent reviews of feedback research have shown that some types of feedback are more powerful than others. Outcome feedback, sometimes called knowledge of results or verification (Shute, 2008), is the simplest and most common type of feedback. Outcome feedback tells students whether they were correct or incorrect, or what their score or grade was. This kind of feedback is useful for some purposes, especially for tasks involving recall. Outcome feedback about correctness coupled with knowledge of the correct response, as for example on the back of math fact flash cards or in some computer learning software, can be effective for memory tasks (Mason & Bruning, 2001).

In contrast, cognitive feedback or elaboration (Shute, 2008) contains information that students can use in their thinking. Cognitive feedback helps students interpret the task, interpret their response or response processes in light of criteria, set goals and monitor progress, address particular errors, and envision next steps; in other words, descriptive, elaborated feedback supports the processes of the formative learning cycle. This kind of feedback generally has more powerful effects on learning (Hattie & Timperley, 2007; Shute, 2008; Van der Kleij et al., 2015). Whether elaborated feedback can be given depends in part on the cognitive demands of the assessment task. Complex tasks typically provide more opportunity for elaborated feedback because student responses include more evidence of student thinking than student responses to recall-level tasks. The quality of the assessment question or task is critical in supporting feedback and thus supporting learning.

Feedback is moot unless students have an opportunity to use it. The opportunity needs to be built into the instructional sequence. Some teachers mistakenly believe that students will make mental notes of their feedback and use it "next time." However, learners need concrete opportunities to use feedback to move learning forward while they are still in the process of moving toward the intended learning goal. Jonsson and Panadero (2018) examined research on students' use of feedback in higher education and described three aspects of context that influence whether and how students will use feedback. One aspect is whether assignments are given in stages in higher education, where there is opportunity for feedback and the possibility of improvement. Similarly, in K–12 education students need formative practice work, timely feedback, and opportunities for revision (Chen et al., 2017). The second aspect affecting students' use of feedback is whether they have been taught how to do that. The third is whether descriptive feedback comes with other evaluative measures, like a score or grade. When feedback is accompanied by grades, students often focus on the grade and do not engage with the feedback (Winstone et al., 2016).

Section Summary

The main theme in this section is that learners need each other, they need content, and they need instruction and assessment that allows them to activate regulatory processes to become active, engaged students. Students need a classroom community in which they can learn with others, in which they feel safe to pursue learning challenges even when they might be wrong, and in which their own background and experience is honored. In this classroom community, optimal learning occurs when students exercise self- and co-regulation of learning by aiming for a learning goal, receiving feedback on how they are doing and suggestions for next steps, and having the opportunity to muster their cognitive, affective, and behavioral skills to use the feedback. The formative learning cycle is a key process.

How do assessments meet learners' needs?

Arguably the most important assessment features to meet these learner needs are high-quality assessment and learning tasks, attention to the formative learning cycle in the structure and sequence of classroom learning, feedback and scores that produce useful information for the students, and where possible the use of learning progressions to design assessments and interpret results. The following sections describe how high-quality assessments meet students' needs now; each section ends by suggesting how these features could be enhanced by new developments in assessment.

Assessment Questions and Tasks

Assessment questions and tasks help students learn when they are educative, engaging, provocative of student thinking, and relevant to the learner's culture and experience. Tasks are educative when they are related to real disciplinary thinking and are rich enough to provide both students and teachers with feedback they can use to improve performance (Heritage & Wylie, 2020, Wiggins, 1998). Such tasks help students in their formative learning cycle by instantiating what it means to understand or be able to do the kind of thinking or work implied by the learning goal. They also aid the formative learning cycle by provoking student responses that give evidence of that thinking and allow feedback on student thinking, not just correctness.

Tasks should be a clear match with the learning goal(s) assessed (Brookhart & DePascale, in press; Heritage, 2013). It does no good to share a learning goal and success criteria with students and then use assessments that do not line up with those goals and criteria. When assessment tasks clearly embody the desired learning outcomes, assessment tasks are also learning tasks (Carless, 2015). One recent study in higher education found the quality of assessment tasks affected the quality of feedback and participation, these variables affected student empowerment, strategic learning and self-regulation, and all variables directly or indirectly affected students' learning transfer (Ibarra-Sáiz et al., 2021).

The concept of coherence (Wilson, 2004) is also relevant here. All assessments used in a school or district should interpret learning goals at different levels—from classroom lesson-sized goals through large-scale standards-based accountability goals—in the same way, so what is taught is what is assessed, and students can recognize that in their learning and performance, from classroom formative assessment through large-scale assessment. This does not mean classroom assessment tasks must be the same as large-scale assessment tasks. In fact, they typically will not be, as classroom assessment tasks usually reflect smaller chunks of learning than large-scale assessment tasks. Rather, it means the underlying construct—what students are trying to learn—must be coherent throughout an assessment system. For example, does understanding the water cycle mean being able to list or draw its steps or write hypotheses about water-related problems (or both)? If a teacher's formative assessment interprets a standard one way and other assessments in the system interpret a standard in another, coherence is lacking and learners are confused.

Furthermore, assessment questions and tasks should be relevant to the learners' culture and experience (Ladson-Billings, 2014; Randall, 2021) in several ways. The performances required of students in performance-based assessment should be responsive to students' cultural differences and connected with students' lives in some way (Hood, 1998; Solano-Flores & Nelson-Barber, 2001). Moreover, assessment questions and tasks should draw on students' linguistic repertoires, as language mediates students' participation in assessment (Bailey & Durán, 2020).

There is plenty of room for improvement in future assessment developments. High-quality tasks that are a direct match with intended learning goals and rich enough to elicit student thinking are very difficult to craft. An improvement in the quality of assessment questions and tasks would be a huge benefit to learners. Needed is ongoing professional development for educators who design classroom assessments and for assessment vendors who design external assessments, as well.

Formative Assessment Processes and Tools

Formative assessment supports learners' processes, motivation, attention, engagement, effort, metacognition, and self-regulation (Andrade et al., 2019; Heritage & Wylie, 2020). As shown in the previous section, educational assessments that serve learners are primarily formative (Brookhart & DePascale, in press; Shepard et al., 2018). "Formative" describes an assessment purpose—in this case informing learning for both students and teachers—and not a particular assessment instrument, since many assessments can be used either formatively (to inform learning) or summatively (to certify or report learning).

The instrument, tool, or method used for formative assessment will be most effective for student learning if the question or task is clearly matched to the learning goal; if the criteria are clear and provided in a form that students can use, with training and instruction on how to use them; and if the assessment, whether formal or informal, is deployed in a process that supports the formative learning cycle. For example, a teacher might pause in an instructional sequence and have students self-assess their work using criteria, then provide opportunity for revision during which students can improve their work and deepen their learning (Chen et al., 2017).

Rubrics are a common way to present criteria and performance level descriptions in a form that students can use for self- and peer assessment and that teachers can use for providing feedback and deciding on next instructional moves. Research in higher education has shown rubrics help make the criteria for good work explicit for students (Jonsson, 2014; Nordrum et al., 2013) and students use rubrics for this purpose (Andrade and Du, 2005; Garcia-Ros, 2011). Similar conclusions have been drawn from research on rubrics in basic education (Brookhart, 2024). Other tools that present criteria in forms students can use include checklists or other lists of criteria, models, and demonstrations. The key seems to be that the criteria are available in a usable form for students, not necessarily that they are rubrics, and that students have instruction and practice in how to use them (Chen et al., 2017; Panadero & Alonso-Tapia, 2013).

Importantly, these reviews and studies have shown that without instruction and guidance in how to use rubrics, and opportunities to do so, students may misunderstand or misuse them. In the terms I have been using in this chapter, the process and tools need to work together to help students regulate their learning.

An obvious future assessment development that would prove useful in improving the use of formative assessment processes and tools is enhanced professional development for both in-service teachers, preservice teachers, and teacher education faculty. There is a lot of rhetoric around formative assessment in schools and teacher education institutions, but programs that do this well are still rare (Brookhart, 2017). Additional developments in the design and use of classroom formative assessment tools and processes would be useful, too, especially focused on strategies where students are the agents of their own assessment (Lee et al., 2021). As the quality of classroom formative assessment information improves, it can be added to the kind of data reviews many schools already do, allowing more specific, equitable, and effective learning diagnoses and instructional remedies than such reviews currently support (Oláh et al., 2010). Safir and Dugan (2021) call this "street data."

Assessment Scores and Feedback

In the previous section, "Feedback and the Opportunity to Use It" was listed as one of the things learners need. Feedback is included in this section as well because it also is something assessment can provide to meet learners' needs. Quantitative scores and qualitative feedback both produce information that is descriptive of current learning status, correlated with learning goals, informational for taking next steps in learning, and connected to the student and their work (Brookhart, 2018; Shute, 2008). Feedback can have a powerful effect on student learning. However, not all feedback is effective in every case (Shute, 2008). In addition, effective feedback can differ markedly depending on the subject matter and the age and level of the student (Smith & Lipnevich, 2018).

Recent reviews of the feedback literature find that in general, the most effective feedback is descriptive information that feeds into the formative learning cycle by helping students understand the current quality of their work and making suggestions for steps they can take to improve (Hattie & Timperley, 2007; Shute, 2008; Van der Kleij, Feskens, & Eggen, 2015). Attention to recent reviews is important because the feedback literature extends back farther in time than many other research literatures in education; early research using behaviorist theoretical frameworks gave way to more cognitive and then sociocognitive and sociocultural models (Lipnevich & Panadero, 2021). As the definition of feedback changed from meaning the simple knowledge of results (right/wrong) needed for behavior-based studies of feedback to including the descriptions and suggestions needed to help students navigate the formative learning cycle, studies began to show increased effectiveness for feedback (Brookhart, 2018).

Future developments in assessment, therefore, should concentrate on equipping educators to provide appropriate feedback—typically descriptive comments based on shared success criteria, but sometimes scores, depending on the learning goal—at moments in an instructional sequence when acting on that feedback would move learning forward. It may be possible to use artificial intelligence to assist in this task. Also, equipping educators to craft and share (or co-create with students) clear success criteria matched to the learning goal will be key to moving forward, because effective feedback is based on those criteria. Future development could also include adding to the repertoire of available self- and peer assessment strategies.

Work is also being done on computer-based cognitive tutoring, which includes feedback to students and also dynamic cognitive modeling to provide feedback to the cognitive tutor itself (Ritter et al., 2007). To serve learners well, feedback from externally-produced learning software deployed to individual students in the classroom will need to be mediated by teachers designing lessons in the context of a classroom learning culture—as opposed to, for example, just having individual students sit in front of computers.

Regarding enhancements in feedback from large-scale assessment that might ultimately serve learners, including attention to formative uses of that feedback, much work is currently being done in developing results reports with multiple pieces of information (Zenisky et al., in press), sometimes suggestive of additional instructional materials (e.g., Smarter Balanced, n.d.). Work is also being done on assessments that use cognitive diagnostic models and report diagnostic classifications based on probabilistic data (Bradshaw & Levy, 2019). The educators who use these reports need professional development to understand them and use them well.

Learning Progressions

Learning progressions are descriptions of hypothesized, and often empirically tested, increasingly sophisticated student understanding that result from ordered steps of instruction in school subjects (Mosher, 2022). Learning progressions help teachers interpret student thinking and learning and engage students in richer, more equitable learning experiences (Alonzo & Elby, 2019; Shepard, 2018). Of course, there are not enough educational psychologists in the world to develop an empirically verifiable representation (Graf & van Rijn 2016) of how children learn and develop in every domain taught in school, but in domains where they are available learning progressions help make interpreting assessment information, providing feedback, and supporting next steps more precise.

Learning progressions are particularly helpful as teachers create or select assessment questions and tasks, give feedback on student work, and plan next steps in instruction. Learning progressions can help teachers understand what differences to expect in students' responses to classroom formative assessments (Confrey, 2019). If the grain size of the descriptions in the ordered steps or levels in a learning progression are small enough to support lesson-sized decisions about what kind of growth typically comes "next" in students' understanding of a concept, assessment can support those decisions. Then students can receive appropriately scaffolded tasks and more targeted feedback, even if they are not aware of the learning progression but especially if they are (Rablin, 2024).

One potentially productive avenue of research and development has been made possible by advances in technology. Because of the internet, classroom walls are more porous than they once were, and this means that school-university partnerships that pair content-area teachers with external assessment researchers are possible in real time. Several programs of research have put together the formative benefits of a learning progression, the assessment design expertise of university researchers, and teacher management of classroom learning (e.g., Confrey & Toutkousian, 2019; Wilson & Draney, 2004; Wilson & Lehrer, 2021), as assessments designed outside the classroom are used inside the classroom, with information flowing both ways. A benefit of this kind of research is that it fosters advances in both learning progressions and assessment at the same time.

Section Summary

To say assessment serves learners when it is formative is almost a tautology, amounting to saying assessment serves learners when assessment informs their learning. Nevertheless, this seems to be a point that needs to be made (Shepard et al., 2018). This section drilled one step deeper, to show how assessment informs learners best when it helps them focus their attention on a learning goal and activate regulatory processes to move closer to it. Specific features of assessment that research has shown to facilitate this include designing high-quality assessment and learning tasks that match learning goals, activating the formative learning cycle in classroom instruction and assessment, providing students feedback that moves learning forward, and where possible using learning progressions to design and interpret instruction and assessment.

How does validation shift as the emphasis in assessment shifts to serving learners?

The phrase "to serve learners" is a purpose statement. Assessment purpose invokes validity because it has to do with inquiring into the appropriateness of interpretations and uses of assessment information (Kane, 2013, 2016). Making claims that students need assessment to function in certain ways and to have certain characteristics raises empirical questions that are at root validity questions.

Validation arguments (Kane, 2013, 2016) typically support interpretation and use (sometimes called meaning and impact, Lederman, 2023) in assessment. Several authors have pointed out that as the emphasis shifts from informing educators and administrators to serving learners, the emphasis in validation shifts proportionately from interpretation/meaning toward use/impact (Hopster-den Otter et al., 2019; Kane & Wools, 2020; Lederman, 2023; Moss, 2016) and, indeed without assessment use score interpretation becomes a moot point (Sireci, 2016).

Lederman (2023) argued more emphasis on use or impact could be incorporated into validation work to disrupt assessment-based racial injustice. Because score meaning may differ for different groups, emphasizing impact is the key to pursuing racial justice in assessment. Some of the same arguments work for supporting learners in assessment, as well. Emphasizing use of assessment information increases the importance of learners' achievement and motivational outcomes relative to score meaning—in other words, it elevates the purpose of serving learners.

When validation expands into questions of use and impact, the contexts of the organizations that use data and the resources of their educational professionals become relevant sources of evidence for validation (Moss, 2016). Moss's argument about educators' use of tests in their schools and districts could easily be extended to students' use of assessments in their classroom learning cultures. If the claim is made that an assessment serves learners, the validity of the assessment information depends in part on students' knowledge and interpretation of that assessment information. For example, for a self-assessment used at a pause point in a classroom unit of instruction, students' understanding of the learning goal the assessment is meant to inform and the criteria by which they will know where they are and where to go next will affect the degree to which they can take the assessment information on board as feedback and use it productively. If students do not have a clear enough concept of what they are trying to learn, information about where they are now will have limited usefulness to them and thus limited validity for supporting them as learners.

Validation can regard students as learners, examinees, or contestants (Dorans, 2012). As assessment shifts from the decontextualized measurement characteristic of conventional large-scale tests to measurement contextualized as part of a student's formative learning cycle—that is, as assessments develop to better serve learners—students shift from being examinees to being learners and sometimes contestants. Dorans (2012) posits that in the case of large-scale assessment, from a contestant perspective—in addition of course to simply wanting to win—students would see assessment interpretation and use as valid if the assessment created a fair race characterized by reliable outcomes, acceptable scoring, clear rules, and empirically verified interpretations. This chapter has shown that, in the case of classroom learning, from a learning perspective students would see assessment interpretation and use as valid if the assessment: (1) was situated in a supportive classroom learning culture that included honoring their own sociocultural context, (2) was situated in a lesson or series of lessons for which they understood clear learning goals and criteria for success, (3) was used as part of their participation in the formative learning cycle, and (4) provided feedback that moved learning forward and was coupled with an opportunity to do so. I would argue that for classroom formative assessment, where results do not need to generalize and often are not scores at all, the latter are the most important validation criteria.

The learner perspective on validity and the contestant perspective on validity are broadly comparable to mastery and performance goal orientations (Maehr & Zusho, 2009). Achievement goal theory posits two potential reasons why students may be motivated to learn. Briefly, students with a mastery goal orientation are motivated because they want to master the content (the learner perspective); students with a performance goal orientation are motivated to demonstrate to others that they are smart, or smarter than their classmates (performance-approach goals, the contestant perspective) or to avoid seeming not to be smart (performance-avoidance). Both mastery goals and performance-approach goals have small positive effects on academic achievement (Maehr & Zusho, 2009; Senko, 2019). Depending on the subject matter and specific classroom culture, students adopt either or both of these goal orientations. Therefore, it seems prudent that as the validation literature expands to include students' perspectives, both the learner and contestant perspectives should be considered.

Summary, Recommendations, and Conclusions

Summary and Principles

In this chapter, I have shown that first, learners need a supportive classroom learning culture that honors their own culture and leverages their own funds of background knowledge and experience. Second, learners need a clear understanding of learning goals and success criteria. Clear and explicit communication of learning goals can be a means for promoting equity because all students have access to the goal and criteria for good work, not just those whose background allows them to infer these things from a lesson where they are only implicit. Principle 6, "Assessment equity requires fairness in design of tasks and their adaptation to permit their use with respondents of different backgrounds, knowledge, and experiences," is clearly implicated.

Third, learners need assessment to be situated in the formative learning cycle, in which students set their sights on a learning goal and actively pursue it using their cognitive, affective, and behavioral resources. Principle 3: "Assessment design supports learners' processes, such as motivation, attention, engagement, effort, and metacognition," is clearly implicated. Fourth, as they are navigating the formative learning cycle, learners need effective, high-quality feedback and opportunities to use it. Principle 5, "Feedback, adaptation, and other relevant instruction should be linked to assessment experiences," is clearly implicated.

I have argued that to meet these learner needs, those who design and use assessments should give attention to the quality of assessment questions and tasks, the use of formative assessment processes and tools, the quality and usefulness of both scores and descriptive feedback that result from assessment, and the use of learning progressions where appropriate. The next section interprets these points into more focused recommendations for those who develop and use assessment to serve learners.

Recommendations for Developing Educational Assessments to Serve Learners

In deciding how to translate the chapter's discussion into recommendations, I had to make a grain size decision. In keeping with the size and scope of the chapter, I offer these recommendations at a middle grain-size level. For example, Recommendation 1 could say "More student voice," which might be too general to be helpful, or it could include a list of many different practical ways to incorporate student voice into the design, interpretation, and use of various different kinds of assessments in various subject matters and grade levels; a how-to list of this sort is definitely needed but is beyond the scope of this chapter. Thus, I offer these recommendations that, to me, follow from the chapter's discussion and are specific enough to at least move the conversation forward and suggest both future research and future professional development for those involved in assessment. Table 1 presents a list of assessment developments that, in my view, will enable the development of educational assessments that serve learners better in the future.

Table 1. *Developing Educational Assessments to Serve Learners*

Recommendation	References
Better understanding of the results of increasing student voice and choice in assessment on the interpretation and use/impact of assessment results	Andrade et al. (2019); Heritage & Wylie (2020); Hood (1998); Ladson-Billings (2014); Randall (2021); Solano-Flores & Nelson-Barber (2001); Taylor & Nolen (2022)
Increase in the quality of assessment questions and tasks, both classroom and external	Brookhart & DePascale (in press); Carless (2015); Chen et al. (2017); Heritage & Wiley (2020); Wiggins (1998); Wilson (2004)
Increase in the quality and usefulness (to learners) of various kinds of scoring schemes and feedback comments and understanding of the conditions under which to deploy them	Hattie & Timperley (2007); Jonsson & Panadero (2018); Shute (2008); Van der Kleij, Feskens, & Eggen (2015)
Increase in the repertoire of available formative assessment strategies for both teachers and students	Bailey & Heritage (2018); Heritage & Wylie (2020); Brookhart (2017)
Increase in the use of classroom assessment results, especially classroom formative assessment, and a concomitant respect and understanding of the place of high-quality classroom assessment in the learning experience for learners	Safir & Dugan (2021)
More judicious use of external assessment results: using accountability and other summative assessment results only to raise questions about learning, not to answer them; developing more diagnostic external "formative" assessments, perhaps using emerging technology and measurement methods	Bradshaw & Levy (2019); Shepard et al. (2018); Zenisky et al. (in press)
More research and development on promising programs that mix internal, situated classroom work with external assessment, if and only if accompanied by the development of learning progressions and other tools, deeply criterion-referenced, and tied to students' learning experiences	Confrey & Toutkousian (2019); Wilson & Draney (2004); Wilson & Lehrer (2021)

Each recommendation in the list is accompanied by some of the citations that inspired and support it. As the references show, some work at least at the concept level has begun for each of these recommendations, and in many cases research and development has begun, as well.

Conclusion

In this chapter, I have tried to shine a light on aspects of educational assessment that might help move the needs of learners closer to the center of assessment. Perhaps it should be surprising that considering the needs of learners has not always been the first principle of all educational assessment. However, it clearly has not (Dorans, 2012; Shepard, 2000). This chapter stands on the shoulders of others who would move the needs of learners into a central place in assessment, and I have tried to cite a wide variety of work to demonstrate that. Often that meant that large bodies of work were just mentioned and cited, or ideas that could be whole chapters in themselves just received a paragraph. I hope this chapter prompts readers to pursue these thoughts further.

References

Alonzo, A. C., & Elby, A. (2019). Beyond empirical adequacy: Learning progressions as models and their value for teachers. *Cognition and Instruction, 37*(1), 1–37. https://doi.org/10.1080/07370008.2018.1539735

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Andrade, H. L., Bennett, R. E., & Cizek, G. J. (Eds.). (2019). *Handbook of formative assessment in the disciplines.* Routledge.

Andrade, H. L., & Brookhart, S. M. (2019). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice, 27*(4), 350–372. https://doi.org/10.1080/0969594X.2019.1571992

Andrade, H. L., Brookhart, S. M., & Yu, E. C. (2021). Classroom assessment as co-regulated learning: A systematic review. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.751168

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation, 10*, 1–11. https://scholarsarchive.library.albany.edu/edpsych_fac_scholar/2

Bailey, A. L., & Durán, R. (2020). Language in practice: A mediator of valid interpretations of information generated by classroom assessments among linguistically and culturally diverse students. In S. M. Brookhart & J. H. McMillan (Eds.), *Classroom assessment and educational measurement*. Routledge. https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9780429507533-4/language-practice-alison-bailey-richard-dur%C3%A1n

Bailey, A. L., & Heritage, M. (2018). *Self-regulation in learning: The role of language and formative assessment.* Harvard Education Press.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Bradshaw, L., & Levy, R. (2019). Interpreting probabilistic classifications from diagnostic psychometric models. *Educational Measurement: Issues and Practice, 38*(2), 79–88. https://doi.org/10.1111/emip.12247

Brookhart, S. M. (2017). Formative assessment in teacher education. In D. J. Clandinin & J. Husu (Eds.), *International handbook of research on teacher education* (pp. 927–943). Sage.

Brookhart, S. M. (2018). Summative and formative feedback. In A. A. Lipnevich & J. K. Smith, (Eds.), *The Cambridge handbook of instructional feedback* (pp. 52–78). Cambridge University Press.

Brookhart, S. M. (2024). Using rubrics in basic education: A review and recommendations. *Estudos em Avaliação Educacional, 35*, Article e10803. https://doi.org/10.18222/eae.v35.10803

Brookhart, S. M., & DePascale, C. A. (in press). Assessment to inform teaching and learning. In Cook, L., & Pitoniak, M. (Eds.), *Educational measurement* (5th ed.). Oxford University Press.

Brookhart, S. M., & McTighe, J. (2017). *The formative assessment learning cycle.* ASCD.

Chen, F., Lui, A. M., Andrade, H., Valle, C., & Mir, H. (2017). Criteria-referenced formative assessment in the arts. *Educational Assessment, Evaluation and Accountability, 27*, 297–314. https://doi.org/10.1007/s11092-017-9259-z

Confrey, J. (2019, May). *A synthesis of research on learning trajectories/progressions in mathematics.* OECD EDU/EDPC(2018)44/ANN3. https://www.oecd.org/education/2030-project/about/documents/A_Synthesis_of_Research_on_Learning_Trajectories_Progressions_in_Mathematics.pdf

Confrey, J., & Toutkoushian, E. (2019). A validation approach to middle-grades learning trajectories within a digital learning system applied to the "Measurement of Characteristics of Circles." In J. Bostic, E. Krupa, and J. Shih (Eds.): *Quantitative measures of mathematical knowledge: Researching instruments and perspectives* (pp. 67–92). Routledge.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4). 438–481. https://doi.org/10.3102/00346543058004438

Dorans, N. J. (2012). The contestant perspective on taking tests: Emanations from the statue within. *Educational Measurement: Issues and Practice, 31*(4), 20–37. https://doi.org/10.1111/j.1745-3992.2012.00250.x

Esteban-Guitart, M., & Moll, L. C. (2014). Lived experience, funds of identity and education. *Culture & Psychology, 20*(1), 70–81. https://doi.org/10.1177/1354067X13515940

Garcia-Ros, R. (2011). Analysis and validation of a rubric to assess oral presentation skills in university contexts. *Electronic Journal of Research in Educational Psychology. 9*(3), 1043–1062.

Graf, E. A., & van Rijn, P. W. (2016). Learning progressions as a guide for design: Recommendations based on observations from a mathematics assessment. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 165–189). Taylor and Francis.

Hattie, J. A. C. (2009). *Visible learning*. Routledge.

Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.

Hattie, J., Crivelli, J., Van Gompel, K., West-Smith, P., & Wike, K. (2021). Feedback that leads to improvement in student essays: Testing the hypothesis that "Where to next" feedback is most powerful. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.645758

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heritage, M. (2013). Gathering evidence of student understanding. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 179–195).

Heritage, M., & Wylie, E. C. (2020). *Formative assessment in the disciplines.* Harvard Education Press.

Hood, S. (1998). Culturally responsive performance-based assessment: Conceptual and psychometric considerations. *Journal of Negro Education, 67*(3), 187–196. https://doi.org/10.2307/2668188

Hopster-den Otter, D., Wools, S., Eggen, T. J. H. M., & Veldkamp, B. P. (2019). A general framework for the validation of embedded formative assessment. *Journal of Educational Measurement, 56*(4), 715–732. https://doi.org/10.1111/jedm.12234

Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2021). The quality of assessment tasks as a determinant of learning. *Assessment & Evaluation in Higher Education, 46*(6), 943–955. https://doi.org/10.1080/02602938.2020.1828268

James, M., Black, P., Carmichael, P., Conner, C., Dudley, P., Fox, A., Frost, D., Honour, L., MacBeath, J., McCormick, R., Marshall, B., Pedder, D., Procter, R., Swaffield, S., & Wiliam, D. (2006). *Learning how to learn: Tools for schools.* Routledge.

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education, 39*(7), 840–852. https://doi.org/10.1080/02602938.2013.875117

Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 531–553). Cambridge University Press.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Kane, M. T. (2016). Explicating validity. *Assessment in Education: Principles, Policy & Practice, 23*(2), 198–211. https://dx.doi.org/10.1080/0969594X.2015.1060192

Kane, M. T., & Wools, S. (2020). Perspectives on the validity of classroom assessments. In S. M. Brookhart & J. H. McMillan (Eds.), *Classroom assessment and educational measurement* (pp. 11–26). Routledge. https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9780429507533-2/perspectives-validity-classroom-assessments-michael-kane-saskia-wools

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia—Social and Behavioral Sciences, 31*, 486–490. https://doi.org/10.1016/j.sbspro.2011.12.091

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: A.k.a the remix. *Harvard Educational Review, 84*(1), 74–84. https://doi.org/10.17763/haer.84.1.p2rj131485484751

Lederman, J. (2023). Validity and racial justice in educational assessment. *Applied Measurement in Education, 36*(3), 242–254. https://doi.org/10.1080/08957347.2023.2214654

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K–12 education: A systematic review. *Applied Measurement in Education, 33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Leighton, J. P., Chu, M-W., & Seitz, P. (2013). Errors in student learning and assessment. In R. W. Lissitz (Ed.), *Informing the practice of teaching using formative and interim assessment: A systems approach* (pp. 185–208). Information Age Publishing.

Lipnevich, A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education, 6*(720195). https://doi.org/10.3389/feduc.2021.720195

Mason, B. J., & Bruning, R. (2001). *Providing feedback in computer-based instruction: What the research tells us.* CLASS Research Report No. 9. Center for Instructional Innovation, University of Nebraska-Lincoln. https://www.researchgate.net/publication/247291218_Providing_Feedback_in_Computer-based_Instruction_What_the_Research_Tells_Us

Maehr, M. L., & Zusho, A. (2009). Achievement goal theory: Past, present, and future. In K. Wentzel & D. Miele (Vol. Eds.), *Handbook of Motivation at School: Vol 1*, (pp. 77–104). Taylor & Francis.

McMillan, J. H. (2018). *Using students' assessment mistakes and learning deficits to enhance motivation and learning.* Routledge.

Mosher, F. A. (2022). *Learning progressions. Routledge Resources Online—Education.* https://doi.org/10.4324/9781138609877-REE115-1

Moss, C. M. (2022). Learning targets and success criteria. Routledge *Resources Online—Education.* https://doi.org/10.4324/9781138609877-REE39-1

Moss, P. A. (2016). Shifting the focus of validity for test use. Assessment in Education: Principles, Policy & Practice, 23(2), 236–251. https://dx.doi.org/10.1080/0969594X.2015.1072085

Nordrum, L., Evans, K., & Gustafsson, M. (2013). Comparing student learning experiences of in-text commentary and rubric-articulated feedback: Strategies for formative assessment. *Assessment & Evaluation in Higher Education. 38*, 919–940. https://doi.org/10.1080/02602938.2012.758229

Oláh, L. N., Lawrence, N. R., & Riggan, M. (2010). Learning to learn from benchmark assessment data: How teachers analyze results. *Peabody Journal of Education, 85*(2), 226–245. https://doi.org/10.1080/01619561003688688

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practice connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology, 11*(2), 551–576. http://dx.doi.org/10.14204/ejrep.30.12200

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019). Using formative assessment to influence self- and co-regulated learning. *European Journal of Psychology of Education, 34*(3), 535–557. https://link.springer.com/article/10.1007/s10212-018-0407-8

Rablin, T. (2024). *Hacking student motivation: 5 assessment strategies that boost learning progression and build student confidence.* Times 10 Publications.

Randall, J. (2021). "Color-neutral is not a thing: Redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educational Measurement: Issues and Practice, 40*(4), 82–90. https://doi.org/10.1111/emip.12429

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher, 16*(9), 13–20+54. https://doi.org/10.3102/0013189X016009013

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review, 14*(2), 249–255. https://doi.org/10.3758/BF03194060

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714

Safir, S., & Dugan, J. (2021). *Street data: A next-generation model for equity, pedagogy, and school transformation.* Corwin.

Senko, C. (2019). When do mastery and performance goals facilitate academic achievement? *Contemporary Educational Psychology, 59*. https://doi.org/10.1016/j.cedpsych.2019.101795

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14. https://doi.org/10.3102/0013189X029007004

Shepard, L. A. (2018a). Learning progressions as tools for assessment and learning. *Applied Measurement in Education, 31*(2), 165–174. https://doi.org/10.1080/08957347.2017.1408628

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018b). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice, 37*(1), 21–34. https://doi.org/10.1111/emip.12189

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sireci, S. G. (2016). On the validity of useless tests. *Assessment in Education: Principles, Policy & Practice, 23*(2), 226–235. http://dx.doi.org/10.1080/0969594X.2015.1072084

Slade, S., & Griffith, D. (2013). A whole child approach to student success. *KEDI Journal of Educational Policy, Special Issue,* 21–35. https://www.kedi.re.kr/eng/kedi/bbs/B0000005/list.do?menuNo=200067

Smarter Balanced Assessment Consortium. (n.d.). *Understanding the Smarter Balanced reporting system for educators.* https://youtu.be/sYMY4CJUO6g

Smith, J. K., & Lipnevich, A. A. (2018). Instructional feedback: Analysis, synthesis, and extrapolation. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge Handbook of Instructional Feedback* (pp. 591–603). Cambridge University Press.

Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching, 38*(5), 553–573. https://doi.org/10.1002/tea.1018

Stobart, G. (2018). Becoming proficient. In A. A. Lipnevich & J. K. Smith, (Eds.), *The Cambridge handbook of instructional feedback* (pp. 29–51). Cambridge University Press.

Taylor, C. S., & Nolen, S. B. (2022). *Culturally and socially responsible assessment: Theory, research, and practice.* Teachers College Press.

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* Jossey-Bass.

Wiliam, D., Fisher, D., & Frey, N. (2024). *Student assessment: Better evidence, better decisions, better learning.* Corwin.

Wilson, M. (2004a). A perspective on current trends in assessment and accountability: Degrees of coherence. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (pp. 272–283). 103rd Yearbook of the National Society for the Study of Education, Part II. University of Chicago Press.

Wilson, M., & Draney, K. (2004b). Some links between large-scale and classroom assessments: The case of the BEAR Assessment System. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (pp. 132–154). 103rd Yearbook of the National Society for the Study of Education, Part II. University of Chicago Press.

Wilson, M., & Lehrer, R. (2021). Improving learning: Using a learning progression to coordinate instruction and assessment. *Frontiers in Education, 6*:654212. https://doi.org/10.3389/feduc.2021.654212

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

Zenisky, A., O'Donnell, F., & Hambleton, R. K. (in press). Reporting scores and other results. In Cook, L., & Pitoniak, M. (Eds.), *Educational measurement* (5th ed.). Oxford University Press.

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and overview. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge.

About the author

Susan M. Brookhart, Ph.D., is Professor Emerita in the School of Education at Duquesne University and an independent educational consultant. She was the 2007-2009 Editor of Educational Measurement: Issues and Practice and is currently an Associate Editor of Applied Measurement in Education. She is the author or co-author of over 100 articles, chapters, and books on classroom assessment, teacher professional development, and evaluation. She was named the 2014 Jason Millman Scholar by the Consortium for Research on Educational Assessment and Teaching Effectiveness (CREATE) and was the recipient of the 2015 Samuel J. Messick Memorial Lecture Award from ETS/TOEFL. Dr. Brookhart's research interests include the role of both formative and summative classroom assessment in student motivation and achievement, the connection between classroom assessment and large-scale assessment, and grading. Dr. Brookhart received her Ph.D. in Educational Research and Evaluation from The Ohio State University, after teaching in both elementary and middle schools.

About the Study Group

The Study Group exists to advance the best of artificial intelligence, assessment, and data practice, technology, and policy; uncover future design needs and opportunities for educational systems; and generate recommendations to better meet the needs of students, families, and educators.

Date of Publication
April 2026

Citation
Brookhart, S. M. (2025). Developing educational assessments to serve learners. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for assessment in the service of learning, Volume I: Foundations for assessment in the service of learning.* University of Massachusetts Amherst Libraries.