Classroom Assessment *For* Student Learning (CASL) Perspective on the JCSEE Student Evaluation Standards

Judy Arter

ETS Assessment Training Institute
317 SW Main, Suite 1200, Portland, OR 97004
503-228-3060
www.ets.org/ati
jarter@ets.org

Classroom Assessment *For* Student Learning CASL)
Perspective on the JCSEE Student Evaluation Standards

I have been asked to do two things in this paper: describe the *Classroom Assessment* for *Student Learning* (*CASL*) professional development program in classroom assessment and relate quality practice, as taught by *CASL*, to The *Student Evaluation Standards* (Joint Committee on Standards for Educational Evaluation, 2003). The purposes of this and other papers in the proceedings are to (1) explore how the student evaluation standards might be revised to reflect current thinking in the field and make them more useful to practitioners, and (2) more broadly, consider how to best ensure that sound classroom assessment practice, especially assessment *for* learning, be implemented by practitioners.

Classroom Assessment for Student Learning (*CASL*)

The *CASL* professional development program centers on the information needs of classroom-level decision makers—teachers, students, and parents. It emphasizes both accuracy—generating dependable information about student learning—and use—using assessment results to promote student motivation and learning. Accuracy is essential for all forms of assessment at all levels, but equally important is helping teachers learn to design the classroom assessment environment to avoid unintended negative motivational consequences surrounding assessment and grading, and in fact, to use the assessment experience as an instructional intervention to maximize student achievement.

Classroom Uses of Assessment Information

CASL makes the distinction between assessment *of* and *for* learning. These two purposes for assessment provide different information to serve the needs of different decision makers at different times. The purpose of assessment *of* learning is to summarize the achievement status of individuals and groups at a point in time. We use the term "assessment *of* learning" synonymously with "summative assessment." Classroom-level summative assessment usually involves a value judgment about the success of the learning; this frequently takes the form of a grade or a statement of level of mastery. *CASL* materials (Stiggins, et. al, 2006, Chapter 10; O'Connor, 2007) deal with grading and reflect current thinking in the field (e.g., Brookhart, 2004, and Guskey, 2001) about sound grading practice: the purpose of a grade is to communicate as accurately as possible the student's level of achievement at the time the grade is assigned without consideration of any other student characteristics.

Although effective summative assessment and grading practices are addressed in CASL, the major focus is assessment *for* learning—using assessment results and materials (e.g., rubrics, item formulas, and items) to maximize student learning before graded, assessment of learning events occur. We place this emphasis on assessment *for* learning because of the research showing that assessment *for* learning practices can dramatically improve student motivation and learning (e.g., Assessment Reform Group, 2002; Black & Wiliam, 1998; Bloom, 1984; Brookhart, 2001; Brookhart & DeVoge, 1999; Brookhart & Darkin, 2003; Butler, 1988; Butler and Newman, 1995; Crooks, 2001; Costa & Kallick, 2004; Dweck, 2001; Dweck, 2006; Dweck, 2007; Hattie & Timperley, 2007; Kluger & DeNisi, 1966; Morgan, 2008; Shepard, 2000; Shepard, 2008; White & Frederiksen, 2988; Zimmerman, 2008).

We do not use the term "assessment *for* learning synonymously" with the term "formative assessment." Traditionally, formative assessment has focused on the information needs of teachers—ensuring that teachers have the information they need, at the time and in the form they need it, to maximize productive decisions about students and instruction. While teacher decision-making is certainly important, we regard students as equally important decision makers about their own learning. Thus, in *CASL*, assessment *for* learning strategies address the information needs of both teachers *and* students.

We emphasize the importance of students as essential users of assessment results because of the relationship between assessment and student motivation. Assessment can have dramatic unintended negative motivational consequences for students. For example, an unintended negative consequence of summative assessment is that it is very susceptible to interpretation by students as a value judgment about their worth as individuals. Such interpretation is not so much of a problem if a student is measured and is found successful, but is distinctly counterproductive for the struggling student. The possibility of assessment results being interpreted in a counterproductive manner is reinforced in the current version of *The Student Evaluation Standards*. The definition of "evaluation" (p. 5) is "the systematic investigation and determination of the worth or merit of an object. In this book, the term *student* will be generically to refer to the object of the evaluation."

We want to use the classroom assessment process to keep students making productive decisions about their own learning. Students respond productively when, upon seeing assessment results, they say (Stiggins, 2009, p. 8):

- I understand these results
- I know what to do next to learn more

- I can handle this

- I choose to keep trying

The counterproductive response, on the other hand, leaves the student saying:

- I don't understand these results

- I don't know what to do next

- I'll probably never get it anyway

- I give up

We regard attention to the emotional dynamics of assessment as crucial to sound classroom assessment practice. If students are essential decision-makers, it behooves us to consider what information they need, in what form, and at what time, to make productive decisions about their own learning. We've found the following statement from Royce Sadler (1989, p. 121) extremely useful in this regard:

> The indispensable conditions for improvement are that the student comes to hold a concept of quality roughly similar to that held by the teacher, is able to monitor continuously the quality of what is being produced during the act of production itself, and has a repertoire of alternative moves or strategies from which to draw at any given point.

In other words, to paraphrase Sadler and Atkin, Black, & Coffey (2001) students are able to make decisions that enhance their own learning to the extent they can answer three questions:

1. Where am I trying to go?
2. Where am I now?
3. How can I close the gap?

In *CASL*, we help teachers enable students to productively answer these three questions by applying what we call *Seven Strategies of Assessment* for *Learning* (see Figure 1). These seven strategies are organized by which of Sadler's three questions the strategy helps students answer. Our various publications (Stiggins, et. al, 2006, Arter & Chappuis, 2006, Chappuis,

2009) provide concrete examples of how these strategies play out with all forms of assessment from selected response to performance.

In summary, *CASL* emphasizes assessment *for* learning strategies because they help teachers:

1. Provide effective targeted, penalty free practice for students.
2. Enable students to take responsibility for their own learning.
3. Improve student motivation to learn. Enabling students to be in control of the circumstances of their own success and giving students targeted, penalty-free practice improves student motivation.
4. Improve student achievement. The previous three result in improved student achievement.
5. Move from a grading culture to a learning culture. Engaging students in AFL activities automatically moves attention away from grading and toward what is already done well and the next steps in learning.

Accuracy

Accurate information about student learning is essential for both assessment *of* and *for* learning. *CASL* materials, therefore, help teachers increase the accuracy of all forms of classroom assessment including selected response, extended written response, performance assessment, and personal communication. We teach reliability and validity by helping teachers:

- Specify the purposes (users and uses) of each classroom assessment.

- Specify the learning targets (student learning outcomes) to be covered in each assessment and make sure that these targets are both the most important ones to cover (alignment with standards) and that they align with instruction.

- Design each assessment to both cover the desired learning targets and satisfy the needs of users. Design has four parts—choosing the best assessment method for each target being assessed, adequately sampling the breadth and depth of each target, writing high quality items, tasks, and scoring guides, and avoiding potential sources of error that can distort results.

One tricky part of classroom assessment is how to be accurate with all students, even those who are struggling, and still obtain productive responses from them. As described one concrete approach is to use assessment for learning strategies such as those describe in Figure 1. Such strategies take the focus off of grading and self-worth and place it on learning, gives students the information they need to be able to take control of their own learning, and thus improves motivation and achievement.

Five Keys to Quality Classroom Assessment

CASL standards for quality classroom assessment—both accuracy and use—are summarized in *Five Keys to Quality Classroom Assessment* (Figures 2 and 3). The first three keys deal with accuracy—clear purpose, clear targets, and sound design.

Keys 4 and 5 deal with use. Key 4 addresses effective communication to all those needing information, including students. Key 5, student involvement connects to all the other keys, but is called out because of its importance in creating large achievement gains. Student involvement connects to Key 1 (clear purpose for assessment) in that students are key users of assessment information and materials. It connects to Key 2 (clear learning targets) because students needs to understand what they are to learn from every instructional activity, to Key 3 (sound assessment design) because students can be partners in designing and interpreting practice assessments, and to Key 4 because it is powerful for students to reflect on and communicate about their learning over time.

ATI's rubrics to evaluate the quality of classroom assessments reflect these Keys (Appendix A) and also provide a framework for the learning outcomes for users of *CASL* materials (Figure 4).

Classroom and Interim Assessment

Many educational systems are implementing more frequent interim assessments in addition to large-scale once-a-year assessment. Although the results of such assessments can be used in formative ways, in *CASL* we make the distinction between classroom assessment and interim assessment. Interim assessments tend to be given in a standardized manner across classrooms at the same time every few weeks. Classroom decision makers, however, need information more frequently. Also, students aren't always in the same place at the same time; therefore, classroom decision makers might need different information about different students at a given point in time. Finally, interim assessments tend to involve only teachers, not students, in making decisions about learning.

CASL Learning Teams for Professional Development

All *CASL* materials are designed to be used in a particular professional development context—that of collaborative learning teams. Collaborative learning by teachers has been shown to increase the chances that innovations are actually implemented in the classroom. In a recent summary of research, Wei, et. al (2009) say:

> "Research ... suggests that professional development is most effective when teachers engage actively in instructional inquiry in the context of collaborative professional communities, focused on instructional improvement and student achievement."

The CASL program involves teachers studying new ideas about assessment, instruction, motivation, and learning, trying out these ideas in the classroom, reflecting on success, and collaborating with each other to improve the success of such interventions. We have developed a handbook to help users set up and facilitate learning teams (Chapppuis, 2007).

Relationship Between *CASL* and *The Student Evaluation Standards*

The following comments are based on observations about the Standards made by several of us at the Assessment Training Institute, including comments sent to the Joint Committee July, 2008 by Rick Stiggins.

In general, the *Standards* and the lessons targeted in *CASL* match up very well. Both address both accuracy and use, focus on the classroom level of assessment, approach the topic from the same philosophical stance, namely classroom assessment in the service of student learning, and have tried to build in features to improve practice. The major differences are in emphasis, organization, and examples. First I'll provide some general observations and then comment on each standard.

First, we strongly recommend either to take "evaluation" out of the title or modify the definition of "evaluation" on page 5. The notion that classroom assessment is the "investigation and determination of the worth or merit of" students runs completely counter to the goal of taking the intimidation out of assessment so that students can focus on the information it provides to improve learning.

Second, even though many standards state that students should be involved in their own assessment, assessment still seems to be presented as something adults do to and for students. In actuality, student involvement is not merely an add-on if there's time, but an integral part of formative assessment essential to create desired achievement gains. There need to be many more examples of how to productively involve students in their own assessment and goal setting.

Third, some of the standards reach outside the classroom; they require action on the part of districts. The standards could be more teacher-friendly by focusing just on the classroom level and moving needed district actions to another place. While *CASL* materials also deal with things

like district assessment policy, hiring, and the need for good curriculum documents, we have

separate materials to guide the district administration's role in ensuring productive student

assessment (Chappuis, et al, 2005). The *Standards* needs to decide what it wants to be—a

comprehensive vision of excellence for student assessment at all levels, or a document designed

for classroom users of student assessment information.

Finally, some of the standards seem to be redundant. Perhaps the document could be

made more user-friendly by condensing and combining. Possible ways to combine are noted

below.

Propriety Standards

The *Propriety Standards* align very closely with the *CASL* approach. They provide the

context for our Five Keys (Figure 3). Many of the specifics fit nicely in our Key 1 (clear and

appropriate users and uses of assessment information), although some fit into Key 2 (clear and

appropriate learning targets), Key 3 (design) and Key 4 (communication).

P1: Service to Students

P1 represents the overall philosophy that the major purpose for student assessment is to

help students learn. We, too, highlight the connection of assessment to teaching and learning. We

agree with all the guidelines, especially E—student involvement. To emphasize the importance

of student involvement, another "Common Error" could be added—Not realizing that the

decisions students make about their own learning are of paramount importance and therefore not

designing the process to meet the information needs of students.

Case 1 (page 31) emphasizes that the teacher "now insists that his students take responsibility for their own learning." He does this by aligning assessments with learning targets and having students "comment on and account for the marks and comments they receive each time they are evaluated." It would be useful to present a larger repertoire of strategies for student involvement resulting in students being able to take responsibility for their own learning. In *CASL*, implementing the *Seven Strategies of Assessment for Learning* (Figure 1) results in students *being able to* take responsibility for their own learning.

Besides the problems noted with Case 2 (page 31), another is that the teacher put grades on practice work (homework) and included it in her final grades.

P2: Appropriate Policies and Procedures

The emphasis here is on district level policy, although two of the cases involve classroom-level policy. As noted above, having everything together makes the document cumbersome.

If district-level policy considerations are retained, it would be useful to actually provide sample policies that reflect best thinking about how to use assessment in the service of student learning. It's important to not *just* have policies in place, it's important to have the *right* policies in place. For example, although the Standards state the need for a policy on student cheating and plagiarism, some policies are better than others: it's *not* a good idea to have a policy that requires recording a zero for the work and incorporating such zeros into a final grade .

P3: Access to Evaluation Information

We don't deal much with this. Perhaps this could be combined with P2—policy.

P4: Treatment of Students

We totally agree with the philosophy behind P4. For us, the way to treat students with respect is through assessment *for* learning because of the relationship between assessment and student motivation. It's not just a matter of ethics; it's also a matter of learning success.

P5: Rights of Students

Perhaps P4 and P5 can be combined.

P6: Balanced Evaluation

We agree that identifying both the strengths and weaknesses of student performance is essential for effective use of assessment results. While such information is certainly important for adult decision-makers, it is especially required for effective use by students.

If students are to be important consumers of balanced evaluation results, the Standards might also note that it is important that strengths and weaknesses be referenced to specific content standards and not referenced to peers, expected growth, or aptitude. Referencing to the latter doesn't provide the information needed by students and can be motivationally counterproductive. Use by students also requires consideration of timing and amount. Students need to get descriptive feedback in time to improve before a graded event. Also, students shouldn't receive more feedback than they can effectively act on at one time. Finally, students should be involved in generating their own descriptive feedback (see Figure 1, Strategy 6).

P7: Conflict of Interest

      CASL deals with conflicts of interest under Key 3—sound design, avoiding potential

sources of bias that might distort result. Might P7 be combined with the accuracy standards?

Utility Standards

U1: Constructive Orientation

      This overlaps nicely with the *CASL* approach. Might U1 be combined with P7 and U7?

U2: Defined Users and Uses

      U2 corresponds with *CASL*, Key 1—the purpose of each assessment must be clear.

Perhaps U2 could be combined with P1 and P6. Also , students need to be kept at the top of the

list of users of assessment information.

U3: Information Scope

      U3 corresponds to *CASL*, Key 3—sound design, sampling. The following might be added

to the list of common errors:

- Not being clear on the learning targets to be assessed.

      With respect to Case 1, page 79, the fourth paragraph describes how, at the beginning of

the evaluation task, the teacher provided students with the criteria for evaluating the quality of

their performances. It would be even better if:

- The learning targets for the assessment were determined, not idiosyncratically by a single teacher, but across teachers based on important content standards
- The criteria for a sound performance were determined across teachers and rubrics defining levels of performance were collaboratively written
- Samples of sound and weak performance were identified to illustrate the various important dimensions of performance as defined by the rubric

- Students were given a student-friendly version of the rubric at the beginning of instruction, and students self- and peer-assessed using the rubric on similar tasks over time

U4: Evaluator Qualifications

U4 corresponds to all of our Keys to Quality Classroom Assessment. To fulfill U4 requires expertise on all the others standards. Does it need to be repeated?

U5: Explicit Values

U5 seems to be redundant with at least P1 and U1. Might they be combined? Also, the two cases illustrating U5 are about large-scale assessment not classroom assessment.

U6: Effective Reporting

U6 corresponds to *CASL*, Key 4—effective reporting.

U7: Follow-Up

Might this be combined with U6, P1 and U1? Also, there might need to be refinements in Case 1, page 102. First, the teacher is "grading" homework. Much homework is practice work and shouldn't be returned with a grade. Perhaps the word "grading" is being used generically, here and elsewhere, to mean "assessing and analyzing." If so, this needs to be made clear.

Second, the teacher is doing all the analysis. It can be even more effective if students do the analysis on practice work (*CASL*, Figure 1, Strategy 4). Finally, it might be useful to say that the teacher did not include previous, out-of-date information in a final grade, but only the most recent and accurate information on student performance.

Feasibility Standards

F1: Practical Orientation

Practicality is an important consideration. Much of the information in F1 seems to overlap with accuracy and use standards. Might they be combined?

Practicality also comes from integrating assessment and instruction—assessment as instruction. Our attempt with assessment *for* learning strategies is to provide concrete assistance with how to integrate assessment and instruction to maximize practicality and student learning.

I think that the main problem with Case 1, page 109, is that the district's solution to improving student writing is to develop another assessment. Instead, why not implement assessment for learning strategies using the state assessment materials, samples, and results?

With respect to Case 2, page 110, there are at least two other problems with the teacher's assessment:

1. The targets to be assessed were not clear. The scenario states that the teacher wanted to see how well his students "understood science" so he developed a performance assessment. If "understanding science" means being able to design and carry out an experiment, then a performance assessment might be best. But if "understanding science" means having knowledge of science facts, then performance assessment is not the most efficient assessment method.

2. The "observation checklist" needs to align with the learning targets to be assessed.

Might this be combined with the other standards on clear users and uses, such as P6 and U7? Also, this one seems bigger than classroom assessment.

F3: Evaluation Support

While true, this one might better be directed at district administration

Accuracy Standards

These correspond very nicely with *CASL*, Key 3—sound assessment design. Many of the accuracy standards, however, seem to be redundant with other standards. Can they be combined?

A1: Validity Orientation

A1 is completely aligned with CASL, Key 3—sound design, choosing the best assessment method, writing good quality items, tasks, and rubrics, sampling adequately, and avoiding sources of bias. A1 also includes information about combining scores which *CASL* deals with in Key 4—effective communication.

Guideline K (page 130) uses the term "graded" in the generic sense of "assessed" or "evaluated." It might be wise to avoid this use of the term "grade." (The same use occurs on page 132.)

In Case, 1 (page 131) there needs to be a rationale for the learning objectives chosen by the teacher; learning targets should not be idiosyncratic. The procedures on page 132 could be enhanced by adding other assessment for learning strategies.

A2: Defined Expectations for Students

A2 corresponds to *CASL*, Key 2: clear and appropriate learning targets. The second bullet on page 138 needs to be changed to reflect the role of content standards in determining the learning targets for each instructional activity. Learning targets should not be idiosyncratic.

A3: Context Analysis

A3 seems to be redundant with A1 and earlier standards on the ethical treatment of students.

In Case 1 on page 145 the student completed a paper instead of a lab. Did these tasks adequately measure the same learning targets? An alternate assessment only works if it provides information on the same learning targets. It's not a matter of an alternative way to earn points, it's a matter of providing evidence of desired learning.

A4: Documented Procedures

A4 is included in *CASL*, Key 4—sound communication. Perhaps an assessment for learning example would be good here, for example, how about a sample letter to parents about not grading homework?

Case 2, page 152 is presented as a good example. Although the teacher had, in fact, documented procedures, the problem is not adequately communicating procedures to all stakeholders.

A5: Defensible Information

A5 seems to be completely redundant with A1 and A4.

A6: Reliable Information

We like this one because it provides practical rules of thumb for teachers. Some of the

information, however, like knowledge of the various types of reliability coefficients, is not useful

for teachers.

A7: Bias Identification and Management and *A8: Handling Information and Quality Control*

These both seem completely redundant with others.

A9: Analysis of Information

A9 corresponds with *CASL*, Key 4—effective communication, both grading and

standardized tests.

In Case 2, the teacher gives a research paper 200 points. This should probably be

changed—it's a scoring reliability issue. Would two teachers give the same performance the same

number of points? Would teachers be able to define the difference between, say, 150 and 151?

Perhaps A9 could benefit from additional guidelines for grading (based on the premise

that the goal of a grade is to represent, as accurately as possible, student achievement as of a

point in time). Additional guidelines might be:

- Don't include zeros for missing work. If there is enough information to make a good judgment about level of achievement without the missing work, assign the grade. If there is not enough information, no grade can be given.
- Don't include non-academic factors, such as effort, into the grade.
- Don't include practice work into a grade; give descriptive feedback on practice work.
- Replace out-of-date information with current information.

A10: Justified Conclusions

A10 seems redundant with the others.

Figure 1: Seven Strategies of Assessment *for* Learning

Where Am I Going?

Strategy 1: Provide students with a clear and understandable vision of the learning target. Students are more likely to learn something if they know what it is they are supposed to learn. Sometimes it's enough to merely tell students what they're supposed to learn, for example "we're learning to read decimals and put them in order." Sometimes learning targets need to be worded in student-friendly language. For complex learning targets, a rubric or scoring guide helps define the target.

Strategy 2: Use examples and models of strong and weak work. "Carefully chosen examples of the range of quality can create and refine students' understanding of the learning goal by helping students answer the questions, 'What defines quality work' and 'What are some problems to avoid?'" (Chappuis, 2009, p. 13). Depending on the learning target, selected response items and item formulas can be useful here as well as samples of work that illustrate various levels of a rubric.

Where Am I Now?

Strategy 3: Offer regular descriptive feedback. "Effective feedback shows students where they are on their path to attaining the intended learning. It answers for students the questions, 'What are my strengths?'; 'What do I need to work on?'; and 'Where did I go wrong and what can I do about it?'" (Chappuis, 2009, p. 13). Effective feedback is descriptive (not evaluative), keyed to what students are specifically to learn, given in time to do something about, and is limited in amount to what students can effectively act on at a time.

Strategy 4: Teach students to self-assess and set goals. "The information provided in effective feedback models the kind of evaluative thinking we want students to be able to do themselves. Strategy 4 teaches students to identify their strengths and weaknesses and to set goals for further learning" (Chappuis, 2009, p. 13) Self-assessment and goal-setting can be done with both selected response and performance assessments.
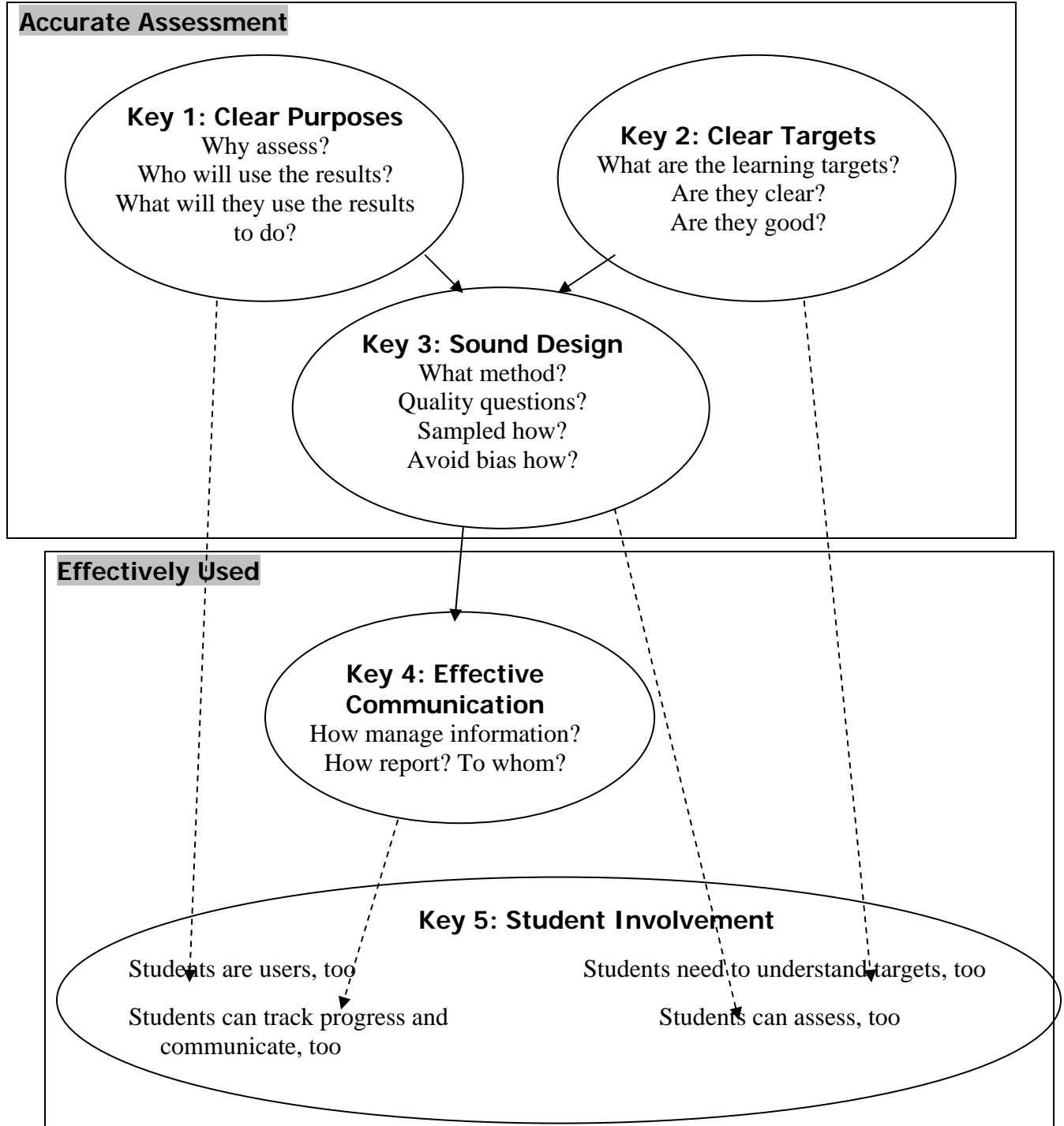
How Can I Close the Gap?

Strategy 5: Design lessons to focus on one learning target or aspect of quality at a time. Strategy 5 involves targeting instruction to the specific needs identified by an assessment.

Strategy 6: Teach students focused revision. "This is a companion to Strategy 5—when a concept, skill, or competence proves difficult for students, we can let them practice it in smaller segments, and give them feedback on just the aspects they are practicing. This strategy allows students to revise their initial work with a focus on a manageable number of learning targets or aspects of quality" (Chappuis, 2009, p. 13).

Strategy 7: Engage students in self-reflection, and let them keep track of and share their learning. "Long-term retention and motivation increase when students track, reflect on, and communicate about their learning. In this strategy, students look back on their journey, reflecting on their learning and sharing their achievement with others" (Chappuis, 2009, p. 14).

Figure 2: Five Keys to Quality Classroom Assessment

Accurate Assessment

Key 1: Clear Purposes
Why assess?
Who will use the results?
What will they use the results to do?

Key 2: Clear Targets
What are the learning targets?
Are they clear?
Are they good?

Key 3: Sound Design
What method?
Quality questions?
Sampled how?
Avoid bias how?

Effectively Used

Key 4: Effective Communication
How manage information?
How report? To whom?

Key 5: Student Involvement

Students are users, too

Students need to understand targets, too

Students can track progress and communicate, too

Students can assess, too

Stiggins, et. al, (2004), p. 13

Figure 3: Keys to Quality Classroom Assessment—Questions to ask of any assessment

Key 1: Clear Purpose
Does the assessment author have a clear picture of how the assessment results will be used and by whom?
- Is it clear who will use the results and how they will be used?
- Is it clear that the assessment can accomplish its intended purpose?
- Do the intended uses fit into a bigger picture of assessments *for* and *of* learning over time?

Key 2: Clear Targets
Does the assessment author have a clear picture of what she or he is trying to measure?
- Are the student learning targets stated?
- Are they clear?
- Is the match between the stated learning targets and what is on the assessment clear?
- Is there a clear connection between the learning targets and state/provincial/district standards?

Key 3: Sound Assessment Design
Has the assessment developer translated the learning targets into assessments that will yield accurate results?
Selecting an Assessment Method
- Will the assessment method(s) chosen accurately reflect the learning target(s) to be assessed?

Sampling
- Do the learning targets represent what was taught/ (Or what will be taught?)
- Does the relative importance of each learning target match its relative importance during instruction?
- Is the sample size large enough to inform the decisions intended to be made, or is it part of a larger plan to gather evidence over time?

Item Quality
- Do the assessment items, exercises, scoring procedure, and scoring guides/rubrics adhere to standards of quality?

Avoiding Potential Sources of Bias and Distortion
- Is there anything in the assessment itself or in the conditions under which it is administered that could lead to inaccurate estimates of student learning?
- Will accommodations made for diverse student characteristics result in any distortions in the final judgment of student learning?
- Are instructions present, clear, and concise?

Key 4: Effective Communication
Has the assessment developer planned for managing information from the assessment adequately and reporting it in ways that will meet the needs of users?
- Are results communicated in a timely manner?
- Will the users of the results understand them and see the connection to learning?
- Do the results provide clear direction for what to do next?

Key 5: Student Involvement
When appropriate:
- Will students be able to use the results to reflect on their learning?
- Are the learning targets written in terms students will understand?
- Are students involved in the assessment process?
- Is the assessment designed so that students can use the results to self-assess and set goals?
- Is there a mechanism in place for students to track their own progress on learning targets and to participate in communicating their status to others?

Figure 4: Learning Outcomes of the *CASL* Program

1. Clear Purposes Assessment processes and results serve clear and appropriate purposes	• Teachers understand who the users and uses of classroom assessment information are and know their information needs. • Teachers understand the relationship between assessment and student motivation and craft assessment experiences to maximize motivation. • Teachers use classroom assessment processes and results formatively (assessment *for* learning). • Teachers use classroom assessment results summatively (assessment *of* learning) to inform someone beyond the classroom about students' achievement as of a particular point in time. • Teachers have a comprehensive plan over time for integrating assessment *for* and *of* learning in the classroom.
2. Clear Targets Assessments reflect clear and valued student learning targets	• Teachers have clear learning targets for students; they know how to turn broad statements of content standards into classroom-level targets. • Teachers understand the various types of learning targets they hold for students. • Teachers select learning targets focused on the most important things students need to know and be able to do. • Teachers have a comprehensive plan over time for assessing learning targets.
3. Sound Design Learning targets are translated into assessments that yield accurate results	• Teachers understand what the various assessment methods are. • Teachers choose assessment methods that match intended learning targets. • Teachers design assessments that serve intended purposes. • Teachers sample learning appropriately in their assessments. • Teachers write assessment questions of all types well. • Teachers avoid sources of mismeasurement that bias results.
4. Effective Communication Assessment results are managed well and communicated effectively	• Teachers record assessment information accurately, keep it confidential, and appropriately combine and summarize it for reporting (including grades). Such summary accurately reflects current level of student learning. • Teachers select the best reporting option (grades, narratives, portfolios, conferences) for each context (learning targets and users). • Teachers interpret and use standardized test results correctly. • Teachers effectively communicate assessment results to students. • Teachers effectively communicate assessment results to a variety of audiences outside the classroom, including parents, colleagues, and other stakeholders.
5. Student Involvement Students are involved in their own assessment	• Teachers make learning targets clear to students. • Teachers involve students in assessing, tracking, and setting goals for their own learning. • Teachers involve students in communicating about their own learning.

References

Arter, J., & J. Chappuis. 2006. *Creating & recognizing quality rubrics*. Portland, OR: ETS Assessment Training Institute.

Assessment Reform Group. 2002. *Testing, motivation and learning*. Cambridge, UK: University of Cambridge, Faculty of Education.

Atkin, J.M, P. Black, & J. Coffey. 2001. *Classroom assessment and the National Science Standards.* Washington, DC: National Academy Press.

Black, P., & D. Wiliam. 1998. Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bloom, B. 1984. The search for methods of group instruction as effective as one to one tutoring. *Educational Leadership, 41*(8), 4–17.

Brookhart, S. M. 2001. Successful students' formative and summative uses of assessment information. *Assessment in Education, 8*(2), 153–169.

Brookhart, S. M. 2004. *Grading*. Upper Saddle River, NJ: Pearson Education.

Brookhart, S. M., & J. DeVoge. 1999. Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*(4), 409–425.

Brookhart, S. M., & D. Durkin. 2003. Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement in Education, 16*(1), 27–54.

Butler, R. 1988. Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation of interest and performance. *British Journal of Educational Psychology, 58*, 1–14.

Butler, R., & O. Newman. 1995. Effects of task and ego-achieving goals on help-seeking behaviours and attitudes. *Journal of Educational Psychology, 87*(2), 261–271.

Chappuis, J. 2007. Learning team facilitator handbook: A resource for collaborative study of classroom assessment for student learning. Portland, OR: ETS Assessment Training Institute.

Chappuis, J. 2009. *Seven strategies of assessment* for *learning*. Portland, OR: ETS Assessment Training Institute.

Chapppuis, S., R.J. Stiggins, J. Arter, and J. Chappuis. 2005. *Assessment FOR learning: An action guide for school leaders*. Portland, OR: ETS Assessment Training Institute.

Guskey, T.R., and J. Bailey. 2001. *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin.

Costa, A. L., & B. Kallick. 2004. *Assessment strategies for self-directed learning.* Thousand Oaks, CA: Corwin.

Crooks, T. 2001. The validity of formative assessments. Paper presented at the 2001 Annual Meeting of the British Educational Research Association, Leeds, UK, September 13–15.

Dweck, C. S. 2001. *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Psychology Press.

Dweck, C. S. 2006. *Mindset: The new psychology of success.* New York: Random House.

Dweck, C. S. 2007. The secret to raising smart kids. *Scientific American Mind,* November 28, 2007. Retrieved November 12, 2008 from http://www.sciam.com/article.cfm?id=the-secret-to-raising-smart-kids&print=true

Hattie, J., & H. Timperley. 2007. The power of feedback. *Review of Educational Research.* Retrieved October 9, 2007 from http://rer.sagepub.com

Kluger, A. N., & A. DeNisi. 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Morgan, A. 2008. *Feedback: Assessment for rather than of learning.* Retrieved September 3, 2008 from http://www.bangor.ac.uk/adu/the_scheme/documents/FEEDBACKJanuary06_000.ppt

Militello, M. 2005. Two much information: A case study of assessment accountability in an urban school district. Paper presented at the annual meeting of AERA, Montreal, Canada.

O'Connor, K. 2007. *A repair kit for grading: 15 fixes for broken grades.* Portland, OR: ETS Assessment Training Institute.

Sadler, D.R.. 1989. Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.

Shepard, L. A. 2000. The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shepard, L. A. 2008. Formative assessment: Caveat emptor. In C. Dwyer (ed.), *The Future of Assessment: Shaping Teaching and Learning* (pp. 279–303). NY: Lawrence Erlbaum Associates.

Stiggins, R.J. 2009. Assessment manifesto: A call for the development of balanced assessment systems. Portland, OR: ETS Assessment Training Institute.

Stiggins, R., J. Arter, J. Chappuis, & S. Chappuis. 2006. *Classroom assessment for student learning: Doing it right—Using it wel*l. Portland, OR: ETS Assessment Training Institute.

Wei, R.W., L. Darling-Hammond, A. Andree, N. Richardson, and S. Orphanos. February 2009. Professional learning in the learning profession: A status report on teacher development in the U.S. and abroad. Dallas, TX: National Staff Development Council.

White, B. Y., & J. R. Frederiksen. 1998. Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3–118.

Zimmerman, B. 2008. Investigating self-regulation and motivation: historical background, methodological development, and future prospects. American Educational Research Association Journal, 45(1), 166–183.

Appendix A: CLASSROOM ASSESSMENT QUALITY RUBRICS

KEY 1: CLEAR PURPOSES		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• The intended users and uses are identified. • It is clear how the intended use fits into an overall plan that includes both assessments *for* and *of* learning over time.	• The intended users and uses can be inferred. • How the intended use fits into an overall plan that includes both assessments *for* and *of* learning is not stated, but can be inferred.	• The intended users and uses are not identified and cannot be accurately inferred. • It is difficult to tell how the intended use fits into an overall plan that includes both assessments *for* and *of* learning.

KEY 2: CLEAR TARGETS		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• Learning targets measured by the assessment are stated. • The learning targets are clear. • The match between stated learning targets and what is on the assessment is clear. • Learning targets are clearly connected to the state/provincial/district content standards.	• Learning targets measured by the assessment are stated or can be inferred easily from the assessment. • Learning targets may be somewhat unclear. • There is a partial match between stated learning targets and what is on the assessment. • Learning targets are partially or loosely connected to the state/provincial/district content standards.	• Learning targets measured by the assessment are not stated and cannot be accurately inferred. • Learning targets are stated, but vague or unclear. • There is no apparent match between stated learning targets and what is on the assessment. • There is no apparent connection between learning targets and the state/provincial/district content standards.

KEY 3: SOUND DESIGN-- SELECTING AN ASSESSMENT METHOD		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• The method(s) chosen is/are capable of accurately reflecting the learning target(s) to be assessed.	• When multiple methods are used in an assessment, sometimes the proper method is selected; sometimes an improper method is used.	• The method(s) used is/are not capable of accurately reflecting the learning targets(s) in question.

KEY 3: SOUND DESIGN-- SAMPLING		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• The learning targets tested represent what was taught. (In the case of diagnostic assessment, they represent what is intended to be taught.) • The relative importance of each learning target on the assessment matches the relative importance given to it during instruction. • The sample size is large enough to inform the decisions intended to be made, or it is clear that it is part of a larger plan to gather enough information over time.	• The learning targets tested partially represent what was taught. (In the case of diagnostic assessment, they partially represent what is intended to be taught.) • The relative importance of each learning target on the assessment partially matches the relative importance given to it during instruction.	• The learning targets tested do not represent what was taught. (In the case of diagnostic assessment, they do not represent what is intended to be taught.) • The relative importance of each learning target on the assessment does not match the relative importance given to it during instruction. • The sample size is not large enough or is substantially greater than needed to inform the decisions intended to be made.

KEY 3: SOUND DESIGN--ITEM QUALITY		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• Selected response and short answer items adhere to standards of quality (See *CASL* pp 138 – 146). • Extended written response items adhere to standards of quality (See *CASL* pp 174 - 177). • Extended written response scoring procedures adhere to guidelines for quality (See *CASL* pp 177 - 179). • Performance assessment tasks adhere to standards of quality (See *CASL* pp 218 – 229). • Performance assessment scoring guides/rubrics adhere to guidelines for quality (*CASL* pp 200 - 218). • Oral examination assessments adhere to guidelines for quality (See *CASL* p 265).	• Selected response and short answer items partially adhere to standards of quality. • Extended written response items partially adhere to standards of quality • Extended written response scoring procedures partially adhere to guidelines for quality • Performance assessment tasks partially adhere to standards of quality • Performance assessment scoring guides/rubrics partially adhere to guidelines for quality. • Oral examination assessments partially adhere to guidelines for quality.	• Selected response and short answer items do not adhere to standards of quality. • Extended written response items do not adhere to standards of quality. • Extended written response scoring procedures do not adhere to guidelines for quality. • Performance assessment tasks do not adhere to standards of quality. • Performance assessment scoring guides/rubrics do not adhere to guidelines for quality. • Oral examination assessments do not adhere to guidelines for quality.

KEY 3: SOUND DESIGN-- BIAS AND DISTORTION		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• There is nothing in the assessment itself or in the conditions under which it is administered that could lead to inaccurate estimates of student learning (See CASL pp 114- 116). • Accommodations made for diverse student characteristics do not result in any distortions in the final judgment of student learning. • Instructions are present, clear, and concise.	• There are a few things in the assessment itself or in the conditions under which it is administered that could lead to inaccurate estimates of student learning. • Accommodations made for diverse student characteristics may result in slight distortions in the final judgment of student learning. • Instructions are present, but not as clear or concise as they could be.	• There are many things in the assessment itself or in the conditions under which it is administered that could lead to inaccurate estimates of student learning. • Accommodations made for diverse student characteristics will result in distortions in the final judgment of student learning. • Instructions are not present, or are difficult to follow.

KEY 4: EFFECTIVE COMMUNICATION		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• Results from the assessment are communicated in a timely manner. • Results from the assessment are communicated so that the intended users of the information understand what they mean and how they connect to learning. • Results provide clear direction for what to do next.	• Results from the assessment are communicated so that the intended users of the information may have questions about what they mean or how they connect to learning. Results do not provide entirely clear direction for what to do next.	• Results from the assessment are not communicated in a timely manner. • Results from the assessment are not communicated so that the intended users of the information understand what they mean or how they connect to learning. • Results do not provide direction for what to do next.

KEY 5: STUDENT INVOLVEMENT		
5—Ready to Use	3—On Its Way	1—Needs Significant Work
• When appropriate, learning targets are written in terms that students can clearly understand. • When appropriate, students are involved in the assessment process. • When appropriate, the assessment is designed so that students can use the results to identify specific strengths and weaknesses, and to set meaningful goals for further learning. • There is a mechanism in place for students to track their own progress on learning targets and to participate in communicating their status to others, if appropriate for the context.	• Learning targets would be somewhat unclear to students, but could be revised to be clear. • Student are minimally involved in the assessment process, when it would be appropriate to do so. • The assessment design would need some work for students to be able to use it easily to identify specific strengths and weaknesses, and to set meaningful goals for further learning. • There is some assistance for students in tracking progress and communicating progress, but either it has to be inferred, or it is not complete.	• Learning targets would have to be entirely rewritten so that students could understand them • Students are not involved in the assessment process, when it would be appropriate to do so. • Students cannot use the assessment results to identify specific strengths and weaknesses, or to set meaningful goals for further learning. • There is not a mechanism in place for students to track their own progress on learning targets and to participate in communicating their status to others, when it is appropriate for the context.