

Assessment Literacy: Assessment Literacy Explanation: Concepts, Principles, Checklist, and Strategies

Grounding Assessment Literacy

Marvin E. Smith, Stefinee Pinnegar, Annela Teemant

Introduction

During the past decade, two major themes have dominated concerns for improving public education in the United States: (1) increases in the diversity of students in U.S. schools and (2) results for American students on international comparisons of student performance. The first theme reflects the changing demographics of the population of the United States and its impact on schooling. Over the past decade, there has been a dramatic increase in the number of English-

as-a-second-language (ESL) students, from both immigrant and international backgrounds, entering all levels of American schools (Rosenthal, 1996; Steward, 1991). More than six million children in the United States do not use English as their native language at home (Rosenthal, 1996).

The second theme began to receive national attention in 1983 with the publication of *A Nation at Risk*. It continued with the development of the National Educational Goals and Goals 2000: Educate America Act (H.R. 1804) (Lam, 1993; Stansfield, 1994). Educational reforms attempting to respond to this concern focus on raising educational standards to a "world class level" (Stansfield, 1994) and implementation of high-stakes assessments targeted at school accountability. As Short noted (1993), "assessment dominates the educational reform dialogue" (p. 630). In fact, national policies have emphasized testing as the primary method for states and districts "to reshape teaching and to effect learning in the schools" (Stansfield, 1994, p. 43).

However, the interaction of these two themes poses a significant problem for reform. The focus on assessment as a strategy for encouraging educational reform can place ESL students at special risk. Bernhardt, Destino, Kamil, and Rodriguez-Munoz (1985) argued these students "are in double jeopardy when confronted with assessment of any type" because they are "forced into demonstrating knowledge in a language over which they have only partial . . . control" (p. 6).

This interaction between content and language presents teachers with the challenge of determining the role of language knowledge and content knowledge in documenting difficulties in student learning (Short, 1993; Rosenthal, 1996). Teachers of ESL students have the added responsibility of using assessment strategies that enable these students to demonstrate what they do know and to make judgments about student performances in ways that support effective teaching and learning.

The purpose of this course is to support teachers of ESL students in gaining knowledge about assessment that can help them respond to the dilemmas of assessment-driven educational reforms among linguistically diverse students.

This knowledge is an essential part of the knowledge base for teaching. More than anything else, the public must be able to rely on the judgment of teachers, and those judgments must be appropriate for all students, including second language learners.

The purpose of this reading is to introduce our view of Assessment Literacy and provide a theoretical foundation for our perspective. The

Literacy Chart includes six principles organized by three concepts. These concepts summarize the imperative:

Assessment must be-

Useful for stakeholders,

Meaningful for its purposes, and

Equitable for all students.

The six principles in the assessment chart define and identify essential elements of the three concepts. The checklist items offer questions teachers can ask themselves to prompt consideration of important issues associated with the six principles. The assessment strategies describe particularly important ways of applying the principles in assessing language minority students.

The remainder of this reading begins with detailed explanations of the meanings and implications of our concepts of Assessment

Literacy Chart. Second, we address the importance of foundational perspectives on knowing, learning, teaching, and assessing that can help us create coherent classroom practices. Third, we provide a comparison of two fundamental models of assessment that are coherent with competing educational perspectives. Finally, we elaborate on assessment strategies that are appropriate for the needs of linguistically diverse students.

Useful

Usefulness weighs the educative value of an assessment against the practical consideration of feasibility and efficiency. Useful assessment is both doable and informative. But an assessment must do more than merely justify an educational decision. It must be educative. It must capture and communicate judgments about student work that show students how to get better at learning the things they are being assessed on. It should also provide teachers with information that will help them improve their teaching and assessment.

Assessment that is useful provides educative feedback. Feedback is educative when it strengthens and supports the learning process rather than interferes with or distorts it. It is often more descriptive than evaluative. When feedback is educative, it identifies for both

the teacher and the student where they must go and what they must do next to move learning forward. Such feedback helps students develop an understanding of and a commitment to what they are trying to accomplish. It also provides a vision of what they should do next to become better at a particular skill, improve their understanding of particular content, or develop more complex thinking.

Educative feedback provides teachers with information about how the assessment itself could be made more useful, meaningful, and equitable. Feedback can also be educative for parents and communities about the substance and quality of teaching and learning occurring in schools.

Educative feedback is useful when it supports teachers and learners in making decisions. Decisions that follow assessment always have educational consequences for both teacher and learner. The decision to move to the next step

or return to an earlier one has consequences for the ultimate learning of the students. Decisions to place students in new groups, contexts, or programs are never insignificant. The more clearly an assessment meets the criteria of usefulness, meaningfulness, and equitability the more likely decisions flowing from the assessment will be sound.

Because teaching occurs in arenas of limited resources and unlimited potential, useful assessments must support teachers in balancing both of these factors. This means assessments must be practical. No matter how brilliant or educative an assessment design, if it is not feasible given the circumstance and situation of an individual teacher the power of the assessment will be limited. When the educative potential is truly significant, it is the teacher's responsibility to determine how it might become feasible: How might processes, performances, or products be altered in ways that make the assessment feasible without altering its usefulness, meaningfulness, or equitability?

Judgments of feasibility are always founded in perceptions of both teachers and learners. These judgments emerge when available resources are weighed against those needed to engage, conduct, or complete the planned assessment. We usually think of feasibility as a teacher judgment concerning a particular format or timing for an assessment. However, feasibility can also be a reason why a learner refuses or only half-heartedly engages in an assessment. The learner's motivation is based on bridging the gap between expected benefits and required efforts. When either the student or the teacher perceives the educative quality and benefit of an assessment to be worthwhile, they are more likely to find a way to make it feasible.

Making assessments practical also requires attention to efficiency. Arguments that an assessment is not practical are often founded in concerns about efficiency. However, adjustments in assessment designs that improve efficiency can occur both inside and outside the assessment. Efforts to streamline various aspects of the assessment process can both improve the educative potential of assessment and reduce assessment costs in time and other resources. When these two competing demands become complementary, assessments can be more useful.

One way of improving the efficiency of testing processes is to streamline reporting procedures so that reports are easily prepared and helpful to both teachers and students. Other ways of improving efficiency might include limiting or guiding choices about what to include in a portfolio. Using a multiple-choice format instead of an essay test or an oral interview instead of a multiple-choice test might improve the efficiency of assessments with ESL students. Ironically, sometimes making a test more efficient for a learner may make a test less efficient for the teacher and consequently less feasible.

Overall, the perceived benefit to the learner, quality of feedback, support for decision making, and strength in meeting learning goals will determine students' and teachers' perceptions of the usefulness of a particular assessment.

Meaningful

Assessment is meaningful when it can guide all stakeholders in the educational process to make decisions that will improve educational opportunities and fully develop student potential. This happens when assessment meets its purposes. In particular, assessment should be meaningful to those most centrally involved in educational improvement—teachers and students. Assessments should provide feedback that can lead students and teachers to accurately identify student progress on learning goals they accept and care about. Assessment should provide teachers with information they find meaningful as they design curriculum and classroom tasks, make judgments about student progress, and guide students to meet learning goals. Educated and thoughtful teacher judgment in the design and use of assessments is a central ingredient for making them meaningful.

Assessment information is meaningful when it is relevant to the goals teachers and learners have set. In designing curriculum teachers have to be concerned about student progress in learning the important concepts, skills, and processes of particular disciplines. They must also be concerned about students' progress in general performance areas like literacy, numeracy, and thinking that cut across discipline boundaries and influence every student performance. In addition, teachers are concerned with whether or not students are developing dispositions and attitudes that

will enable them to participate successfully as adult members of communities beyond the classroom. Therefore, meaningful assessments will provide teachers relevant information about where students are in their growth and development in content knowledge, literacy, numeracy and thinking skills, and character development. In the language of Inclusive Pedagogy, assessment will provide information that is relevant for each of the critical learning domains: cognitive, academic, social, affective, and linguistic.

The content of assessments should provide insight and information about each of these areas. Because teachers will not be able to assess everything in every area all the time, they must carefully select the focus of particular assessments and plan for a collection of assessments that provide a complete picture of students' learning. Teachers have various resources available to help them identify important content goals, including national, state, and local standards for content areas, for special population students, and for other learning goals such as character development. Teachers must think through the big ideas they think are the most worthy aims in the education of students of a particular age in a specific content area. Once teachers have thought through all that they might teach, they must select those things that are most worthy of everyone's efforts in their classrooms. These big ideas represent the core goals for their curriculum, instruction, and assessment.

Teachers teach students by engaging them in tasks. They make judgments about how students are progressing by observing their performance on those tasks. Just as learning tasks must be relevant, assessment tasks must also be relevant. The challenge is to develop tasks that engage students with language and content in ways that allow teachers to make accurate judgments about their progress, proficiency, and performance in ways that link back to the identified learning goals.

One way to improve the links between important goals, engaging learning activities, and valuable assessment information is to use authentic tasks for both learning and assessment. Authentic tasks can develop and assess student understanding in contexts and situations that make students' performances both highly realistic and interesting. Students may be asked to solve real-world problems, predict unknown outcomes, or identify examples and situations from their own lives. Simulations, experiments, service learning, and activities based on adult work in a particular field are all examples of authentic tasks. However, authenticity alone is not enough. To

be useful in promoting learning, assessment tasks should provide feedback that allows students and teachers to adjust their responses and make informed decisions about next steps. The feedback should help them determine whether or not they are meeting or will meet their goals for learning. Tasks should provide evidence of knowledge of the content, appropriate use of methods, development of skillful craftsmanship, growing sophistication of general and specific skills, and other specific benefits of the learning experience.

Even when content and tasks are highly relevant, assessments are only meaningful when the feedback from them is accurate. Assessment is accurate when results are both valid and reliable. Reliability refers to the dependability of the data upon which judgments about student performance are based. For teacher made paper and pencil assessments, teachers can improve reliability by creating a table of specifications that identify concepts to be tested, tasks for testing them, and thinking levels and language skills required. In this way teachers can check the specifications against their learning goals and use them to guide the construction of assessment. In addition they can make certain several items assess each big idea and that tasks are carefully constructed. Using longer tests and more consistent testing conditions for all test takers provides more reliable results. However, this requirement can be satisfied by allowing all students to have plenty of time and all of the useful tools that might benefit some students. Restricting time and tools to the minimum provides consistent conditions but does so in ways that discriminate against some students. For complex authentic assessments, analytic rubrics and checklists that provide detailed guides for scoring performances improve the reliability of the data.

Reliability is a characteristic of the data on which interpretations and judgments are made. Reliability of assessment data can be jeopardized by the health, mood, motivation, test-taking skills, or general abilities of students. Reliability can also be compromised by the quality of the directions, ambiguities of language, distracting conditions in the environment, interruptions during administration,

biases of the observer, scoring sheet errors, or even bad luck. Teachers can reduce the impact of these factors by attending to conditions that can make assessments more reliable.

Validity is concerned with the claim, judgment, or interpretation made about the student's performance. It refers specifically to the appropriateness of the conclusions, uses, and consequences that follow from an assessment. Validity is always a matter of degree and is always determined in relationship to adequacy of particular evidence for a particular purpose. When making judgments based on assessments, teachers improve validity when they make certain the evidence behind their judgment is sound; try out alternative interpretations or look for disconfirming as well as confirming further evidence; and determine whether, given the consequences, the judgment is reasonable and evidence-supported. When teachers suspect students have difficulties in general learning skills like literacy or numeracy or that they have had only limited opportunities to develop these proficiencies, they should make additional observations and collect additional data using assessment tools that are not so dependent on general skills. Validity includes the trustworthiness of the judgments we make about our students, our curriculum and our instruction. When our judgments are trustworthy they will be more meaningful.

In the real world, we are repeatedly assessed on our ability to do challenging work in unfamiliar contexts and situations. In those settings we are able to ask questions about the purposes, audience, standards, and criteria for our performances. We can quiz and will be quizzed about isolated facts as well as our general comprehension of difficulties or needs or successes. These assessments typically occur both during and at the end of completed projects. In schools, students rarely experience these kinds of assessment. Sometimes they question the purpose of the work we ask them to do. They may not see how assessments relate to their learning and growth. In fact a teacher's assessments and grading system may make students unwilling to put forth needed effort because they are afraid they might look stupid. Or, they may feel success is simply a matter of luck or teacher preference. Some students may be so afraid of failure or looking stupid that they act apathetic or disinterested. By focusing assessments on relevant content and tasks and utilizing educative feedback systems, students increasingly see how to monitor and adjust their performance to reach goals they value. Teachers need to make certain that they select content, learning tasks, and assessment tasks worthy of students' attention. Authentic tasks can help open the learning process to students so that they become aware of their own growth and development. Teachers and students should collect evidence of their learning that is dependable so that relevant and valid feedback and decisions can emerge. When this happens assessment is meaningful.

Equitable

Equitable assessment is clearly fair, but in a different way than most people expect when thinking about fairness. Fairness in education is not like fairness in competitive sports. It does not mean that everybody plays by rules that favor some students over others. It does mean that everybody should be using rules that give every student the same probability of success. In teaching, this means that every student is supported by a more capable other within his or her own zone of proximal development. In assessment, this means that every student has access to assessment tasks that allow them to show what they know and can do. For example, students with limited English writing skills can be assessed on their understanding of important concepts orally, using gestures and movement, or with pictures. This provides them with the opportunity to show learning and to receive comprehensible feedback about how to improve the quality of their learning. Equitable assessment ought to enable all students to achieve classroom goals. Assessments that are equitable promote equal opportunities for all students to grow and develop and encourage improvements in teaching to support their learning.

Open assessment happens when students understand how and on what they will be assessed. Through disclosure of assessment procedures, teachers involve and empower students to engage and succeed in assessment. However, for assessment to be genuinely open, teachers should invite students and others to fully participate in the assessment process. Students can be involved in identifying goals and developing criteria for judging products, thus clarifying exactly what the requirements are and committing to the learning and assessing process. In addition when students

participate in authentic real-world tasks, experts from the community can be invited in to the classroom to make decisions about the quality of student work and provide students with authentic feedback to improve performance.

Appropriate assessment makes certain that content and tasks are meaningful and that feedback and judgments are educative. Assessment clearly based in learning goals and that provides students with feedback that guides their performance is more likely to be equitable and appropriate. However, teachers must also consider fairness

and impact when evaluating their assessment processes. This often requires attending simultaneously to cognitive, academic, social, affective, and linguistic learning goals and how assessment tasks balance those potentially conflicting goals to appropriately meet the needs of students. For example, increasing the authenticity of a task may simultaneously increase the cognitive and linguistic load of a task. Accommodations may be needed to ensure ESL students have access to the task so that the task remains appropriate for all students.

In order to manage a classroom, teachers often make collective judgments about groups of students that enable them to efficiently set behavior boundaries and educational goals. To avoid expectations that are unfair and inappropriate, teachers need to articulate to themselves, perhaps in a journal or log, just what they expect from their students. In this manner teacher expectations become explicit and open to personal reflection and discussion among peers.

Fairness requires that assessment tasks, language, and processes are respectful of gender, culture and linguistic differences present in the classroom. Materials and contexts need to be meaningful to students of all backgrounds. If it appears that only one group of students is showing learning growth, teachers must examine the accuracy of their assessment and teaching strategies for inequities and to identify the causes of unequal outcomes by group.

Impact has to do both with the feedback teachers receive from their assessments and the decisions they make. Assessments always have cognitive, academic, social, affective, and linguistic consequences for students. These consequences constitute the impact of the assessment. For example, teachers may use assessment information to adjust the difficulty of the curriculum, make various accommodations, or fundamentally redesign the assessment. They may find that the structure or nature of a commonly used assessment has taught students to become disinterested in certain valued learning or to react in other unexpected ways. Teachers may see a need to consider how particular assessments produce other positive or negative consequences when they plan future assessments.

When assessments are equitable, negative consequences are minimized and positive ones are emphasized.



This content is provided to you freely by EdTech Books.

Access it online or download it at https://edtechbooks.org/ell_tools/assessment_literacy_o.