# InSTEP

**Integrating Strategies and Technology in Education Practice**

## Conducting Case Studies:
## Guidelines for Classroom Observations

**Protocol B**

February 2004

These scales were based on several sources including: (a) the scales for authentic instruction and the scoring of student work developed by the Center for Organization and Restructuring of Schools at the University of Wisconsin-Madison, (b) scales developed/adapted by Brendefur & Hernandez (1996), and (c) scales developed/adapted by the "Science for All" project at the University of Miami (Luykx & Lee, 2003).

**InSTEP Case Studies**

**Guidelines for Classroom Observations**

## Areas of Inquiry

Four areas of inquiry were identified based on a review of the InSTEP model and constructivist ideas for teaching, learning, and assessment: (a) teaching strategies (inquiry/problem-based), (b) use of technology, (c) classroom environment (student-centered), and (d) assessment strategies.

## Indicators

To keep the inquiry focused and in-depth, up to three indicators are used to guide the observations for each area of interest.

1. Inquiry/Problem-Based Teaching Strategies

    - Disciplinary understanding
    - Inquiry focus
    - Value beyond the class

2. Use of Technology

3. Classroom Environment
    - Locus of authority
    - Substantive conversation
    - Student engagement

4. Assessment Strategies

## Scoring

To generate numeric data, each of the indicators is scored based on 5-point Likert scale.

## Evidence

Observation notes are required to substantiate scores and provide descriptive evidence of numerical indicators.

## General Guidelines

### Spirit of the Scales

When applying these scales, you should think about the spirit in which they were developed. For most scales, the numbers go up based on two things: (a) the intensity or frequency with which something is taking place and (b) the number of students who are engaged in doing that thing. As a rule of thumb—and only that—it may help to think of the numbers as the following:

1. The stereotype of undesirable environment;
2. Minimal intensity; could be limited to the teacher or to a few students;
3. Greater and/or uneven intensity; includes some students;
4. Substantial and intense; includes many to most students;
5. Very intense; includes most to almost all students

### Meanings of Terms

| *Term* | *Description* |
|---|---|
| **Almost all** of the time, of the students | 90% or more of the time, of the students |
| **Most** of the time, of the students | From 50% and less than 90% of the time, of the students |
| **Many** times, students and/or much (of the time) | More then 20% and less than 50% of the time, of the students |
| **Some** of the time, of the students | More than 10% to 20% of the time, of the students |
| **A few** times, of the students | 10% or less of the time, of the students |

### Observation Notes

- Observation notes will be treated as the primary source of evidence justifying how the scales were scored. Note taking on paper or laptop must be comprehensive to support ratings and descriptive to allow others understand what went on in the classroom.
- Observe an entire class and, at the start of the observation, note the time, number of students, and general demographics of the class, lesson goals, and any other relevant information.
- Keep typing/writing as much as you can but don't make the teacher your only focus of attention. Student behaviors and interactions are all relevant to their engagement in the lesson and the general classroom atmosphere. Use names whenever possible even if you don't know enough about the individual students and the relevancy is not apparent at the moment. For example, this information may helpful later when you are trying to determine the extent of student interactions with the teacher.

## Summary of Classroom Observations

During classroom observations, the observer takes field notes about classroom activities, academic tasks, and teacher and student interactions. Ratings for the scales can be completed at the end of the classroom observation. Based on field notes, ratings can be modified at the end of the day when reviewing observations notes.

Date: _____ Case study: o Internal o External

School: _____ Grade: _____

Teacher: _____ Subject: _____

Observers:_____ Rater: _____

**Summary of Scores**

| | | | | | | |
|---|---|---|---|---|---|---|
| **Inquiry/Problem-Based Teaching Strategies** | | | | | | |
| 1. Disciplinary understanding | | 1 | 2 | 3 | 4 | 5 |
| 2. Inquiry focus | | 1 | 2 | 3 | 4 | 5 |
| 3. Value beyond the class | | 1 | 2 | 3 | 4 | 5 |
| **Use of Technology** | | | | | | |
| 4. Use of technology | NA | 1 | 2 | 3 | 4 | 5 |
| **Classroom Environment** | | | | | | |
| 5. Locus of authority | | 1 | 2 | 3 | 4 | 5 |
| 6. Substantive conversation | | 1 | 2 | 3 | 4 | 5 |
| 7. Student engagement | | 1 | 2 | 3 | 4 | 5 |
| **Assessment Strategies** | | | | | | |
| 8. Assessment strategies | | 1 | 2 | 3 | 4 | 5 |

## Inquiry/Problem-Based Teaching Practices: Disciplinary Understanding

TO WHAT EXTENT DO STUDENTS DEMONSTRATE A DEEP UNDERSTANDING OF THE SUBJECT MATTER?

For students, **knowledge is deep** when they develop relatively complex understandings of the lesson's concepts. They also may produce new knowledge when they connect disciplinary concepts or topics to one another. Further, they apply concepts to explain natural phenomena or real world situations. Instead of reciting only fragmented pieces of information, students develop relatively systematic, integrated, or holistic conceptual understandings. **Knowledge is shallow**, thin, or superficial when concepts are taught in isolation from related ideas, personal experiences, or real world phenomena, providing students with only a surface acquaintance with their meaning. This superficiality can be due, in part, to instructional strategies, such as when teachers cover a large quantities of fragmented ideas and bits of information that are unconnected to other knowledge. Evidence of shallow understanding by students exists when they do not or cannot use knowledge to make clear distinctions, build arguments, solve problems, or develop more complex understandings of other related phenomena. **In scoring this item**, you should note that depth of knowledge and understanding refers to the substantive character of the ideas that students express as they consider topics of interest. It is possible to have a lesson containing substantively important and deep knowledge, but students fail to show understanding of the complexity or the significance of the ideas. Observers' ratings should reflect the depth to which *students* pursue the content.

### Scale

1. Knowledge is superficial. Concepts are taught in isolation from related ideas, personal experiences, or real world phenomena. Students are mainly required to memorize information.

2. Knowledge remains superficial. Underlying or related concepts and ideas might be mentioned or covered, but only a superficial understanding of these ideas is evident.

3. Knowledge is treated unevenly during instruction; there is deep understanding of some scientific concepts and ideas, but superficial understanding of some other ideas. At least one idea is presented in depth and its significance may be grasped by some students (10%-20%), but in general the focus is not sustained.

4. Knowledge is relatively deep. Students provide information, arguments, or reasoning that demonstrate the complexity of one or more ideas. The teacher structures the lesson so that many students (20%-50%) do at least one of the following: sustain a focus on a significant topic for a period of time; demonstrate understanding of the connections between concepts, and between these and personal experiences or real world phenomena; demonstrate understanding of the problematic and incomplete nature of information; or demonstrate understanding by making reasoned and well-supported arguments.

5. Knowledge is consistently deep. The teacher successfully structures the lesson so that most students (50%-90%) do at least one of the following: sustain a focus on a significant topic for a period of time; demonstrate understanding of the connections between concepts, and between these and personal experiences or real world phenomena; demonstrate understanding of the problematic and incomplete nature of information; or demonstrate understanding by making reasoned and well-supported arguments.

## Inquiry/Problem-Based Teaching Practices: Inquiry Focus

TO WHAT EXTENT DO STUDENTS USE INQUIRY PROCESSES?

The scale is intended to measure the extent to which students engage in inquiry processes. There are two dimensions to this construct. First, **inquiry occurs when students conduct an investigation or an experiment**. Inquiry involves generating questions, designing investigations and planning procedures, carrying out the investigations, analyzing and drawing conclusions, and reporting findings. Inquiry is not a linear process; instead, aspects of inquiry interact in complex ways. For example: (a) asking a question about objects, organisms, and events in the environment; or asking a question that can be answered through a scientific investigation; (b) planning and conducting a simple scientific investigation; (c) using appropriate tools and techniques to gather, analyze, and interpret data; (d) using data to construct a reasonable explanation; or developing descriptions, explanations, predictions, and models using evidence; (e) communicating scientific procedures, investigations, and explanations, and (f) using mathematics in appropriate aspects of scientific inquiry. Second, **inquiry can be thought of as higher order thinking** that involves thinking that goes beyond recording or reporting facts, rules, and definitions or mechanically applying concepts. Inquiry involves searching for patterns, making hypotheses or inferences, and justifying those with evidence. Inquiry also includes organizing, analyzing, synthesizing, evaluating, predicting, arguing, making models or simulations, and inventing original procedures. A lesson can be low in scientific inquiry when students' activities are limited to repeating information provided by the teacher or text, or following a scripted set of procedures that does not require them to engage in higher order thinking. **Note:** Inquiry might take place almost accidentally or, seemingly, as an aside to the main flow of the lesson. For example, the teacher may ask a rhetorical question whose posing, if the question were taken seriously, would provide evidence of inquiry.

### Scale

1. Students receive, recite, or perform routine procedures. In no activities during the lesson do students engage in inquiry.

2. Students primarily receive, recite, or perform routine procedures. Students conduct a scripted investigation without higher order thinking. Or at some point during the lesson, students engage in higher order thinking as a minor diversion.

3. There is at least one significant activity involving scientific inquiry in which some students (10%-20%) demonstrate higher order thinking and/or conduct a non-scripted investigation. Or higher order thinking occurs sporadically.

4. There is at least one major activity in which many students (20%-50%) engage in higher order thinking and/or conduct a non-scripted investigation. This activity occupies a substantial portion of the lesson.

5. Most students (50%-90%), for most of the time (50%-90%), are engaged in scientific inquiry through an investigation and/or other activities involving higher order thinking.

## Inquiry/Problem-Based Teaching Practices: Value Beyond the Class

TO WHAT EXTENT IS THE ACTIVITY CONNECTED TO COMPETENCIES OR CONCERNS BEYOND THE CLASSROOM?

Whether a lesson has value and meaning beyond the instructional context is important because it is relevant to student understanding of the real world importance/applications of what they are studying. A lesson gains in value beyond the class the more it is connected to the larger social context within which students live. Students may solve relevant problems by applying knowledge to a variety of different situations and contexts. Two areas in which student work can exhibit some degree of connectedness are: (a) a real world public problem; e.g., students confront an actual contemporary issue or problem, such as applying statistical analysis in preparing a report to the city council on the structural concerns of near by bridges; (b) students' personal experiences, situations, or aspirations; e.g., the lesson focuses directly or builds on something students have actually gone through. **High scores can be achieved when the lesson** entails one or both of these. In a lesson with little or no value beyond the class, activities are deemed important for success only in school (now or later), but for no other aspects of life. Student work serves only to certify their level of competence or compliance with the norms and routines of formal schooling.

### Scale

1. Lesson topic and activities have no clear connection to anything beyond itself; the teacher offers no justification beyond the need to perform well in class.

2. Students work on a topic, problem or issue that the teacher tries to connect to students' experiences or to problem situations. For example, the teacher tells students there is value in the knowledge being studied because it relates to the world beyond the classroom. They are told that architects do scale drawings to justify the study of ratio and proportion.

3. Students study a topic, problem or issue that the teacher succeeds in connecting to students' actual experiences or to problem situations. Students recognize some connection between classroom knowledge and problems outside the classroom, but they do not explore the implications of the connections, which remain abstract or hypothetical. For example, to establish a connection between ratio and proportion and scale drawings, teachers might show architectural drawings and scale models to students.

4. Students study/work on a problem or issue that the teacher and students see as connected to their personal experiences or contemporary public situations. Students recognize the connection between classroom knowledge and the external problem situation. Students may apply and/or develop new knowledge within contexts that create personal meaning and significance for that knowledge. For example, in trying to understand ratio and proportion, students work with scale drawings and other architectural works.

5. Students study/work on a topic, problem or issue that is directly connected to their personal experiences or actual problem situations. Students recognize the connection between classroom knowledge and this particular situation. They explore that connection in ways that create personal meaning and significance for the knowledge addressed in this issue. Students create tangible, concrete evidence of their attempts to understand this topic or solve the problem. For example, students use ratio and proportion to design and create their own scale drawings and models for a development that is actually being proposed or debated in their town.

## Use of Technology

TO WHAT EXTENT IS TECHNOLOGY INTEGRATED INTO THE CONSTRUCTIVIST LEARNING ENVIRONMENT?

Instructional technologies run a wide gamut: from the process of instructional design to instructional tools that deliver, augment, supplement, or support teaching and learning. An instructional tool is anything that amplifies the capability of the teacher or learner (e.g., from pencil/chalkboard to computer software and hardware). This scale measures the integration of electronic/computer-mediated technology tools in learning activities. Teacher's technology literacy and comfort-level are key components to technology integration within the classroom along with a teacher's use of technologies for instructional preparation, delivery, and management. However, the goal of technology integration is to elicit student ownership of technologies as productivity, communications, research, and problem-solving and decision-making tools. **When fully integrated**, control and direction of technology usage has moved from teacher-directed to student-directed. When new technologies are incorporated within a classroom, they are often accompanied by an excitement due to novelty effects. Once technology tools are integrated within a classroom's practice, observed excitement, intensity, or engagement during tool usage is due to the enhanced capabilities afforded by the technology, as opposed to the technology itself. The degree of students' independent proficiency with the tool can help an observer classify the type of excitement. Technology integration is often accompanied by student-initiated collaborative communication and activities. When scoring this scale, mention (a) type of technologies (see list of tools, next page), and (b) how the tool is being used in instructional activities.

### Scale

NA Technology integration would be inappropriate within the context of the lesson.

1. There is no evidence of technology usage. Technology tools may be present in the room, but there is no evidence that teacher or students use them on a regular basis.

2. Minimal use of technology tools. The teacher may use one or more technology tools but use is primarily centered on classroom management, and instructional delivery and/or preparation. If any, student use is restricted to mechanical/rote activities (e.g., typing information).

3. Moderate use of technology tools. At least one portion of the lesson involves most students (50%-90%) in the use of one or more technology tools (individually or in groups) as a means to facilitate student understanding of concepts, communication, creative productivity, self-directed research, problem solving and/or decision-making (e.g., use of calculator). However, it is apparent that many (20%-50%) learners are unfamiliar with the tool and assistance in working with it. Student excitement may be due to the novelty of the tool.

4. High use of technology tools. Usage of one or more tools permeates most (50%-90%) of the lesson with some evidence of technology integration. Some usage may be student-directed (10% - 20%) and may involve tools in creative productivity, self-directed research, problem solving and/or decision-making. However, most (50% or more) usage still involves teacher-guided activities. When groups use tools, one or two members (20% to 50%) control the tools, while the rest support or record/watch the effort. Only a few (10% or less) who actually drive the tool request support in tool usage.

5. Integrated use of technology. Various technology tools appear to be used routinely and fully integrated within classroom practice. The role of technology tools is primarily centered on facilitating student understanding of concepts, communication, creative productivity, self-directed research, problem solving and/or decision-making. When used, most students (50%-90%) are engaged most of the time (50%-90%) in directing their own use of one or more technology tools. When used in groups, most (50%-90%) members of the group move easily in and out of the role of driver of the tool. Only a few (10% or less) who actually drive the tool request support in tool usage.

**List of Technologies Used in the InSTEP Program**
(Based on Summer 2003 Workshops)

- Microsoft Office
- Nikon Digital Cameras
- Garmin E-Trex Global Positioning Units
- ImagiProbes
- Palm Zire Handhelds
- Margi Presenter to Go
- Laptops (Gateway)
- Dreamweaver MX
- Inspiration
- Kidspiration
- Geometer's Sketchpad (were offered, but none chosen as tech tool)
- Blackboard
- Signals of Spring Curricular and Resource Materials (web-based, NASA funded)
- Getting InSTEP with Lewis and Clark: Exploring the Possibilities (web-based curriculum and resource)
- Palmscale, My Weigh, Digital Scales

**Other Electronic/Computer-Mediated Technologies Observed**

- _____
- _____
- _____
- _____

## Classroom Environment: Locus of Authority

WHAT IS THE EXTENT OF SHARED AUTHORITY IN THE CLASSROOM?

This scale highlights the extent to which a lesson supports a shared sense of authority and responsibility for validating students' reasoning. When students take on responsibility for justifying their own reasoning, they develop stronger understandings of the content and are more likely to make meaningful connections across disciplinary content and/or to the real world. **To score high on this scale**, the teacher and students hold each other accountable for convincing themselves and each other that their reasoning is sound and that their answers are correct. **Low scores are given either when** the authority for determining whether something is right or wrong rests with the teacher or the text, or (as occasionally happens) when neither the teacher nor students have a means for determining whether their reasoning is valid or not. This scale is not intended to measure students' control over the content of a lesson. The teacher still must decide what is worthwhile content and when a particular activity is not worth exploring in all of its details. In other words, the teacher makes curricular decisions; but those decisions should not undermine the sharing of authority within the class.

### Scale

1. For the most part, students accept an answer as correct only if the teacher says it is correct or if it is found in the book, and seldom challenge information from either of these sources. If stuck on a problem, students almost always ask the teacher for help. **OR**, there is no clear authority for determining whether someone's reasoning is valid. The teacher does not indicate whether students' answers are right or wrong, becomes flustered when queried about a topic, or is at a loss as to how to find out the answer, instead of suggesting possible resources to students.

2. Students rely on the teacher and some of their more capable peers as the legitimate source of authority. The teacher often relies on a few students to provide the right answer when pacing the lesson or to correct an erroneous answer. As a result, other students often rely on these students for correct solutions, verification of right answers, or help when stuck.

2. Many students (20% - 50%) share authority among themselves. They tend to rely on the soundness of their own arguments for verification of an answer. However, they still look to the teacher as the authority for making final decisions. The teacher sometimes asks students to provide their own arguments or hypotheses (e.g., by asking them, "What do you think?" or "How do you know?"), but intervenes with the answer in an effort to speed things up when students seem to be getting bogged down in the details of an argument.

4. Most students (50% - 90%) share in the authority of the class. Though the teacher might intervene when students are getting stuck, she usually does so with a question that focuses their attention or helps them to see a contradiction that they were missing. The teacher often answers a question with a question, though from time to time she provides the students with an answer.

5. Almost all the students (90% or more) share in the authority for the class. Students rely on the soundness of their own arguments and reasoning. As a rule, the teacher answers a question with a question or provides instrumental help (as opposed to just giving the answer) for students to make their own decisions. It is not uncommon to see students leaving a class still arguing about one or more points in their lesson.

## Classroom Environment: Substantive Conversation

TO WHAT EXTENT IS CLASSROOM DISCOURSE CHARACTERIZED BY SHARED UNDERSTANDINGS?

This scale assesses the extent to which talking is used to learn and understand important concepts in the classroom. In classes characterized by high levels of substantive conversation, there is considerable teacher-student and student-student interaction about the ideas of a topic; the interaction is reciprocal, and it promotes coherent shared understanding. (1) The talk is about subject matter and includes higher order thinking such as making distinctions, applying ideas, forming generalizations, raising questions; not just reporting of experiences, facts, definitions, or procedures. (2) The conversation involves sharing of ideas and is not completely scripted or controlled by one party (as in teacher-led recitation). Sharing is best illustrated when participants explain themselves or ask questions in complete sentences, and when they respond directly to comments of previous speakers. (3) The dialogue builds coherently on participants' ideas to promote collective understanding of the lesson's theme (which does not necessarily require an explicit summary statement). In classes where there is little or no substantive conversation, teacher-student interaction typically consists of a lecture with recitation where the teacher deviates very little from delivering a preplanned body of information and set of questions; students typically give very short answers. Because the teacher's questions are motivated principally by a preplanned checklist of questions, facts, and concepts, the discourse is frequently choppy, rather than coherent; there is often little or no follow-up of student responses. Such discourse is the oral equivalent of fill-in-the-blank or short-answer study questions. NOTE: the use of subject matter terminology does not guarantee the existence of discourse; indeed, the inappropriate use of terminology may actually interfere with the development of collective understandings and shared meanings. Terms, when used, should be meaningful and appropriate, and they should help support the conversation. In a whole class setting, students could participate in substantive conversation by listening and being attentive to the conversations that take place. Students do NOT have to all take turns talking on each and every point of a lesson; such turn taking may interfere with communication. Rather, students may selectively make comments when they have something to add. In small group settings, communication is likely to be more broadly spread throughout the group. In both cases, the issue is one of balance: no one person should dominate the conversation, but also, there does not need to be the ritualistic taking of turns where everyone speaks, even when there is nothing new to add to the conversation. The teacher and students behave as if there are agreed upon rules for taking turns and talking about the lesson's theme.

### Scale

1. Virtually no features of substantive conversation occur during the lesson.

2. Sharing and the development of collective understanding among a few students (or between a single student and the teacher) occur briefly.

3. There is at least one sustained episode of sharing and developing collective understanding about the theme that involves (a) a small group of students or (b) a small group of students and the teacher. Or, brief episodes of sharing and developing collective understandings occur sporadically throughout the lesson.

4. There are many sustained episodes of sharing and developing collective understandings about the theme in which many students participate.

5. The creation of and maintenance of collective understandings permeates the entire lesson. This could include the use of a common terminology and the careful negotiation of meanings. Almost everyone participates.

## Classroom Environment: Student Engagement

TO WHAT EXTENT IS STUDENTS' ATTENTION FOCUSED ON THE LESSON?

On-task behavior signals a serious psychological investment in class work. Ideally, it includes not only being attentive and doing the assigned work, but also showing enthusiasm for this work by participating verbally, contributing to group tasks, or helping peers. To score high on this scale, students must be deeply engaged in the lesson's actual content; that is, they cannot be reading non-school material or material related to another school subject, talking with peers about topics unrelated to the lesson, or doing yesterday's or tomorrow's homework (unless that homework is the topic at hand). **Note:** The ways in which students demonstrate engagement are culturally mediated and thus may vary by cultural background.

Off-task behavior is signaled by indications of boredom or a lack of effort by students. These include sleeping, daydreaming, talking to peers about non-class matters, making noise or otherwise disrupting the class. It is assumed these behaviors indicate that students are not taking seriously the substantive work of the class.

**Note:** Students can be on-task with regard to content that, to the observer, seems to be contrived, trivial, uninteresting, and boring. Alternatively, students might be off-task even with regard to content that strikes the observer as exciting, authentic, and interesting. Put such considerations aside when using this scale. The substance of the content is scored in other scales. The focus of this scale is on whether the class environment is or is not one of student engagement.

### Scale

1. Students are frequently off-task as evidenced by gross inattention or serious disruptions by many students (20%-50%); this is the central characteristic during much of the class.

2. Students appear lethargic and are only occasionally on-task carrying out assigned activities; for substantial portions of time, many students (20%-50%) are either clearly off-task or nominally on-task but not trying very hard.

3. Students are sporadically or episodically on-task; most students (50%-90%), some of the time (20%-50%), are engaged in class activities, but this engagement is uneven, only mildly enthusiastic or dependent on frequent prodding from the teacher.

4. On-task behavior is widespread; most students (50-%-90%), most of the time (50%-90%), are on-task pursuing the substance of the lesson; most students seem to be taking the work seriously and trying hard.

5. On-task behavior is the general rule throughout the class; almost all students (90% or more) are deeply involved, almost all of the time (90% or more), in pursuing the substance of the lesson.

## Assessment Strategies

TO WHAT EXTENT ARE ALTERNATIVE ASSESSMENT STRATEGIES USED TO INFORM BOTH TEACHING AND LEARNING?

Alternative assessment calls for instructional goals to be explicitly and holistically reinforced in teaching and learning. While alternative assessments can provide summative measures, their formative nature allows for ongoing feedback of learning progress to both the learner and teacher, and enhances learning and informs teaching. Usually, assessment instruments have been composed of selected response (e.g., multiple choice), short answer, and free response (essay) items. While it is possible to construct traditional assessment instruments that measure higher-order conceptual understanding, complex procedural knowledge and problem-solving; traditional assessment instruments have tended to measure less profound changes in mental models. Alternative assessment attempts to counteract the proclivity to measure rote learning. Alternative assessments are often performance-based, authentic tasks: products, performance tasks, or process-focused activities (e.g., interviews, think aloud, learning log, etc.). They are scored through rubrics, rating scales, checklists, and written or oral comments. Alternative assessment evaluation criteria guide student understanding of the aspects of quality. They serve as guides to instruction and evaluation. Evaluative criteria should be advertised within the classroom and available to students in advance. When scoring this scale, be sure to mention (a) nature of task, (b) procedures for monitoring progress, and (c) expectations/criteria. **Note:** Use data from pre-interview conversations and/or interview follow-ups to augment your classroom observations.

**Scale**

1. Students complete a traditional assessment instrument OR there is no evidence of formal alternative assessments within the classroom culture. Any informal observations by the instructor during the session were documented for, at most, a few students (10% or less).

2. Instructor uses an alternative assessment for summative evaluation of the unit supported by this lesson. There is no evidence of routine incorporation of evaluation expectations. Evaluation criteria are not posted on walls/ board or in students' materials. Almost all of the time (90% or more), neither teacher nor students engage in classroom discourse modeling evaluation criteria tied to instructional goals. Almost all of the time (90% or more), there is no evidence of instructor modeling formative assessment practices that students might use to structure peer review. If peer review occurs in the day's activities, most of the time (50%-90%) most of the students (50%-90%) are not on task.

3. Teacher has begun to incorporate formative evaluation. Some of the time (10%-20%) the instructor models discourse language containing evaluation criteria. Some (10%-20%) students' discourse follows the instructor's discourse model some of the time (10%-20%). Some students (10% - 20%) conduct peer review when assigned and seek feedback on ideas, procedures, and/or products. Instructor's formative assessment feedback, observations, and/or scorings are documented within instructor's records some all of the time (10% - 20%). Assessments connected with this lesson replicate authentic student activities or represent a sample of authentic student product some of the time (10% - 20%).

4. Teacher models scaffolding for ongoing formative assessment, incorporating evaluation expectations into her discourse most of the time (50%-90%). Grading criteria are prominently posted or advertised (e.g., rubrics, rating scales, checklists, at upper middle school and high school, inside student folders or unit packets). Instructor documents scoring of summative alternative assessments and may document a few formative observations. Assessments connected with this lesson replicate authentic student activities or represent a sample of authentic student product much of the time (20% - 50%).

5. Ongoing formative assessment is fully integrated within classroom practice. Instructor documents scoring of formative/summative alternative assessments. Evaluation criteria align with instructional goals and are routinely and prominently advertised (e.g., rubrics, rating scales, checklists, at upper middle school and high school, inside student folders or unit packets). Most of the time (50%-90%) the instructor models discourse language containing evaluation criteria. The discourse language of most (50%-90%) students follows the instructor's model most of the time (50%-90%). Many students (20% - 50%) conduct ongoing informal peer review and seek feedback on ideas, procedures, and/or products. Almost all (90% or more) of this discourse is focused and task-centered. Instructor's assessment feedback, observations, and/or scorings are documented almost all of the time (90% or more). Assessments connected with this lesson replicate authentic student activities or represent a sample of authentic student product almost all of the time (90% or more).