

SECTION III

Defining Good Assessment

We must arrange school learning so that it encourages more varied achievement goals than the narrow set of values often associated with competitive excellence and high standardized test scores at all costs. We must also learn to respect alternative ways for attaining excellence — for the sake of the group, for tradition, and for honor. Moreover, these changes must be made without doing violence to the fundamental academic mission of all schooling, that of providing students with the subject matter skills necessary to thrive, not merely survive. Finally, in the process of reform, we must not ask students to give up their cultural and ethnic identities.

Martin V. Covington, *The Will to Learn: A Guide for Motivating Young People*, 1998



Section III focuses on the criteria for high quality assessment. Five keys to quality are described with detailed explanations, examples, and classroom strategies. The following vignette provides an example of a situation in which the teacher has assessment questions and could benefit from understanding the keys to good assessment.



**VIGNETTE:
MR. JONAH'S CLASSROOM ASSESSMENT CHALLENGE**

It's October, and Mr. Jonah, an eighth-grade science teacher, is planning a unit of study on electrical energy. Last year's visit to the local power plant provided a real-life context for the conceptual understandings the students were developing and also served as a starting point for investigating the impact of electricity on the daily lives of the community. His assessment plan last year included keeping electrical energy learning logs, recording observations at the power plant, and drafting, then refining, questions to use while interviewing older members of the community. At the end of the unit, students presented their learning at an Energy Fair through oral presentations and written reports with models and charts.

This year, Mr. Jonah plans to involve his students in developing criteria for evaluating their learning logs and interview questions. He's also pondering ways to gather good information about several new students who have recently enrolled. He's had a difficult time interpreting their written work and is worried that they may be

unable to complete the reports that were a central feature of last year's unit assessments. He's fairly sure that two of the students have very limited experience writing in English. Another has lived in this country all her life, but her writing, while lengthy, seems unorganized with lots of starts and stops and ideas that seem to disappear. A fourth student's writing is often unrelated to the topic. Mr. Jonah sees some difficult assessment challenges ahead.



THINGS TO CONSIDER:

- How can Mr. Jonah make sure that these students understand what's expected?
- How will he interpret and use the work they produce? What decisions can he make about adjusting instruction and judging their progress?
- Which forms of assessment will give them the best opportunity to show their learning while providing good evidence about the expectations he's set?
- How many samples of their work will he need to make confident judgments about their learning?
- What are potential barriers to good assessment: features of the assessment that contain biases or lead to mismeasurement?

Like Mr. Jonah, many educators struggle on a daily basis to determine what their students truly know and are able to do. They want to use their assessments to provide insights into the skill and knowledge of their students and to guide their next steps in classroom instruction and assessment. Before they can move their students toward improved knowledge and performance, educators must be able to assess effectively the work of their students.

Building on the groundwork laid in Sections I and II, this section provides a foundation for ensuring that any assessment we use to gain insights into the learning of our students is of high quality and adheres to standards for assessment defined in the work of Dr. Richard J. Stiggins, author of *Student-Involved Classroom Assessment, 3rd Edition*

(Merrill, 2001), and his colleague, Dr. Judith Arter, both of the Assessment Training Institute, Portland, OR. Dr. Arter was principal author of *Improving Classroom Assessment: A Toolkit for Professional Developers: Toolkit98*, a previous publication of the nation's regional educational laboratories (See Section VIII). In addition to information about each standard, which we are calling Keys to Quality Assessment, we add sociocultural implications. For each key we provide an example of cultural and linguistic issues to consider, guiding equity questions, and things to try for sharpening mastery of these keys to high quality.

FIVE KEYS TO QUALITY ASSESSMENT*

KEY 1

What?

Clear and
Appropriate
Learning Targets

KEY 2

Why?

Clearly Focused
and Appropriate
Purpose

KEY 3

How?

Appropriate Match among
Targets, Purposes, and
Method of Assessment

KEY 4

How Much?

Sufficient Sampling of
Student Work to Make
Sound Inferences
about Learning

KEY 5

How Accurate?

Fairness and Freedom
from Biases that Distort
the Picture of Learning

* Source: Stiggins, R.J. (2001). *Student-Involved Classroom Assessment, 3rd Edition*. (Columbus, OH: Merrill, an imprint of Prentice Hall.) Adapted by permission.

KEY 1

What?

Clear &
Appropriate
Learning Targets

Clear and Appropriate Learning Targets. *WHAT* do we want to assess? What knowledge, skills, reasoning ability, products, and habits of mind are essential for student success? Do students understand what is expected of them? Is the knowledge or skill that students are expected to demonstrate in the assessment influenced by cultural and linguistic factors?

The first key to good assessment is having clear and appropriate learning targets. A learning target is a chunk of content clear enough and specific enough to both define what to teach and to let students know what they need to learn. Learning targets can be generated by the teacher or by others outside the classroom. They are called many things: content standards, benchmarks, outcomes, learning goals, objectives, essential learning requirements, competencies, learning expectations, and more. Effective learning targets are characterized by a strong relationship between targets and state or district standards, a match between targets and a specified amount of curriculum, and a clarity that allows students and others to understand what the targets mean. If we're not clear about our learning targets, our students will not be clear in their responses. Our job is to define in crystal clear terms **what** is to be learned.

As we develop targets, it's important to consider:

- Is the target clear enough that a group of teachers would agree on the range of knowledge, skills, and performance implied by the target?
- Would they agree on what to teach and what to assess?
- Have we clearly defined our expectations not only for ourselves, **but for our students?** (Regional Educational Laboratories, 1998)

For example, it's easy to agree on a target like "communicates well," but what does this mean? What types of communication? In what contexts? Effective student learning targets are specific enough to enable everyone — teacher, student, and parent — to share the same understanding of what knowledge and skills demonstrate mastery.

But this requires just the right amount of specificity. For instance, as an appropriate classroom learning target, "geography of our state" is too broad — it's really a topic that's likely to have multiple targets within it — and could be a whole year course! By contrast, "the length of the Columbia River" is too narrow a target — not of high enough importance to warrant much teaching time (although this depends on your overall goals for student learning and what core knowledge is an essential foundation for additional learning).

Types of Learning Targets*

There are many different ways to categorize learning targets for students. Categorizing helps us thoroughly think through what we want students to know and be able to do (in other words, clarify targets) and helps us determine if we have a good mix of learning targets.

- **Knowledge and Understanding:** What facts do students know outright? What information can they retrieve? What do they understand? The assessment challenge is to develop knowledge and understanding targets that are at the heart of a discipline — those worth learning and assessing.
- **Reasoning Proficiency:** Can students analyze, categorize, and sort into component parts? Can they generalize and synthesize what they've learned? Can they evaluate and justify the worth of a process or decision? The assessment challenge for reasoning targets is to define the difference between doing these things well and doing them poorly.
- **Skills:** We have certain skills that we want students to master such as reading fluently, working productively in a group, making an oral presentation, speaking a foreign language, or designing an experiment. The assessment challenge for skills targets is to define, in clear words, what it means “to do something well.”
- **Ability to Create Products:** Another kind of learning target is student-created products — tangible evidence that the student has mastered knowledge, reasoning, and specific production skills. Examples include a research paper, a wooden table, or artwork. The assessment challenge for product targets is to describe and define the characteristics or dimensions that make for a quality product.
- **Dispositions:** We also frequently care about student attitudes and habits of mind, including attitudes toward school, persistence, responsibility, flexibility, and desire to learn. The instructional challenge is to generate a classroom environment where students actively choose to engage in the learning and perform at their best. The assessment challenge is to define these and know how to use results to motivate students to want to learn.

* Source: Stiggins, R.J. (2001). *Student-Involved Classroom Assessment, 3rd Edition*. (Columbus, OH: Merrill, an imprint of Prentice Hall.) Adapted by permission.



VIGNETTE: THE STORY OF A MISUNDERSTOOD TARGET

In the example below, the teacher has a specific learning target in mind, but her students do not seem to be aware of what the teacher is looking for as she assesses them.

Ms. Schwartz is a high school English teacher. Her students are studying a novel. Periodically, as her students work through the novel, she asks questions in class. She uses this questioning as an informal assessment to gauge their understanding. During the questioning, the students give her very general answers; they do not provide details from the novel that reflect significant knowledge about the book.

As a means of checking their knowledge through another type of assessment, this week she gives her students an essay question. When she examines the results, she finds that students were able to make solid generalizations, but they still do not back up these generalizations with supporting details. Her conclusion, based on her two assessment processes, is that these students lack understanding of the novel either because they aren't reading the assignments or they read without comprehension. She is not sure what strategy she needs to use to get her students more deeply into the reading — or get them to read it at all! She decides to bring it up at the next grade-level meeting.

In her work session with other grade-level teachers, Ms. Schwartz discusses her concern about this lack of effort and skill to the other teachers. Several of the teachers begin to discuss the cultural background of many of the students in this class. Ms. Schwartz had not been aware that the students she teaches are from a culture that prizes communicating meaning with as few words as possible. In fact, in this culture, providing extensive detail is viewed as disrespectful of others, because it implies that the writer assumes that the audience can't "get it" without extra help. To be a powerful communicator in these cultures is to be able to communicate the gist of the idea or concept with as few words as possible.



THINGS TO CONSIDER

In the case above, the learning target was not explained to students. The teacher’s unspoken expectation that clear understanding is characterized by a display of extensive detail is at odds with the definition of what it means to “communicate well” in their culture.

The underlying challenge for this teacher was not just to address the different operational definitions of demonstrating understanding, but to help students bridge between the communication patterns in their home/culture and the expectations in the classroom — to arm students with clear portraits of the learning target, and in so doing, enable them to select the appropriate communication pattern for the assessment.

Many times a target that we consider straightforward and clear is just the opposite. Unless we take the time to get to know our students and their personal context, we may incorrectly assess the level of our students’ understanding.

Table 5: Reflective Questions for Key 1

Equity Questions	Some Implications for Action
Are my learning targets clear to all of my students?	Discussing your targets with other teachers may help you know if you are using strategies that are ineffective because of culture. Also, talk to your students about their assessments. If their response is not what you expected, ask them why.
Is my learning target important? To all of my students? Is it clear to my students and me how I will know when they achieve the target?	Ask yourself if you have helped students see how the target fits into the big picture. Connect the target with external expectations and life outside the classroom. What words and processes have you used to communicate your expectations?
Is there prior knowledge that can help my students connect to, and make meaning of, the learning targets?	Help students connect a given target with their experience. Give and ask for examples to clarify meaning and check for understanding. Use conceptual organizers (webs, concept maps) to convey the meaning of the target. Ask students to brainstorm what they know about the target.



THINGS TO TRY

Think of a learning target to improve
(i.e., “communicates well”).

When you see this term or phrase, what does it tell you needs to be taught? Is it absolutely clear what you would need to teach?

Is it clear what would be assessed? What would students do to show their learning? When they are masters of the target, what would students know, be able to do, and create/demonstrate? What would you expect to see in students’ behavior?

What could you add to the target statement to strengthen and clarify it for other teachers and your students?

Revised/Refined Learning Target:

- **An option for checking common meaning:** Invite a colleague to review your draft target for clarity and appropriateness. Do you share the same definition of the revised target and its importance to student success?
 - **An option for student involvement:** When students complete the assignment for the target, have students rate the answers and look at differences and similarities in answers. Ask if some of the ideas are ones they might not have thought of, or if there are some that everyone agrees upon. This can be both an opportunity for insights into differences across cultures and a bridge to shared meaning within the context of the classroom.
-

KEY 2

Why?

Clearly Focused
& Appropriate
Purpose

Clearly Focused and Appropriate Purpose.

WHY are we assessing? How will the assessment information be used? By whom? To make what decisions? Are there cultural and linguistic expectations that might conflict with the intended purpose of the assessment?

The second key to good assessment is having a worthwhile purpose. Some assessments give students information about their own performance. Students use the results to decide what they'll study, how much to study, whether it's worth studying, what they're good at (or not), and how they might be able to earn their living as adults. Other assessments help teachers change and improve instruction. Still others give administrators, parents, and community members information about individual students or schools that they use in different ways. For example, parents may want to know how their own children are doing while state department officials may use statewide assessments to decide how to distribute resources.

In addition to providing useful information, good assessment does not harm or unduly frustrate our students. Assessment should not be so expensive, time-consuming, or painful that the negatives of the process outweigh the benefits. Ultimately, we want students to understand how they learn, how to evaluate their own performance, and how to undertake the learning necessary to improve their work.

Looking at the variety of users and uses also underscores the crucial importance of balancing large-scale assessments and high quality classroom assessments. After all, it is the day-to-day classroom assessments that most affect decisions made by teachers, parents, and students. What happens if classroom assessments are not well thought-out and executed?



VIGNETTE: THE WRITING CONFERENCE

The following example shows what can happen when a teacher is not clear about the purpose of an assessment.

Mrs. Albert is a seventh-grade language arts teacher. This week she is beginning a unit on essay writing. She has decided that she will help her students to develop their essays through writing conferences. The students will write three drafts on the same topic, and she will conference with them personally about each draft. On the first rough draft, she asks the students to focus on the main ideas for their paper. However, she cautions them to add enough detail so that she knows they can further develop the main topics.

One of her students, Nguyen, is looking forward to this conference. She used brainstorming to come up with her topics and then used clustering to group topics and put them in order. She is excited about the subject matter, and she knows that while she is still behind her class in the mechanics of English, she is very good at coming up with ideas.

When Mrs. Albert hands her paper back, the first thing she notices is all of the marks on the paper for spelling and grammar. In fact, there are so many errors marked on the paper that she is unable to focus on what Mrs. Albert is saying to her. When Mrs. Albert asks her if there are any questions, Nguyen quietly picks up her paper and returns to her seat.



THINGS TO CONSIDER

Using writing conferences to improve student writing is a wonderful tool in the classroom. It is a strategy that allows the teacher and the student to interact one-on-one about the student's work. However, because Mrs. Albert did not communicate to his students all that she would be assessing, this process failed Nguyen. Her pride in her work was gone within seconds. The dismay that Nguyen felt is common when a teacher's purpose changes or is miscommunicated. This causes students to feel confused and frustrated. It is vital that students understand the purpose for their work and that teachers carefully express and follow through on the information they give their students.

Before an assessment can be used to reflect the current capabilities of the students, it must be written and designed in such a way that students can respond effectively with the knowledge that they possess so that the results of the assessment have meaning and can be useful for decisionmaking. If the purpose of the above assignment was to determine how well students are able to express their ideas, then the conference should have focused on ideas. Nguyen was unprepared to discuss spelling and grammar conventions and may have, in fact, believed that she would be able to work on those parts of the paper later. She may have focused exclusively on ideas; therefore, an assessment of her spelling and grammar abilities is unlikely to show what she is truly able to do when she is prepared and tries her best in those areas.

Table 6: Reflective Questions for Key 2

Equity Questions	Some Implications for Action
Why do I use this assessment with my students?	Consider turning your assessment purposes into questions to clarify the kinds of decisions you wish to make on the basis of an assessment. Example: Has my focus on word choice resulted in more vivid words in my ELL students' oral and written stories? Is there a consistent pattern of errors in my students' mathematical problem solving? Which students can benefit from math lab?
What decision(s) will be made on the basis of this assessment? Are my students clear about how the assessment will be used?	Students need to know when assessment is intended primarily to let you and them adjust and improve and when there will be other users of the information. They need to know what kinds of decisions will be made, particularly when high stakes may be attached.
How can I make sure that users outside the classroom don't misuse assessment information about my students? For example, do the users of the information understand the validity issues posed by assessing the content learning of English language learners?	There are many different audiences for assessment information. It is important to consider what information to present and how to present it. Some assessments should be used exclusively within the classroom to promote learning and some should be shared. Users must realize that, for students whose language and/or background experience do not match those of the dominant group, inferences about those students may not be valid.
When do I use assessment activities solely to build student self-assessment skills? How do these occasions differ, if at all, from those in which I assess students to find out what they know and can do?	Helping students understand how they are assessed and allowing them to be a part of the process helps them understand expectations and learn to recognize how to improve their own performances. These are important skills, and research shows this process improves student motivation and achievement.



THINGS TO TRY

The list below was developed by a group of teachers seeking to define the major uses of assessment in their classrooms.

We want our assessments to

- show growth in students' work
- allow all students, including those of various cultural and linguistic backgrounds, to demonstrate what they are able to do
- promote creative and critical thinking and problem solving
- give insights into progress toward our targets
- enable students to manage their own learning
- help determine student grades
- communicate with parents
- provide information to improve instruction
- let us see where our students' work is in relation to standards of excellence
- help us evaluate the effectiveness of program, materials, and approaches
- enable students to show what they know and can do in a variety of ways

(Compiled from the work of teachers at Waialae Elementary School, Honolulu, and Garapan Elementary School, Saipan)

Select 2-3 items from the list that you believe match YOUR primary purposes for assessment, or feel free to add to the list. Compare and discuss your priorities with those of your colleagues, parents, students, or others.

My primary purposes for assessment

KEY 3

How?

Appropriate Match
Among Targets,
Purposes, & Method
of Assessment

Appropriate Match Among Targets, Purposes, and Method of Assessment. *HOW* will we assess the targeted learning? Which methods of assessment are most appropriate for each kind of target? Is the method used to address each target designed with consideration of the cultural and linguistic traits of the students?

The third key to high quality assessment is to match targets, purposes, and methods. Good assessment means choosing the best assessment method for the learning target and the population being assessed. Stiggins has provided us with a description of four methods of assessment (*Student-Centered Classroom Assessment, 2nd Edition, 1997*). Each method has strengths and provides insights into one or more kinds of targets. Each also has limitations for use with certain targets and varies in applicability to different purposes.

Assessment Methods

- **Selected Response:** The selected response method provides students with a set of possible answers from which to choose. Common selected response formats include multiple choice, true/false, and matching. In addition, we have included short answer formats here because, like selected response items, they require specific correct answers that can be scored yes/no or right/wrong. The strength of this method is its ability to establish whether students have the knowledge on which further learning can be built. It tends to match best with knowledge and reasoning targets. A limitation is that this method's frequent focus on vocabulary and sentence structure can be inordinately challenging for English language learners.
- **Essay:** This method generally involves using writing as the tool for demonstrating content knowledge, conceptual understanding, and reasoning. Typical formats include traditional content area essays and writing essays. The strength of this assessment method is that it requires students to pull together the bits and pieces that they have learned into a coherent written whole. Again, however, writing may not always be the best way for English language learners to demonstrate their content knowledge.
- **Performance Assessment:** Performance assessment is assessment based on observation and informed judgment. It focuses on what students can do with what they have learned. It generally falls into two categories: products and performance skills. Products include such things as posters, graphs, drawings, videos, models, projects, rebuilt engines, maps, diagrams, computer visuals, etc. When

writing itself is the learning target, performance assessment is used to create products that demonstrate mastery of forms of writing such as narration, exposition, persuasion, and so forth.

Performance skills that might be the focus of a performance assessment include oral presentations, plays, debates, songs, or lab skills. A strength of performance assessment is that it enables us to assess powerful learning that is often not well assessed using other methods. A limitation is that, without care, students might be asked to perform in contexts that are culturally unfamiliar.

- **Personal Communication:** This method offers us opportunities to view student learning by talking with, and listening to, students. This method helps to uncover student misconceptions, as well as confirm their reasoning, attitudes, and behaviors. Strategies for this method might include asking questions, interviewing, conferencing, discussing, and listening for *their* questions. The strength of this method is that it often allows students to express knowledge without having to worry about the mechanics of writing. Personal communication is often an appropriate way for teachers to follow up and probe beyond the learning revealed in products and performances. A limitation in this method is, again, cultural — do we interpret nonverbal and interactional information correctly?



**VIGNETTE: MR. WASHINGTON'S STORY:
"WHY DON'T THE GIRLS SPEAK?"**

In his ninth-grade geometry classroom, Mr. Washington has just started a new unit on circumference and diameter. He decides to assess students' prior knowledge of the topic by asking the class several questions and writing their answers on an overhead entitled K-W-L. K is for what students already know, W is for what they want to know, and L is for what they will learn by the end of the unit. While he asks questions and calls on the students who raise their hands, he makes notes about who answers in his grade book.

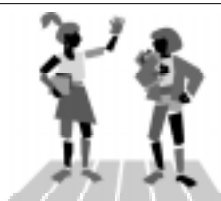
At the end of class, he looks at his notes and realizes that almost all of the answers came from the boys in his class. He is dismayed that the girls do not seem to know anything about the topic.



THINGS TO CONSIDER

Simply asking questions of the whole class and calling on students who raise their hands may not always be the best way to assess what students know. Researchers have frequently found that boys tend to put their hands up faster than girls do when they are asked a question. Many girls do not raise their hands until they have organized their thoughts, often not until after a boy has already answered. Or something else entirely may have been going on in Mr. Washington's class to create bias that he was not aware of. Teachers need to be aware of these potential gender differences, as well as other sources of bias, in order to create mechanisms to prevent them from happening when they create assessments.

Perhaps allowing everyone to share with a neighbor while he walked around the class and made observations or allowing everyone to write their prior knowledge on note cards would have given Mr. Washington a better sense of the class's prior knowledge and interest level. Finding the appropriate match between target learning, purpose, and method can be challenging but is an important key to effective student assessment.



DID YOU KNOW?

Studies have shown that teachers typically wait only nine-tenths of a second for students to answer questions. This rapid pace favors males, as males are more likely to jump into classroom conversations first. If teachers increase wait time to three to five seconds, it will give girls and other students more of a chance to participate. The wait time also allows for more right answers and more creative answers. Research also shows that the time between a student's answer and the teacher's reaction to that answer is only a split second. Increasing this second wait time will give the teacher a better opportunity to respond more meaningfully to girls' contributions. (For more information about gender issues in the classroom, see the Web site: www.girlscount.org)

Table 7: Reflective Questions for Key 3

Equity Questions	Some Implications for Action
As I think about a specific assessment I've given recently, was there a good match between my purpose, the kind of learning target, and the assessment method? Would I change anything?	Thoughtful reflection on your own practice is one of the most effective ways to make changes that will help your students and give you better information. As you continually question whether you chose the best assessment methods for your targets and purposes, your assessments will improve.
Do I use a good mix of methods? Are there any methods that I use very infrequently? Why?	Some students will respond to some methods better than others will. In particular, linguistic or cultural issues may make some methods more effective than others. It is a good idea to use a wide variety of methods throughout your teaching.
Which methods might not be effective for English language learners? Why?	Assessments that rely heavily on reading and writing English will be challenging for English language learners. Using them is fine if you are assessing English language development. If you are assessing for other content areas, however, students may know much more than they are able to express. Helpful guidelines may be to provide many opportunities for clarifying instructions and to create prompts that are easy to read and stated in simple syntax.



THINGS TO TRY

Think of a learning target. (For example, “Knows the Causes of the Civil War”)

Learning Target:

How would you normally assess this target (selected response, essay, performance assessment, or personal communication)? What kind of information about the students’ knowledge of this target would this assessment give?

How?

What information:

Choose an assessment method that you do not normally use. What kind of information about the students’ knowledge of this target would this assessment give?

How?

What information:

An option for checking common understanding:

Discuss your ideas with colleagues. They may have different ideas about how to best match up targets, purposes, and methods. Discuss your results with students to see if they agree with your choice of methods and the types of information you can gather from the assessment.

KEY 4

How Much?

Sufficient
Sampling of
Student Work to
Make Sound
Inferences
About Learning

Sufficient Sampling of Student Work to Make Sound Inferences About Learning. *HOW MUCH* will we collect? Do we have enough varied samples of student work to make good judgments about current proficiency related to the target learning? Have we chosen sufficient and varied assessment examples that allow the students to take advantage of their cultural and linguistic strengths?

Assessment is a broad term that should refer to more than a single test or a single sample of work. To be able to assess accurately what a student knows and is able to do for any given learning target, teachers should analyze enough samples to make confident inferences about student learning. In fact, larger, more complex learning targets may require more samples. For instance, a single assessment such as a single sample of writing or a single math problem would not enable a teacher to infer writing or problem-solving competence. Much evidence would need to be gathered to draw confident conclusions.

In addition to choosing a sufficient number of samples, a teacher should also consider the ways culture and language influence student performance. For example, although we frequently think of selected response items as the way to test for knowledge, perhaps students can exhibit the same knowledge through visuals, webs, maps, drawings, or graphs. Perhaps some work can be produced collaboratively. Depending on the target, samples from multiple occasions and use of different methods help contribute to a clear picture of student performance. A good assessment plan offers a representative sampling from a variety of tasks that is large enough to permit the teacher to make accurate inferences about student learning.



VIGNETTE: MINERS AND THE GOLD RUSH

In the vignette below, a teacher uses only one assessment to assess student learning and finds that he may not have the information he needs.

Mr. Salena, a fourth-grade teacher, has developed a unit called the “Story of the Gold Rush.” He and his students have spent weeks studying the routes to the California Gold Country, the supplies needed by the early settlers,

and the impact of the Gold Rush on the population of California. They have read books and examined artifacts from the time period. Finally, Mr. Salena has asked his students to write a story from the viewpoint of a miner.

When Mr. Salena reads the answers of the students, he is very disappointed in their stories. They reveal very little about the journey to California; on the other hand, they are unusually creative, and many of the stories are actually parables that present lessons in life to the reader. Mr. Salena thought he had been very clear in explaining to the students that he wanted a story about the Gold Rush. He is disappointed that he is going to have to reteach the material and develop a new assessment.



THINGS TO CONSIDER

When Mr. Salena made this assignment, he was trying to garner the interest of his students as well as assess their knowledge about the California Gold Rush. But he failed to consider that his students might see this assignment from the viewpoint of their culture rather than as a social studies project. In the culture of his students, storytelling is used as a tool of communicating expectations, religion, and etiquette to children; stories are used to communicate ideas and concepts, not facts. The children did tell stories; however, they created stories that fit their experiences with storytelling, not the stories that Mr. Salena expected.

Because this is a single assessment and because the students have responded in a manner he didn't expect, Mr. Salena does not have a clear idea of the depth of understanding that his students have about the Gold Rush or its impact on California. In reality, his students may know much more about the Gold Rush than they have displayed in their assessment, but it is impossible to know without additional samples. If this project was one of several assessments, he would have a clearer vision of his students' knowledge.

These stories do serve a purpose that Mr. Salena did not intend. They have provided him with an opportunity to

get to know more about the way his students think, and therefore how they learn. Though this one event will not provide him with all the information he needs, it will help him to know about his students' sociocultural perspectives and strengths and how these should figure into future assessment plans. These stories might also provide a bridge to create an environment of sharing and interest that will help him in future assessments.

Table 8: Reflective Questions for Key 4

Equity Questions	Some Implications for Action
Do I remember a time when I felt uncomfortable because a decision was made based on too little information?	This is an indication that multiple assessment opportunities are important. A student could be sick, tired, or distracted on any given day, so that even if the assessment was created perfectly, they may not perform to the best of their abilities. We certainly want to give them another chance.
Are my students given multiple opportunities to show what they have learned, and through multiple kinds of assessments?	For students from nondominant language and cultural backgrounds, getting an adequate sample of performance is a challenge because it is likely that without significant changes, many assessments will yield less information about these students than they do for "mainstream" students.



THINGS TO TRY

Below is the start of a plan for gathering samples of student work related to a reading standard. After examining the standard, its key components, and the beginning of the plan, add your ideas about additional sources of evidence that will provide a good sampling of information about student progress toward the standard.

Reading Standard	Making Sense of Standards	Key Features of the Standard	Performance Indicators
Read a range of literary and informative texts for a variety of purposes	<p>“Range” is the ability to read a wide variety of texts for a variety of purposes. Reading a range of texts helps students deepen their knowledge of reading and of their world.</p>	<p>Types of text:</p> <ul style="list-style-type: none"> • Literary • Informative • Functional <p>Purposes:</p> <ul style="list-style-type: none"> • Literary experience • Gain information • Perform a task 	<p>There is evidence of breadth. The student has read:</p> <ul style="list-style-type: none"> • At least 25 books • A balance of literary and non-literary works • At least 3 different genres or modes • At least 5 different authors

(Adapted from National Council of Teachers of English, 1996)

Sampling Plan

Preliminary list of assessment sources for collecting samples:

- ✓ Reading interview
- ✓ Booklist/log
- ✓
- ✓
- ✓

Some Cultural and Linguistic Considerations

How will students contribute to the evidence of their reading range?

What are some texts that reflect the language and culture of your students?

What sources in the family and community can contribute to the picture of students' reading range?

Evaluate Your Plan: Does this plan ensure that you'll gather enough information of the right kind, so you can draw confident conclusions about student achievement? What else would you want to know, and what evidence would you need?

KEY 5

How Accurate?

Fairness and Freedom from Biases that Distort the Picture of Learning

Fairness and Freedom From Biases That Distort the Picture of Learning. *HOW ACCURATE* are the

assessments? Do they really assess what we think they're assessing? Is there anything about the way a target is assessed that masks the true learning of a student or group of students? Do we know the strengths that students bring to learning and use those strengths in our assessments?

Even when we design our assessments so that we know exactly what we are assessing and why we are assessing, we've picked just the right way to assess these things, and we know how much to assess; unfortunately, it's still possible to end up with assessments that don't work — for your purposes and/or for your students. Have you ever tried to engage students in an instructional activity and found that they did not learn what you expected them to learn? The instructions weren't clear, there wasn't enough time, students didn't have all the prerequisite skills, or students did not react to the learning sequence the way you envisioned they would.

The same thing can happen in assessment. These “things that go wrong” are called sources of mismeasurement, bias and distortion, or invalidity. The result is that the information from the assessment doesn't mean what we think it means. Such biases can have serious consequences because a grade or certification of competence is only as good as the assessments upon which it is based. What happens if the ability to read the instructions interferes with a student's ability to demonstrate math skills? Or, the necessity to write a response interferes with how well a student can demonstrate the skills needed to set up a scientific experiment? These are serious potential sources of bias and distortion.

When the first four keys are addressed equitably, bias and distortion can be limited. However, as can be seen from the questions and vignettes above, any assessment can contain sources of bias and distortion, not just the so called “objective” instruments. We tend to think of standardized tests as containing the greatest bias and distortion in the classroom because these assessments do not take into account students' experiences, nor do they often correspond closely with local curricula. However, as can be seen from the questions above, any assessment can create bias and distortion, not just objective instruments. The following table indicates some of the most common sources of bias and distortion.

Table 9: Possible Sources of Bias and Distortion in Student Assessments

Sources related to the assessment itself	<ul style="list-style-type: none">• Too much reading or writing on an assessment designed to assess something besides reading or writing• Unnecessarily difficult or unfamiliar vocabulary used in instructions• Assessment in a language that is unfamiliar to the student• An assessment method that doesn't allow students with different learning styles to do their best• Unclear instructions• Attempts to make a problem more "real-life" that result in a context more familiar to some groups of students than others• Rater bias or untrained raters for performance assessment• Performance criteria that don't cover the most important aspects of performance or that are vague• Irrelevant clues to the right answer, more than one right answer, or unnecessarily convoluted questions on multiple choice tests• Narrowly defined criteria that recognize only one way to be excellent
Sources related to the student	<ul style="list-style-type: none">• Student is tired, hungry, sick, or distracted for some other reason• Student is not used to the format, timing, or other logistics of the testing situation
Sources related to the environment	<ul style="list-style-type: none">• Noisy or distracting environment• Assessment administrator who projects a negative attitude toward the assessment



VIGNETTE:
MS. PONCIANA'S STORY: "SAVING THE TURTLES"

Ms. Ponciana and her students are in the midst of a unit that is structured around this core question: How can we preserve and care for turtles found on our reef and encourage others to do the same? One of the targets for this unit is to strengthen students' knowledge of turtle habitat and how it's changing as a result of human use of the reef environment. Her hope is to also move her students to care deeply about the preservation of the turtle and, because of that caring, to be disposed to use their knowledge to guide the actions they take themselves.

It is early in the unit, and Ms. Ponciana wants to gauge the current attitude of her students toward the unit topic. She has been using personal communication — questions that she's given small work groups to discuss and report — as the method of gathering initial information about student knowledge of, and disposition toward, endangered turtles.

But, as she focuses in on one group, she realizes that there's a student who has not spoken at all. Rowena and her cousin, Charlie, are newcomers from the Federated States of Micronesia. He's taking part. Rowena appears to be disengaged from the discussion. In the midst of the discussion, she is silent. Ms. Ponciana is frustrated. Surely Rowena, who comes from an island where turtles are greatly valued, has ideas to add to the group's work. Why won't Rowena contribute?



THINGS TO CONSIDER

Judging silence as disinterest or lack of knowledge can result in biased assessment. Different cultures have different expectations of behavior, communication, and respectful action that encourage different behaviors. For example, in the scenario above, Charlie and Rowena are from a culture where females do not speak when they are in a group with a male relative. Grouped with other girls or in a mixed group without a male relative present, Rowena would have the opportunity to express her ideas during the discussion. If Ms. Ponciana had changed the composition of the work groups in the classroom, Rowena would have been able to express herself.



DID YOU KNOW?

Grouping by gender can play a role in assessment accuracy in settings where the expectations and experiences of girls result in limited participation. For example, there is evidence from recent efforts to encourage more girls to go into mathematics and science that, when grouped together for hands-on science activities, the males in the group tend to dominate and girls' contributions are little recognized.

Table 10: Reflective Questions for Key 5

Equity Questions	Some Implications for Action
If the target is an academic skill or concept, can it be separated from the student’s language as a target of assessment? Do I use words in defining the learning targets that are outside the experience of some of my students or that might have different meanings for them?	Consider that some targets may be language-dependent. Second-language learners may function differently vis-à-vis the target from native speakers. Check for multiple meanings of key words in your targets — ask students, colleagues, community and family members for help. Stick to the simplest language that can convey clearly the meaning of your targets.
How does culture impact assessment methods?	Like language, culture can have a profound effect on assessment results. Teachers need to be aware of the norms for the different cultures in their classes, so that they do not create assessments that will confuse or frustrate students. Some students may respond better through group work, demonstrations, or exhibitions versus paper and pencil tasks, and some choice over when they will show what they know.
Do I know enough about my students’ cultures and the strengths they bring to learning?	Invite families, cultural organizations, and students themselves to describe and give examples of the strengths in their culture.
Am I unconsciously biased against particular students because of the ways they use language, their dialect, or their pronunciation?	Try to look beyond the words of the students and focus on what they are communicating. Collect multiple papers from ELLs and compare them to find trends and traits. Use this information to help in assisting the students and in determining when the student has more knowledge that his/her paper reveals.



THINGS TO TRY

Think of a time when you were not able to get the information you needed from an assessment.

Your assessment:

What were the possible biases or sources of mismeasurement on this assessment?

What do you think went wrong, and how do you think you could change your assessment in order to get the information you need?

Summary

Think back to Mr. Jonah’s classroom. He was concerned about new students whose writing was very limited or appeared rambling and unorganized. A simple diagnosis of “having problems” is the common reaction of most teachers; however, if we take the time to look at the students, we find that there are sometimes other issues that affect student success. We should always ask ourselves if there are cultural or linguistic expectations and patterns that result in students failing to do well on an assessment. Perhaps the topic of the writing is outside their experience. Perhaps it’s the need for expressing knowledge in written form when having little mastery of English writing skills is the barrier. Mr. Jonah’s written science report requirement may actually mask the learning of the two students who are English language learners. Requiring an essay written in English will not give accurate information about the science targets if the students do not have the necessary vocabulary and writing skills. Though it may be true that a child may be slow to develop writing skills, it is still important for us to look at the whole picture of the student.

For us, good assessment is about more than creating solid evidence of the status of student work. Although we must pay attention to the quality of our assessments, we must also attend to the impact assessments have on student motivation to learn. Good assessment, and the feedback that it can provide, leads to improved instruction and opportunities for students to learn.

Good assessment at any level is characterized by

KEY 1

What?

Clear &
Appropriate
Learning Targets

Key 1: Clear and Appropriate Learning Targets.

WHAT do we want to assess? What knowledge, skills, reasoning ability, products, and habits of mind are essential for student success? Do students understand what is expected of them? Is the knowledge or skill that students are expected to demonstrate for the assessment influenced by cultural and linguistic issues?

KEY 2

Why?

Clearly Focused
& Appropriate
Purpose

Key 2: Clearly Focused and Appropriate Purpose.

WHY are we assessing? How will the assessment information be used? By whom? To make what decisions? Will the cultural and linguistic traits of the user interfere with the intended purpose of the assessment?

KEY 3

How?

Appropriate Match Among Targets, Purposes, and Method of Assessment

Key 3: Appropriate Match Among Targets, Purposes, and Method of Assessment. *HOW* will we assess the targeted learning? Which methods of assessment are most appropriate for each kind of target? Is the method used to address each target designed with consideration of the cultural and linguistic traits of the students?

KEY 4

How Much?

Sufficient Sampling of Student Work to Make Sound Inferences About Learning

Key 4: Sufficient Sampling of Student Work to Make Sound Inferences About Learning. *HOW MUCH* will we collect? Do we have enough varied samples of student work to make good judgments about current proficiency related to the target learning? Have we chosen sufficient and varied assessment examples that allow the students to take advantage of their cultural and linguistic strengths?

KEY 5

How Accurate?

Fairness and Freedom from Biases that Distort the Picture of Learning

Key 5: Fairness and Freedom From Biases That Distort the Picture of Learning. *HOW ACCURATE* are the assessments? Do they really assess what we think they're assessing? Is there anything about the way a target is assessed that masks the true learning of a student or group of students? Do we know the strengths that students bring to learning and use those strengths in our assessments?