CHAPTER 6

Assessing Small-Group Learning

INTRODUCTION AND LEARNING OBJECTIVES

Assessment plays a key role in educational accountability. Being able to assess the outcomes of students' learning is very important, and probably more so for such pedagogical practices as cooperative learning where responsibility for learning is devolved to the group and where teachers act as facilitators of learning rather than instructors of knowledge. With this approach to learning, teachers need to be able to assess how students are managing the learning process (process learning) as well as what they are achieving (outcomes of learning) if they are to make changes to how they teach and how students learn. This is particularly important given the accountability requirements of the No Child Left Behind (NCLB) Act of 2001 that requires schools to close the gap between high- and low-performing students not only overall but also between minority and nonminority students and between disadvantaged students and their more advantaged peers (Kim & Sunderman, 2005; see also discussion in Chapter 1 of this volume). While the research on the academic and social benefits of cooperative learning is unequivocal for students, generally, and specifically for those in minority groups such as second-language learners, the ethnically diverse, and students with special learning needs (Cohen, 1994; Johnson & Johnson, 2002; Putnam, Markovchick, Johnson, & Johnson, 1996; Sharan, 1990; Slavin & Cooper, 1999), effective assessment practices require that these benefits be documented so teachers can communicate them to parents, students, and reporting authorities. Moreover, by doing so, teachers are able to reflect on their own teaching practices and determine what they may need to adjust or change to promote improvement in students' learning. This is important because research clearly indicates that teachers become committed to new practices after they have actively engaged in using

them in their classrooms and seeing the changes in students' achievements for themselves (Guskey, 2002; Warfield, Wood, & Lehman, 2005). The purpose of this chapter is to present an overview of different assessment practices and the types of rubrics that can be used to assess learning, as well as examples of rubrics used to assess different higher-order thinking and problem-solving skills that students demonstrate.

When You Have Finished This Chapter You Will Know:

- The difference between formative (process) and summative (outcomes) assessments
- How different types of formative and summative assessments can be used to inform the teaching and learning process
- How to design different rubrics to assess the processes and outcomes of learning
- The importance of conducting authentic assessments
- How to design rubrics to assess higher-order thinking and problem-solving skills

Case Study 6.1

Teachers' Reports on How They Assess Small-Group Learning

I: Can you tell me a bit about how you assess students' learning during small-group work?

T1: I use a variety of checks to monitor how they're working. I do a lot of my work informally where I cruise around the room and keep an eye on what's happening in each group. I can tell at a glance who's contributing and who's loafing. Sometimes I'll get them to do specific tasks or I'll give them different colored pens so I can see each person's contribution. I find the informal checking that I do really helps me to get a handle on what's going on.

I: What about you, Dani?

T2: I have some checklists that I've developed like, who's contributing, who's helping others, and so on to give me a way of monitoring what's happening. I encourage the students to work together with everyone having a chance to present ideas, work and so on, so I like to see that happening when they're working together. I also have checklists that I give the students—every now and then so they can evaluate themselves. When they do this, we spend time in class debriefing each other so we can hear what's happening and what others are doing.

T3: I've used both those approaches (referring to the comments by T1 & T2). I also think it's important to keep a tight handle on what they're actually learning as they go along so I have little checks that I do on their learning. Sometimes, I'll check on where they're up to—as a bit of a check on their progress with the task. . . . When they finish, I always get them to give a group presentation to the class so we can all see what they've accomplished. It's pretty hard to hide then because each person has to present a piece of the work, so they've got to know it.

T1: Yes, I agree. I do a lot of the things that have been mentioned and I also do things a bit more formally. For example, when they do their group presentation, I give them a *rubric* of what they're going to be evaluated on. We go over this before they start so they know what's expected. The sorts of things I look for include making sure they've addressed the topic . . . clear presentation of all the content and the presentation itself has to be interesting and designed to capture the interest of others. So for a PowerPoint display, I'd be looking at the visual presentation and what they did to capture my interest.

T2: One of the best ideas I read about was to get the class to provide the group with feedback on their presentation—how interesting it was, did they cover all the key areas, did everyone contribute, and so on. The kids like doing this, and I've found that provided they have clear guidelines on how they are to provide feedback . . . so it doesn't become destructive, they manage it OK and the groups seem to listen to what their peers have to say.

Note: I = Interviewer; T = Teacher

In Case Study 6.1, the teachers report that they use a number of different approaches to assessing students' work during their small-group learning activities. For example, the first teacher (T1) uses different informal approaches to check on her students' learning (Turn 2), while the other two teachers (T2 & T3) report using more formal checklists or procedures to help students evaluate their learning (Turns 3 & 4). Interestingly, T1 indicates that she also

presents her students with a rubric that outlines what they are going to be evaluated on (Turn 5), while T2 provides her students with opportunities to give constructive feedback to their peers (Turn 6). As can be seen from the above discussion with the interviewer, the teachers are using a variety of approaches to assess both how their students are learning and what they have achieved.

Formative Assessment

The informal approaches adopted by T1 often allow teachers to obtain some quick feedback on how students are managing the learning process so changes can readily be made to task requirements or additional scaffolding can be provided if students appear to be experiencing difficulties. This is in contrast to the more formal approach adopted by T2, where she is using checklists either to keep a record of who is contributing to the group or to encourage her students to evaluate themselves. Both these approaches are acceptable and are designed to provide information on how students are managing their learning. The informal approach used by T1 and the more formal ones adopted by T2 are generally referred to as *formative assessments* because they are designed to provide information that informs the ongoing teaching and learning process to ensure its effectiveness and to ascertain whether other approaches should be adopted or whether expectations for the task should be changed (McInerney & McInerney, 1998).

There are many different types of formative assessments that can be used to provide information on students' learning, from those that consider how students interact, the types of discourse they use to facilitate understanding and learning, through to those that tap students' perceptions of how they are learning, and so on. These assessments generally require teachers to be observant and move around the room so they can listen to the discussions occurring in groups and when appropriate intervene to challenge students' perceptions or scaffold their understandings. Both actions require teachers to be active in the learning process to ensure that misconceptions are corrected or developing understandings are consolidated.

Curriculum-Based Assessments

Another type of formative assessment is *curriculum-based assessment* (CBA). This approach uses a variety of probes or stimulus questions to gauge students' understandings or progress at some point during their participation in the

curriculum. Paris and Hoffman (2004), in a review of the different types of reading assessments teachers use from kindergarten through third grade, found that they ranged across observations, anecdotal evidence, informal inventories, and work samples as the main source of evidence for students' reading achievement and progress. Interestingly, although teachers had access to some very sophisticated standardized reading tests, most regarded informal measures or formative assessments that they design, select, and embed in the curriculum as more useful for teachers, students, and parents than commercial instruments. Paris and Hoffman argue that because a single reading assessment cannot capture the variety of skills and developmental levels of children in most K–3 classes, teachers use multiple assessments and choose those that fit their purposes.

Graves, Plasencia-Peinado, Deno, and Johnson (2005) used a series of curriculum-based measures of reading (randomly selected first-grade reading passages that had been standardized to ensure equal levels of readability) and a list of nonsense words to measure the oral reading fluency of a group of English learners in Grade 1 and found that all students (high, medium, and low readers) were able to demonstrate improvements on both measures. These findings were consistent with findings from other studies examining the growth in oral reading of first-grade students who are not English learners. Graves et al. argued that the curriculum-based measures used in this study provide valuable information to first-grade teachers on the progress of English learners. Moreover, the authors argued early use of these curriculum-based progress measures can signal the need for intense preventative intervention that can assist teachers in avoiding the erroneous placement of some English learners in special education.

Similarly, Weissenberger and Espin (2005), in an evaluation of curriculum-based measures (students were required to write two stories in response to a sentence prompt) used to assess students' writing proficiency for Grades 4, 8, and 10, found moderate to strong positive correlations between the curriculum-based measures used and the standardized Wisconsin Knowledge and Concepts Examinations (WKCE) scores for language arts and writing assessment, providing evidence of the validity of the curriculum-based measures they used. Furthermore, the alternative-form reliability correlations of the two writing samples (i.e., the writing samples were compared) were high, suggesting that teachers can have confidence in the stability of these curriculum-based measures. These findings led the authors to propose that these curriculum-based writing measures can be used to assess the developing writing skills of students at the elementary, middle, and high school levels.

Curriculum-based assessments have also been used successfully in mathematics. In a study that investigated the extent to which teachers' use of different instructional practices related to improved performance in mathematics for

Grade 10 students, McCaffrey et al. (2001) found that the use of *reform practices,* as opposed to traditional practices, were positively related to achievement gains for students. The reform practices that the authors identified included the following:

- Students worked in small groups.
- Students used manipulatives in solving problems. Students were encouraged to consider alternative methods for solving problems.
- Students were required to explain reasons for their answers.
- Students worked on solving real-world problems.
- Students worked in small groups on extended mathematics investigations of projects.
- Students made formal presentations to the class of their group's project.
- Students wrote reflections on what they learned from their experiences.
- Students were tested using open-ended responses requiring them to provide descriptions and explanations.
- Students were required to complete performance tasks both individually and in groups.

These reform practices were in contrast to the traditional practices for teaching mathematics where teachers predominantly explained problems to the class and students worked from textbooks or worksheets (often individually) to practice mathematics facts, rules, or formulas. In these classrooms, tests were predominantly short answer, true or false, or fill in the blanks.

In short, McCaffrey et al. (2001) demonstrated that when student assessment practices are curriculum based and are linked to teaching practices that are student-centered and inquiry based, students obtain higher achievement gains as measured on traditional standardized tests than peers who are taught by traditional practices. The authors argued that the results should alleviate teachers' concerns that reform-based teaching will adversely affect students' performances on standardized tests, because that did not happen in this study.

Peer Assessment

A further type of formative assessment may involve peer assessment where students are asked to consider the level, value, or quality of a product produced by their peers. This type of assessment appears to be most helpful if it provides rich and detailed qualitative feedback about strengths and aspects that need to be improved rather than a mark or a grade. Topping, Smith, Swanson, and

Elliot (2000) used a formative peer assessment process to provide feedback to postgraduate university students on the quality of the academic reports they had compiled as a requirement of their educational psychology training program. Students were paired so that each acted as assessor and assessee to the other and were provided with the criteria staff had previously used to assess the academic reports they were evaluating. The results indicated that there was a balance between positive and negative statements between staff and peer assessments, and although only half of the detailed formative assessment statements made showed some degree of correspondence between staff and peers, Topping et al. (2000) concluded the reliability and validity of this type of assessment appeared adequate. Moreover, the overlap in detail between staff and peer assessment suggested that peer assessments add value to the feedback provided to students. In turn, students reported that although the process of providing formative feedback to their peers was time-consuming, intellectually challenging, and socially discomforting, it was effective in improving the quality of their own subsequent written work and it helped them to develop other transferable skills that would generalize to their own future writing. It appeared that the peer assessment process prompted self-assessment among many of the students, and this coupled with the obvious pressure to spend time scrutinizing and clarifying the assessment criteria contributed to a greater understanding of effective ways to learn new content, the importance of structure and organization, and the reader's perspective in receiving a written report.

Computer-Supported Peer Assessment

Formative peer assessment can also be conducted in a computer-supported collaborative (CSCL) environment. Prins, Sluijsmans, Kirschner, and Strijbos (2005) report on such a study where students, enrolled in an online virtual seminar series, worked in multidisciplinary teams investigating sustainable development and enlargement of the European Union. As part of the seminar series, students were required to discuss the assessment criteria they had been given and adjust it to their group's needs, assess a draft of their own report and a fellow group's report using the assessment criteria, and, finally, write a reply to a fellow group indicating how they had revised their report according to the feedback they had received. The formative assessment criteria the students used as a template (which they were able to adjust to their group's needs) included both content-related criteria (e.g., application of result, quality of language used, links to different knowledge bases, summary) and process-related criteria (e.g., planning research, planning individual tasks, cooperation within the

group). Results showed that students' attitude to peer assessment was positive and many felt quite comfortable about providing feedback to students they had never met. Prins et al. noted, however, that the quality of the assessment reports was low, with more negative statements rather than positive ones, probably because the majority of students did not apply the provided criteria and feedback rules. This resulted in some students' having difficulty accepting critical feedback, although they also indicated that peer feedback was valuable for the revision of their report. These findings led Prins and colleagues to suggest that students not only needed rules for providing formative feedback but they also needed more support or explicit rules for receiving or accepting feedback and coping with it.

Weinberger, Ertl, Fischer, and Mandl (2005) used both epistemic and social scripts to support students' dialoguing in computer-supported collaborative learning. *Epistemic scripts* specify and help sequence knowledge construction activities. They are designed to guide learners toward specific aspects of the task and toward specific task-orientated activities, while social scripts are designed to help students learn how to elicit information from each other during collaborative learning. Palincsar and Herrenkohl (2002) used an epistemic script to guide the audience's feedback to the group on the presentation of their explanations about the scientific phenomena they had been investigating. The *cognitive roles* different members of the audience adopted required some to provide feedback on the group's presentation on the clarity of the relationships between the predictions and the theories; others were responsible for ascertaining clarity of the summary of the findings, and others were responsible for determining if the reporter discussed the relationships among the group's predictions, theory, and findings. Palincsar and Herrenkohl reported that the use of these cognitive roles and the associated tools (scripts) for providing feedback promoted student dialogue, advanced students' theorizing, influenced their thinking, and promoted conceptual understanding.

The epistemic script students used in the Weinberger et al. (2005) study involved students' discussing a series of case problems that was presented online as text and jointly preparing an analysis of each. In order to facilitate the learning process, students were provided with a series of prompts or questions that suggested they apply specific theoretical concepts to the problem cases. In addition, students were provided with social scripts that aimed to foster critical negotiation and to discourage quick decisions and arriving at a false consensus. Hence, each student was assigned two roles; one as an analyst of one of the cases discussed, and the other as a constructive critic on the remaining cases. Both these roles were supported by online social script prompts. Interestingly, Weinberger et al. found that the epistemic script

actually impeded the individual acquisition of knowledge whereas the social script facilitated it. In a second study where students worked online in pairs and teleconferenced about the task they were assigned, the results once again showed that students using the epistemic script learned less than those without it, while the social script enhanced learning. Weinberger et al. concluded that in both studies, the epistemic scripts appeared to hinder the individual acquisition of knowledge if students were not sufficiently motivated to engage in joint elaborative discussions of the learning task. In contrast, social scripts appear to change the interaction pattern and motivate learners to critique the contributions of the learning partners more strongly; as a consequence, they acquire more individual knowledge than learners who do not have this type of support. These findings prompted Weinberger et al. to suggest that careful consideration needs to be given to the type of epistemic scripts that are used in virtual environments and how they can be used to foster elaboration of the group task.

Given the above findings, what are the implications for teachers who need to be able to assess how students learn and what they achieve as they work cooperatively? The following summarizes the research on formative assessments:

- Process learning involves monitoring how students are managing to learn in their groups.
- Formative assessments can involve both informal and formal approaches to assessing the learning process.
- Informal formative assessments may include observations, *anecdotal records*, and work samples.
- Formal formative assessments may include checklists, inventories, and progress reports.
- Curriculum-based assessments have high validity as they are designed to match closely what students are learning in the curriculum.
- Peer assessments are most helpful if they provide rich and detailed qualitative feedback.
- Peers learn from being able to assess their own and other's work.
- Computer-supported peer assessment presents difficulties unless peers are prepared to engage with the task and receive training in how to provide feedback.
- Epistemic and social scripts can be used to guide students' learning and feedback during computer-supported learning; however, social scripts are more successful at promoting learning, possibly because they may support interaction of learners, which, in turn, appears to facilitate individual knowledge acquisition and learning.

Practical Activity

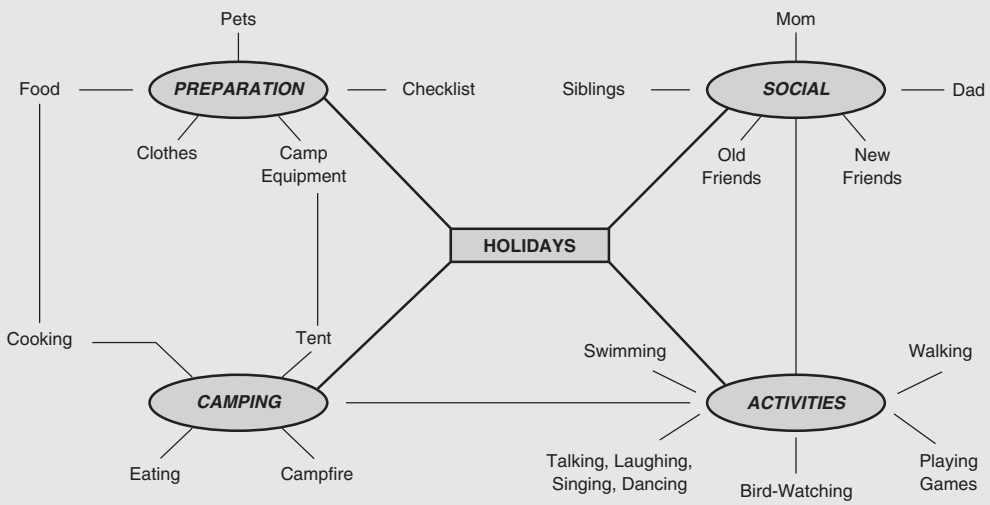
Conducting Formative Assessments of Small-Group Learning

Elementary School

★ Teacher self-reflections on student learning. Teachers will often ask the following types of formative evaluation questions as they monitor groups: How well are the students working together? What are they learning? What can they do? What do they need to know? How are they providing feedback to each other? What else might facilitate their learning?

★ Team word webbing and how to assess the group's efforts. Students are provided with different color pens for this task. The topic for discussion is written in the center of a large piece of butcher paper. The students take turns contributing their ideas to the web, focusing on key ideas and supporting elements. The completed web enables the teacher to have a ready tool for assessing the group's functioning, and because individual members have contributed their ideas in color, it is easy to identify each student's contribution. Requirements can be established for each of the component parts so that students may be required to identify a minimum of three or four key ideas (preparation, camping, social, activities) and a similar number of supporting elements (food, clothes, pets, camp equipment). It is possible to construct a rubric that acts as a checklist for the students and a tool for the teacher to evaluate the group's productivity.

Word Webbing



★ Student reflections on their learning. The following Know-Want-Learned (KWL) is designed to help structure the group's planning so students can clearly identify what they need to learn in the context of what they currently know and what they want/need to learn. This exercise is very helpful in getting students to make links between current knowledge and future needs to know.

Know: *What do we know?*	***Want:*** *What do we want to know?*	***Learned:*** *What have we learned?*

★ Strategic Questioning (King, 1999). The following questions are designed to help students focus on the task so they can plan, monitor, and evaluate their progress. The questions can be constructed as a rubric so that children have the questions down the left side and their responses on the right, or they can be given the questions on prompt sheets or cards:

Planning:

- What is the problem?
- What are we trying to do here?
- What do we know about the problem so far?
- How can this help us?
- What is our plan?
- Is there another way to do this?
- What should we do next?

Monitoring:

- Are we using the best plan?
- Do we need a different plan or strategy?
- Has our goal changed?
- Are we on the right track?

(Continued)

(Continued)

Evaluating:

- What worked?
- What didn't work?
- What would we do differently next time?

★ Group's action plan. This action plan is designed to help students plan the task, including who will do what and by when. The evaluation section is to enable students to keep a record of where they are and what they have achieved.

Group's Action Plan

Goal: Our goal is:

_____

_____

_____

_____

_____

_____

★ Student reflections on how their group handled the process of learning. The following two surveys are designed to tap student' perceptions of how their groups worked.

Task	*Subtasks*	*Who Does What?*	*Time*	*Evaluation*		
				Finished	Not Finished	Not Attempted
Comments:						

This Is How My Group Worked Today

1. How well did your group share the work today?

Right on Pretty good OK Not so good Missed the mark

2. How well did your group stay on the task today?

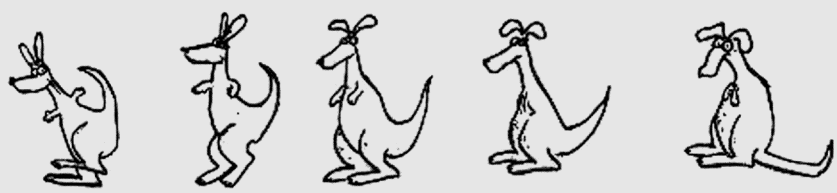
Right on Pretty good OK Not so good Missed the mark

3. What did you do to help one another? Did you . . .

(a) try to recognize ideas in others?

Right on Pretty good OK Not so good Missed the mark

(b) listen to one another?



Right on Pretty good OK Not so good Missed the mark

(Continued)

(Continued)

(c) try to help other members share their ideas?

Right on | Pretty good | OK | Not so good | Missed the mark

(d) take turns in talking and listening?

Right on | Pretty good | OK | Not so good | Missed the mark

4. How did you show others that you were listening to them?
(Check [☐] the things that you did.)

- Nodded to show that you were listening. ☐
- Said: "That's a good idea!" or "That's good" when you liked an idea. ☐
- Asked questions. ☐
- Listened and tried to answer questions. ☐
- Tried sometimes to add information to another member's thoughts or ideas. ☐

5. Overall, how would you rate your group's performance today?

Right on | Pretty good | OK | Not so good | Missed the mark

Thoughts About Group Work Questionnaire

Read each statement below. Rate your response from 1 (almost never happens) to 5 (almost always happens)

1. Group members give each other time to talk and make suggestions.	1	2	3	4	5
2. Group members treat each other nicely.	1	2	3	4	5
3. The ideas of others are important.	1	2	3	4	5
4. Group members often use the ideas of others.	1	2	3	4	5
5. Group members offer help to each other when it is needed.	1	2	3	4	5
6. Group members seek help from each other before asking the teacher.	1	2	3	4	5
7. Group members feel free to talk and make suggestions.	1	2	3	4	5
8. Decisions are made by the group.	1	2	3	4	5
9. Group members do the best they can.	1	2	3	4	5

Middle School and High School

★ Ways of assessing process-related criteria for a project. The following rubric is an adaptation of one used by Prins et al. (2005) to assess the learning process during an online virtual seminar series. Students use the rubric to reflect on their group's learning process.

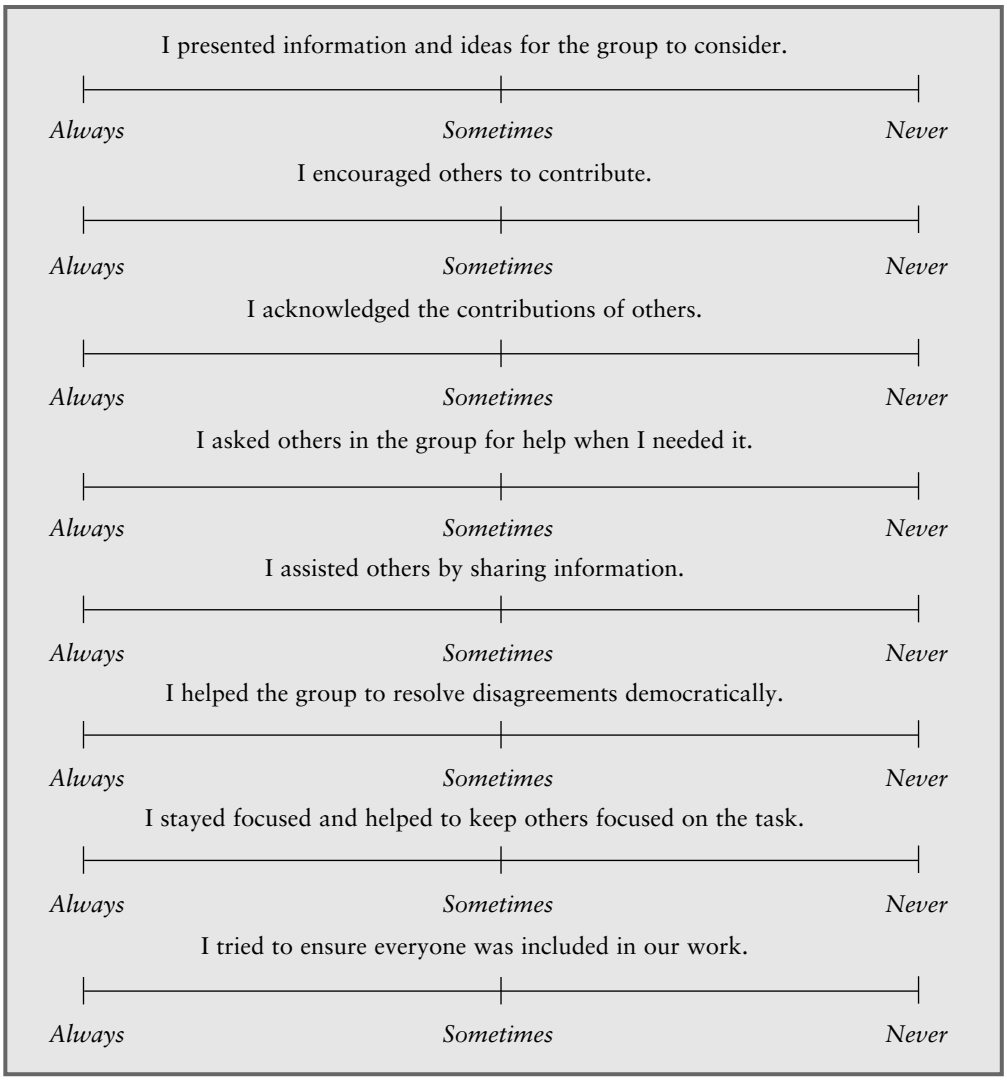
Process-Related Criteria	*Above Average*	*Average*	*Below Average*
Planning	Work was spread out over the time of the project.	Most of the work was done at the end.	One or more of the deadlines were not met.
Individual tasks	The division of labor was fair.	The division of labor was fair but not every student had a reasonable load.	The division of labor was not clear.

(Continued)

(Continued)

Process-Related Criteria	*Above Average*	*Average*	*Below Average*
Cooperation	The discussion boards were used, so decisions are traceable.	Decisions were made, but not every group member participated in the decision.	The group did not work as a group, so decisions were not discussed.
Participation	Each group member contributed to the online discussions and the final report.	Not every group member participated equally to the online discussions and the group report.	Not every group member participated equally, and input into the report differs strongly.
Respond to feedback	The team dealt with comments given by other team members and staff.	The team dealt with the comments received in a way that fits with the rest of the report.	The team did not deal with the comments they received.

- ★ Teacher reflections on how the students are learning in their groups often involve asking the following types of questions: How are the students managing the task? Have they identified the component tasks and planned their schedule to completion? How are they dealing with diverse ideas and opinions? Are they able to work independently as a group or do I need to monitor and intervene when needed? Does everyone have opportunities to express opinions and contribute? How are they utilizing their resources, both personnel and material?
- ★ Samples of the group's planning and small-task sheets will also provide the teacher with ongoing anecdotal information on how the students are managing the learning process.
- ★ Students' perceptions of how they cooperated in their group.



Summative Assessment

A *summative assessment* is designed to measure what students have learned or have accomplished at the end of a period of instruction. In small-group learning, summative assessment provides information on what the group has accomplished by working together on the task; it may include such assessments as a group presentation, a group assignment, or the presentation of a group

product. Summative assessments rarely involve the administration of standardized tests simply because such tests are not usually considered to be valid measures of what students have learned from such an experience. Most summative assessments that are used to measure the outcomes of group learning are criterion referenced and are designed by class teachers to see if students have achieved a predetermined goal, based on specific learning experiences from their curriculum. In group work, criterion-referenced assessments can include curriculum-based measures (discussed previously) and a variety of authentic assessments such as the development of a portfolio of information or ideas, the presentation of a multimedia display, the demonstration of specific skills or outcomes (e.g., presentation of a play), and so on. In addition, teachers may decide to develop a scoring matrix to assess whether the component part of the criteria has been met and if so, how adequately. The scoring matrix may also be weighted so that consideration is given to the processes of learning (discussed previously) as well as to the achievements obtained. In this way, teachers are able to gain some well-documented, highly valid information on the achievements of the group. This is particularly important given the need to be able to demonstrate students' achievements in practices such as cooperative learning where teachers act to facilitate learning and students are expected to be self-directed and active as learners in the construction of knowledge (Hackmann & Schmitt, 1997).

Criterion-Referenced Assessments

In criterion-referenced assessments, achievement goals are clearly specified in a set of criteria so that students understand what they are expected to accomplish. The criteria are usually based on expectations in the different syllabi for students' achievements for specific core learning outcomes and the associated knowledge, processes, and attitudes needed to understand key concepts that have been identified as critical to these core learning areas. Hence, students are assessed on whether they have met the required criteria or not, although many teachers prefer to refer to a criterion as developing (not quite fully developed), developed, or highly developed to provide them with the diagnostic information they need when they are determining what additional learning experiences students may require in order to be able to determine that the criterion is developed.

Teachers often prefer to conduct criterion-referenced assessments because they provide information on what students can or cannot do, which can be very helpful in communicating with parents about a student's progress.

Moreover, as many educational objectives are written in performance terms, the criterion-referenced approach appears more suitable for measuring the achievement of these performance objectives (McInerney & McInerney, 1998). In addition, criterion-referenced assessments can be very useful in helping students to understand what they have accomplished and what else they may need to accomplish in a specific subject or core area.

Authentic Assessments

Another type of criterion-referenced assessment is an *authentic assessment*. Authentic assessments are designed to measure student learning in a real-life context against specific performance criteria. According to Woolfolk (1998), authentic assessments have the following characteristics:

- Involve real-life tasks
- Are contextualized
- Are intellectually challenging
- Involve student's own research
- Assess student's higher-order thinking
- Are engaging and educational
- Require collaboration with others
- Are criterion referenced
- Include student self-reflections
- Use a multifaceted approach to evaluation
- Are compatible with schoolwide aims
- Allow for different student learning styles
- Aim to support and help student learning as necessary

Herrington and Oliver (2000) proposed that authentic assessments need to occur in authentic learning environments if they are to have veracity, integrity, and fidelity. Their argument is that these types of assessments need to reflect the types of problem situations students are likely to encounter in real-life learning environments where problems are generally complex and ill-defined and there are no set procedures to follow. Moreover, in order to solve these problems, students need to have access to a range and diversity of materials that will allow them to explore topics in depth and to apply sustained thinking to their resolution, have access to the critical insights of others (this usually occurs when students work collaboratively), and present their findings as both oral and written reports to their peers. The authentic assessment that was

discussed was based on the results of the investigation the students conducted and was assessed against specified performance criteria that had previously been discussed with them. Herrington and Oliver believe that this type of assessment has high validity and reliability as it is based on the task that the students have completed. Moreover, as the assessment is seamlessly integrated with the task, there is no need for follow-up testing, because the oral and written presentation provides teachers and peers with multiple indicators of the learning that has occurred.

Building on Herrington and Oliver's (2000) work, Gulikers, Bastiaens, and Kirschner (2004) developed a five-dimensional framework for authentic assessment that included (a) the assessment task, (b) the physical context, (c) the social context, (d) the assessment form, and (e) the assessment criteria. According to Gulikers et al., the authentic task, on which authentic assessment is based, is generally complex, ill-structured, and involves multiple solutions, similar to Cohen's (1994) definition of a complex learning task that has no definitive answers or clear-cut solutions. It must be located in a real-life context and collaboration with others must be required to work it out. Moreover, as Herrington and Oliver have stated, the learning task needs to resemble the assessment task, and students need to understand how they are going to be assessed as this will guide their learning. While this five-dimensional framework was developed in the context of professional practice for college students training to be nurses, Gulikers et al. proposed that, overall, it provides a good description of the dimensions and elements that should be taken into account in an authentic assessment.

Using Authentic Assessments in Different Contexts

Darling-Hammond and Snyder (2000) discussed the use of authentic assessments in preservice teacher education programs that were designed to help novice teachers move from an intellectual understanding of issues to the enactment of practice. The authors argue that without an appreciation of the intense and dynamic realities of classroom life and for the multidimensional problems and possibilities posed by different students, it is difficult for those with sound theoretical knowledge to apply what they know in practice. Moreover, authentic assessments not only better reflect the complexity of teaching and provide valid data about competencies, they also help teachers improve the caliber of their work.

Based upon emerging research and their own experiences, Darling-Hammond and Snyder (2000) identified four aspects that make up an authentic assessment of teaching. These include assessments that

- sample the actual knowledge, skills, and dispositions desired of teachers as they are used in teaching and learning contexts
- integrate multiple types of knowledge and skills as they are used in practice
- use multiple sources of evidence collected from different situations and over time
- are evaluated by individuals with recognized expertise in the field against accepted performance criteria

The types of authentic learning that capture the aspects identified above are case studies, portfolios or examples of practice, exhibitions of performance, and problem-based inquiries.

Case Studies

While Darling-Hammond and Snyder (2000) argue that no single type of authentic assessment represents the totality of teaching, each does assess important aspects of teaching that enable novice teachers to integrate different areas of learning and to apply them in different ways. For example, *cases* or *case studies* allow for the investigation of particular issues through theories, principles, and practices as they occur in the real world. A case study of a child with behavioral difficulties may involve an exploration of theories of child development and learning, an investigation of classroom context and curriculum, and an explication of management strategies employed by the class teacher. Opportunities to reflect on the information obtained in the light of feedback from mentors and peers often lead to a greater understanding of how different knowledge and theories can be better integrated and applied in classroom situations with students who manifest behavioral difficulties. In other words, novice teachers often learn to arrive at more meaningful understandings of what factors may be contributing to the difficulties students are experiencing when they have had the opportunity to undertake real-life investigations that involve the analysis of multiple sources of information and their application to the problem at hand and to receive feedback from others with expertise in the field.

Although the research above refers to case studies in the context of helping teachers to learn more about the students they teach, case studies can also be used in classrooms to help students gain a better understanding of the real-life experiences of individuals who have made various contributions to humanity and the sciences. For example, students may develop a greater appreciation of the political contributions that Abraham Lincoln made through a case study of

his life. This type of case study would provide students with the opportunity to study features of the economy during Lincoln's presidency, the historical significance of different events, the sociodemographic makeup of the community, and so on. Such a case study would provide students with opportunities to develop a rich and in-depth knowledge and understanding of this president, an appreciation that might not occur if they were left to read about him in a text.

Portfolios

A second type of authentic learning proposed by Darling-Hammond and Snyder (2000) involves the use of *portfolios*. Portfolios are collections of exemplars of teachers' practices from multiple sources. They may include various curriculum documents such as lesson plans; units of work; and samples of students' work such as DVDs, CDs, and other multimedia presentations as well as samples of their written work. They may include detailed descriptions of specific lessons; analyses of students' work; and teachers' timetables, logs, or journals as well as feedback obtained from their own teaching practice or reflections on it. In addition, various class and school policy documents and procedures for working with students with diverse learning and adjustment needs are often included. In this sense, portfolios allow novice teachers to collect and assemble a rich variety of documents relevant to a range of teaching practices that can help to inform their own developing pedagogical skills. Novice teachers are often required to submit teaching portfolios to demonstrate their learning and involvement with specific cases or classroom experiences. Moreover, as part of the assessment process, they are often required to write reflections on these experiences: what they have learned and the insights they have gained both from the practical experience they have had and the background research they have undertaken. Opportunities to share these reflections with supervising teachers or peers often lead to additional insights and learning as novice teachers learn to justify decisions, actions, and interpretations in the light of the information they have presented and the feedback they have received from experienced practitioners in real-life contexts.

In classroom situations, teachers will often use portfolios to collect examples of students' work over a period of time. Portfolios can be assembled by teachers on students, or students can construct their own portfolios. Either way, they are designed to be an ongoing record of work students have attempted or completed and as such can provide insights into progress. They can be used to facilitate communication with parents, students, and other teachers so that successes can be celebrated and difficulties can be identified

and strategies developed to help overcome them (Gillespie, Ford, Gillespie, & Leavell, 1996).

Micklo (1997) used mathematics portfolios with elementary students and found that when students understood the purpose of a portfolio, it became a tool to help students learn to monitor their own learning by setting goals for themselves and checking their progress toward reaching them. Moreover, Micklo argued such portfolios provide insights into students' abilities to communicate mathematically, to solve problems, to demonstrate mathematical reasoning, and to make connections between different concepts and relationships in mathematics. This enables teachers to make more appropriate decisions about an instructional program based upon students' knowledge and capabilities.

Exhibitions of Performance

A third type of authentic assessment involves exhibitions of performance (Darling-Hammond & Snyder, 2000). In preservice teacher training programs, exhibitions of performance allow teachers to demonstrate particular competencies in ways that simulate teaching contexts or events. These exhibitions can draw upon video observations, case study analyses, observations of simulated events, lesson plans, or group activities to demonstrate what teachers do when solving particular problems. The advantage of exhibitions is that performances can be evaluated in relation to prescribed standards of practice that have been previously discussed. Moreover, when novice teachers have to demonstrate competencies through performance, it sharpens their awareness of the knowledge and skills they must integrate in order to meet required standards. Similarly, students who are required to give a performance in a specific field (i.e., music, drama) must also learn to integrate their knowledge and skills if they are to demonstrate their competencies.

Problem-Based Inquiries

The final type of authentic assessment referred to by Darling-Hammond and Snyder (2000) involves the establishment of *problem-based inquiries*. With problem-based inquiries, teachers design and conduct investigations into issues arising from their work. For example, an investigation into the apparent high rates of truancy in a high school may lead investigators to form hypotheses about the causes of the absenteeism as well as possible ways of resolving it. This may require them to consider relevant research on engagement–disengagement and motivation

to learn in adolescents as well as an examination of current data on rates of absenteeism in the school and compare them with previous years' records or absenteeism rates in similar neighborhood schools. Furthermore, as part of the inquiry, the investigators may need to interview students, parents, and school personnel to ascertain their perceptions of the problem and potential solutions. In this sense, the inquiry has to be very detailed and comprehensive and it must include all stakeholders as well as all available information that can be used to inform the investigation. Investigators have a responsibility to be able to analyze and synthesize the available data and provide evidence for the conclusions they have drawn so they can propose suggestions that may help to solve the problem. The sharing and critiquing of the research that follows can help to transform teaching as investigators and teachers collaborate to work on the problem. Darling-Hammond and Snyder maintain that having novice teachers engage in this type of inquiry can help prepare them both as consumers of research and producers of knowledge. Moreover, it can give them the tools to make sense of their practice and help them to think systematically and analytically about the problems they confront.

Problem-Based Learning Using Formative and Summative Assessments

Zimitat and Miflin (2003) report on the development and application of a four-step assessment task (4SAT) that was used to gather both formative and summative assessment information on medical students engaged in problem-based learning. The impetus for the instrument was to align assessment with the philosophy and process of the problem-based learning (PBL) curriculum that aims to help students better explore clinical cases, engage in more effective hypothetico-deductive reasoning processes about the presenting problem, share their findings with their peers, and probe and challenge each other's understanding as they further refine their hypotheses and gather additional information.

Students worked initially on the problem case by themselves, identifying key features and formulating potential hypotheses (Step 1), before working with their group members to share their analyses and hypotheses. With the presentation of new clinical data by the tutor, students were forced to reconsider their hypotheses and discuss possible scenarios before submitting a written summary of their deliberations and their learning objectives and tasks for the week. A list of the top 10 learning issues identified by the different groups was then collated and e-mailed through to students, which Zimitat and Miflin believe made the learning genuinely reflective of students' learning needs (Step 2). It was during this step that information was collected on both the learning product (written

report and learning issues) the students had generated from their discussions and the group process.

The group's written response was assessed according to the extent they were able to identify key presenting symptoms, formulate hypotheses, utilize prior knowledge, justify the need for further information, reconsider hypotheses in the light of new information, and develop learning issues related to the case. The group process was assessed according to whether group members participated in the discussion, maintained a supportive and nonjudgmental learning environment, considered the ideas of others, recognized and resolved conflicts and disagreements, negotiated members' tasks, and reviewed the group's progress with the case.

During Step 3, students were encouraged to work collaboratively together on their learning tasks and independent study and to share resources. Step 4 consisted of a written examination based on the top 10 learning issues identified by the groups in Step 2. Case scenarios were constructed (similar to Step 1), and students were expected to generate hypotheses, interpret data, and answer a variety of questions based on these cases.

Zimitat and Miflin (2003) propose that the advantage of the 4SAT was that it mirrored the PBL process; students were therefore tested on the learning issues that the groups had discussed and examined. As such, the 4SAT was an integrated assessment tool that provided both formative and summative information on the learning that had occurred and the group processes involved. Moreover, it had high face validity (it appeared to measures what it is supposed to measure, as it was closely linked to the curriculum), and an analysis of its psychometric properties indicated that it is moderately correlated to other summative assessment instruments used to test students' basic and clinical science skills in medicine. Furthermore, as the 4SAT was criterion referenced, students had clear guidelines on what they had to do to achieve specific goals in the PBL curriculum.

Classroom teachers can use problem-based inquiries not only to extend students' learning and foster their interest in a topic, but also to gauge the extent to which students are constructing understandings in their groups. The advantage of this approach to assessment is that it is closely linked to problem issues that students have been investigating, so the assessment is a valid reflection of what the children have learned.

Key Points on Summative Assessments and Their Purposes

- Summative assessments are designed to measure outcomes of learning.
- Most summative assessments for group learning are criterion referenced.

- Criterion-referenced assessments specify clear criteria for achieving the goal.
- Teachers often prefer to use criterion-referenced assessments to communicate with teachers and students and to set learning goals.
- Summative assessments can involve authentic assessments.
- Authentic assessments are criterion-referenced assessments.
- Authentic assessments are designed to measure progress.
- Authentic assessments involve real-life tasks.
- Authentic tasks can include cases, portfolios, performances, and problem-based inquiries.
- Authentic assessments can be used to measure progress in different learning contexts.

Practical Activity

Conducting Summative Assessments of Small-Group Learning

Elementary School

★ Criteria to assess group's PowerPoint presentation on the drug nicotine

Criteria	*Still Developing*	*Developed*	*Highly Developed*
PowerPoint presentation:	Sometimes uses the following design features: • Eye catching • Same size font except when in headings • Appropriate picture • Appropriate use of sound effects • Effective use of colors • Includes transition • Similar background	Usually uses the following design features: • Eye catching • Same size font except when in headings • Appropriate picture • Appropriate use of sound effects • Effective use of colors • Includes transition • Similar background	Consistently uses the following design features: • Eye catching • Same size font except when in headings • Appropriate picture • Appropriate use of sound effects • Effective use of colors • Includes transition • Similar background

Criteria	*Still Developing*	*Developed*	*Highly Developed*
Describes nicotine	Unable to write one or two sentences describing nicotine	Writes a paragraph describing nicotine	Writes a paragraph or more describing and comparing it to alcohol
Analyzes how nicotine affects the body	Unable to write one or two sentences on how nicotine affects the body	Writes a paragraph on how nicotine affects the body	Writes a paragraph or more on how nicotine affects the body and compares it to alcohol
Identifies what the body does to process nicotine	Does not identify what the body does to process nicotine	Writes a paragraph and identifies what the body does to process the nicotine	Writes a paragraph or two and identifies what the body does to process nicotine and compares it to alcohol
Identifies other groups of people affected by nicotine exposure and analyzes the consequences	Does not identify other groups of people affected by nicotine exposure and does not analyze the consequences for these groups	Writes a paragraph to identify other groups of people affected by nicotine exposure and analyzes the consequences for these groups	Writes a paragraph or two to identify other groups of people affected by nicotine exposure and analyzes the consequences for these groups
Explains why nicotine is a drug	Does not explain why nicotine is a drug	Writes a paragraph to explain why nicotine is a drug	Writes a paragraph or two to explain why nicotine is a drug
Identifies some interesting facts about nicotine use	Does not identify interesting facts about nicotine use	Writes a paragraph on interesting facts about nicotine use	Writes a paragraph or two on interesting facts about nicotine use
Language:	• Spells most familiar words accurately as well as some new/topic-specific words	• Spells most familiar words and new/topic-specific words accurately	• Spelling is nearly error free

(Continued)

(Continued)

Criteria	*Still Developing*	*Developed*	*Highly Developed*
	• Sentence structure: mostly simple sentences • Paragraphing: attempts to group sentences containing related information into paragraphs • Punctuation contains some common punctuation	• Sentence structure: mostly simple sentences with some complex sentences • Paragraphing: generally groups sentences containing related information into paragraphs • Punctuation: consistent use of most common punctuation	• Sentence structure: demonstrates control of simple, complex, and compound sentences • Paragraphing: consistently groups sentences containing related information into paragraphs • Punctuation: error free and appropriate punctuation
Student comments:			
Teacher comments:			
Level of support: Check appropriate comment	Lots of teachers help	Some teacher help	Very little teacher help

★ Group has been asked to discuss a problem (e.g., getting rid of plastic bags because of the difficulties with disposal) and write a report on their proposed solution(s). The following questions have been given to the group to help stimulate their discussion on the topic:

What do I know about the problem?

Why is it worth solving?

How will solving it help us?

What are the minuses of solving this problem?

How do I feel about this problem?

What are some ways of solving this problem?

How could the solution be put into practice?

The criteria to evaluate the information presented in the report are listed below:

Criteria	*Example of Criteria*	*Stimulus Questions*	*Score*
1. Knowledge review	Able to recall three basic facts relevant to the topic	What do I know about the problem?	
2. Comprehension	Able to understand the need to solve/deal with the problem using two cause-effect relationships or a plus and minus of the problem	Why is this worth solving? How does it help us? What are the minuses of solving this problem? What are some ways of solving this problem?	
3. Connecting	Able to demonstrate at least one more complex understanding of cause and effect or make links between thoughts and feelings or connect ideas	How do I feel about this problem?	
4. Application-extending knowledge	Able to demonstrate how to apply known information to possible practice	How could the solution be put into practice?	
5. Thinking about thinking	Able to generate additional information relevant to the problem based on connecting ideas and constructing new knowledge or idea or generating question to seek more information	Are there any other questions I need to ask to help with this problem?	

Scoring: 1 = acceptable response; 0 = unacceptable response.

(Continued)

(Continued)

★ Students have been asked to think about a problem and depict their thinking in the following way:

<table>
<tr><td colspan="6">Problem: What are some ways of solving the water shortage problem in city X?</td></tr>
<tr><td colspan="2">Suggested solution 1:</td><td colspan="2">Suggested solution 2:</td><td colspan="2">Suggested solution 3:</td></tr>
<tr><td colspan="2">List Consequences + and –</td><td colspan="2">List Consequences + and –</td><td colspan="2">List Consequences + and –</td></tr>
<tr><td>+</td><td>–</td><td>+</td><td>–</td><td>+</td><td>–</td></tr>
<tr><td colspan="6">Best solution:</td></tr>
<tr><td colspan="6">Reason it's the best solution:</td></tr>
<tr><td colspan="6">Create a slogan, design a logo, or design a poster for your group that communicates the main message arising from your best solution.</td></tr>
</table>

The criteria for evaluating the group's response to the problem are presented below:

Level	*Criteria*	*Score*
1. Suggested solutions	Able to identify two possible solutions to the problem	
2. Consequences	Able to identify at least one consequence for each solution	
3. Best solution and reason	Able to identify the best solution and provide at least one reason. The reason needs to be well argued so it shows cause and effect for at least two different groups (people, animals, industry).	
4. Thinking about thinking	Able to communicate the main message through a logo or poster. Message needs to be well reasoned and not basic recall of detail; successfully combines various thinking abilities.	

Scoring: 1 = acceptable response; 0 = unacceptable response.

★ The group has been asked to develop a television advertisement to promote a product of their own choosing. Below are the criteria used to assess the television advertisement:

Criteria	*Self-Assessment*					*Teacher Assessment*				
Slogan, jingle	1	2	3	4	5	1	2	3	4	5
Logo										
Packaging										
Pricing										
Persuasive techniques used										
Target audience, appropriate use of language										
Critique										
Storyline sequencing										
Performance										
Overall success of the advertisement										
Scoring: 1: Low; 5: High Group members: Comments:										

Middle School and High School

★ Ways of assessing content-related criteria for a project. The following rubric is an adaptation of one used by Prins et al. (2005) to assess the learning that occurred from a group project that was the outcome of an online virtual seminar series.

★ The group's task is to investigate an alternative fuel source and write a report providing information on their chosen alternative fuel and justifying their choice. As part of the report, students must explain how their energy source works and its advantages and disadvantages, as well as the implications of this alternative energy source being the sole provider of energy in today's society. Possible choices of energy include: hydroelectric, nuclear, biomass, tidal, solar, wind, ocean thermal, wave, geothermal.

(Continued)

(Continued)

Content-Related Criteria	*Above Average*	*Average*	*Below Average*
Topic under discussion is made operational	Students give a definition used in their report and suggest practical tools to measure their solutions.	Students give a definition but do not suggest practical tools to measure their solutions.	Students do not give a definition and do not suggest practical tools to measure their solutions.
The different aspects of the topic are used in coherence	The different aspects of the topic are balanced, and coherent arguments are given priority.	Not all aspects of the topic are used, but when they are, they are balanced and coherent.	The different aspects of the topic are not balanced or coherent.
Consistency in content and awareness of gaps in knowledge	The content of the report is cogent and consistent. Insight is given into gaps in knowledge.	The content of the report is cogent and consisten. Gaps in knowledge are not recognized.	The content of the report is not cogent or consistent. Gaps in knowledge are not recognized.
Integration of disciplinary contributions (sources of knowledge)	The different disciplines are integrated in each chapter of the report, not just at the end.	The different disciplines are integrated only at the end of the report.	The different disciplines are not integrated.
Relationship between problem definition, analysis, and solution	The report is scientifically written. Questions are answered, and recommendations are logical.	The report is scientifically written. Not all questions are answered.	The report is not well written scientifically. Not all questions are answered. Recommendations do not flow logically.
Creativity	Different understandings and knowledge are linked to each other in a creative way. Recommendations are provocative and sharp.	Different understandings and knowledge are linked to each other in a creative way.	No new insights are provided because different knowledge and understandings are not linked.
Summary	A two- or three-page summary is added containing: background research, target group, recommendations, and possible ways to implement solution. The summary is sharp and provocative.	Summary is lacking one of the following: background research, target group, recommendations, and possible ways to implement solution. The summary leaves room for interpretation.	Summary is lacking two or more of the following: background research, target group, recommendations, and possible ways to implement solution. The summary leaves room for interpretation or no summary is added.

Criteria	*Highly Developed*	*Well Developed*	*Satisfactory*
Knowledge and understanding	Comprehensively explains concepts and interrelationships	Explains concepts and interrelationships	Adequately explains concepts and interrelationships
	Uses a wide variety of sources to obtain information	Uses a substantial variety of sources to obtain information	Some evidence of variety of sources to obtain information
	Extensive use of key terms and concepts	Good use of key terms and concepts	Adequate use of key terms and concepts
	Comprehensive use of scientific concepts to explain processes and changes that occur	Good use of scientific concepts to explain processes and changes that occur	Adequate use of scientific concepts to explain processes and changes that occur
Investigating	Collects wide variety of relevant and reliable quantitative and qualitative data	Collects a variety of relevant and reliable quantitative and qualitative data	Collects a range of quantitative and qualitative data
	Comprehensively identifies and analyzes patterns of similarities and differences	Identifies and analyzes patterns of similarities and differences	Adequately identifies and analyzes patterns of similarities and differences
Communicating	Excellent use of written genre to explain terms and meaning	Good use of written genre to clarify terms and meaning	Adequate use of written genre to clarify terms and meaning

BRINGING IT ALL TOGETHER: UNDERSTANDING THE RESEARCH

There is no doubt that there has been increasing pressure on schools as a consequence of the high-stakes testing regime for teachers to "teach to the test" or to spend inordinate amounts of time and energy preparing students for the *standardized testing* that occurs (Posner, 2004). Given that the funding for schools and school districts depends on the numbers of students who can meet required proficiency standards overall and within major subgroups within a school (i.e., economically disadvantaged, major racial and ethnic minorities, students with disabilities, and students with limited English proficiency), it is not surprising that a great deal of attention is devoted to ensuring students meet the required standards (Kim & Sunderman, 2005).

Moreover, given that schools that fail to demonstrate adequate yearly progress (AYP) for 2 consecutive years are identified as "in need of improvement" (Kim & Sunderman, 2005, p. 3) and subject to a series of sanctions that increase in intensity the longer a school remains delinquent, teachers are undoubtedly feeling strongly pressured as they teach. When this happens, some teachers question the value of other types of nonstandardized tests, such as criterion-referenced formative and summative assessments, that often take time for the teacher to develop and generally do not provide students with experience in responding to the types of multiple-choice and short answer questions that generally appear on standardized tests. In contrast, other teachers often express concern over the relevance of standardized testing when it ignores the process of teaching and learning that occurs in classrooms, arguing that it does not provide them with the guidance they need to be able to improve how they teach (Stiggins, 2002).

In an effort to investigate both these concerns, Black and Wiliam (1998a) conducted a synthesis of the results of evidence on formative assessments published in more than 250 articles from a number of different countries and concluded unequivocally that formative assessments do raise students' achievements and that these findings hold for students from kindergarten to university undergraduates and across different subject areas. Furthermore, formative assessment was more beneficial for low achievers because it helped to close the achievement gap while raising achievement for all students overall. Black and Wiliam (1998b) attributed this improvement to the frequent use of feedback, which helped students to understand what they needed to do to achieve or to be successful. When students are involved in the assessment process, they not only learn what they must do but also how to monitor their own learning more closely, and this enhances their cognitive and metacognitive thinking. Furthermore, when formative assessments are conducted in a culture of success and the belief that all students can succeed, students are more likely to persevere with their tasks than they would otherwise.

The success, however, of formative assessments is also dependent on the type of task students are given, the emphasis given to the procedural aspects of knowledge acquisition, the quality of the discourse between teachers and students and students with each other, and the ongoing adjustments that are made to teaching to accommodate changes in students' learning (Black & Wiliam, 1998a).

There is no doubt that the task students are given on which the formative assessment is based needs to be novel and varied, intellectually

challenging, and meaningful if students are going to be motivated by it. It also needs to provide students with opportunities to focus on how to solve the problem rather than just declare what they know about it, and this usually occurs when students are provided with opportunities to discuss their ideas with their peers and receive feedback from them. Students are remarkably honest about the feedback they provide about both themselves and each other. When they have been taught how to provide feedback in a socially acceptable manner, many students are receptive to the opinions of their peers as these are an important complement to self-feedback and they will learn from it (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998a).

Teacher discourse is also very important, particularly when it is used to challenge students' thinking and scaffold their learning. Probing students to think about issues and encouraging them to consider alternative positions is likely to promote more self-reflection and higher-level thinking than engaging them in *initiation-response-feedback (IRF) exchanges* where teachers typically ask closed questions to which students often already know the answer (Rojas-Drummond & Mercer, 2003). Moreover, when teachers engage in discourse that promotes thinking and scaffolds learning, students' own thinking becomes influenced by these communication exchanges and they begin to model them in their interactions with each other. It seems that the responses expected in this type of discourse appear to create an expectation in recipients to reconcile their understandings with those of others, clarify misunderstandings, and provide justifications that others will accept as well reasoned and valid. Furthermore, this situation, coupled with the security of the small group, appears to provide a psychological environment that motivates students to be more willing to reconcile contradictions between themselves and others, try out their ideas, and work to construct new understandings (Gillies, 2004b).

Other ways in which students can learn from formative assessments include having students develop a scoring rubric that focuses their attention on the criteria required to produce a quality product (Black et al., 2004). This helps to make explicit the expectations for meeting the criteria or achieving a quality outcome. Moreover, engaging students in peer-assessment and self-assessment encourages students to be active in their learning, prepared to acknowledge their mistakes, and responsible for monitoring when they are learning and when they are not.

Feedback is very important for students' learning. In a comprehensive review of studies on feedback to students, Kluger and DeNisi (1996) found

that feedback improved performance in the majority of these studies. Feedback that focuses on what needs to be done is more helpful than the awarding of a mark or a grade. Furthermore, students are often prepared to accept the feedback they receive when they understand its purpose and how it can be used to help them overcome difficulties they may be experiencing. Feedback that focuses on what needs to be done can encourage students to believe they can improve. Such feedback can enhance learning as students are motivated to invest in the effort required (Black et al., 2004).

While there is strong evidence that the use of formative assessments or assessment for learning leads to higher quality learning and enhanced learning outcomes, the effects of this type of assessment on summative assessment outcomes is less clear because of the limited research on this topic. Wiliam, Lee, Harrison, and Black (2004), in a study of the use of formative assessments with high school students, found that when teachers made time to develop and use formative assessments there was strong evidence of tangible benefits in terms of students' achievements on mandated standardized tests. Similarly, Engel, Pulley, and Rybinski (2003), in a study of the use of authentic assessment activities in the curriculum for students in Grades 2 and 4, found that these activities helped students to do well on standardized tests. The authors argued that performance tasks involving hands-on activities that were completed before the standardized tests enabled students to apply those skills to the standardized assessment. Moreover, Black et al. proposed that it is possible to improve classroom teaching practice and hence students' learning when summative assessments are used for formative purposes. This happens when students

- Engage in reflective reviews of the work they have completed to allow them to study more effectively
- Set questions and practice responses to gain a deeper understanding of the assessment process
- Use peer- and self-assessment to help them better understand what they need to do

Assessment is so important that Stiggins (2005) argues that students need to be inside the assessment process so they understand what they have to achieve and the scaffolding that will support them as they learn. When this happens, students can see themselves grow as they learn to monitor their progress, feel in control of their learning, and believe that success is possible if they keep trying. Moreover, Stiggins maintains when these practices are consistently carried out in classrooms, they have been consistently linked to gains on high-stakes testing, with the largest gains made by low achievers.

CHAPTER SUMMARY

The research on assessing small-group learning suggests

- Formative and summative assessments can be used to assess student learning in small groups.
- Students learn more when these assessments are criterion referenced.
- Criterion-referenced assessments are designed to help students understand what they must do to achieve a specific outcome.
- Formative assessments do raise standards of achievement overall.
- Formative assessments are more beneficial for low achievers.
- Changes in teaching practice need to occur as a consequence of feedback students receive from formative assessments.
- Formative assessments lead to tangible benefits on summative assessments such as standardized achievement tests.
- Students benefit when summative assessments are used for formative purposes or to help students understand what they need to know and how they can learn.
- Students benefit from both formative and summative assessments when teachers use the information to make changes in their teaching.

ACTIVITIES

1. Discuss with students your expectations for the completion of a small-group activity/project they are about to undertake. Work with the students to develop a rubric that reflects both the group process and the performance criteria required. Negotiate with the students levels of demonstrated performance, such as: beginning, developing, accomplished, and exemplary. Spend time at the completion of the group activity/ project to discuss how the students perceived they met the various criteria for both the process of learning and the performance achieved.

2. Help students to construct a journal of their learning experiences. Provide guidelines for what they might like to include in their journal. For example, information on what they have learned from a specific experience, what they may still need to do to accomplish a special goal, and so on. Make time to have an individual conference with each student to discuss his or her perceptions of his or her progress and what goals the student might like to develop as a consequence.

3. Discuss with the class how they can provide feedback to a group of students at the completion of their presentation. Suggestions may include having members focus on discussing clarity of the presentation, links established between points/issues raised, quality of the background research, presentation of thesis/key points, and so on. Clear performance criteria need to be established so all groups understand the assessment process. Feedback can be used for both formative and summative purposes. It may be advisable to help students construct a rubric of the criteria they will use to assess the group's performance so all students are aware of the criteria that are going to be applied to their presentation. The advantage of using a formalized rubric is that it focuses on the group's presentation and not the individuals within the group.
4. Interview two teachers about how they assess cooperative group work. Focus on both formative and summative assessments and try to ascertain how these assessments are conducted. For example, are these assessments embedded in a group activity or task? What types of tasks do they use? How do they assess the process and outcome of these group tasks? How do they provide feedback to the group? Does anyone else provide feedback? Check to see if the teachers are satisfied with this approach to assessment. If they are, investigate the strengths. If not, ask them what they would like to do differently next time they give feedback to a group. These are just some of the questions that can be asked. Collate your information and discuss the similarities and differences with your study group. What have you learned about teachers' attitudes to assessment as a consequence? Reflect on how you believe the information you have collected and discussed will inform your own approach to assessment.

SUGGESTIONS FOR FURTHER READING

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 8.

Stiggins, R. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan, 83*(10), 758.

Stiggins, R. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan, 87*(4), 324–330.