dents' thinking beyond their written (and drawn) responses to a task. The models themselves provide a context in which the students can clarify their thinking and refine their models in response to the critiques, to make more explicit claims to explain what they have observed. Thus, this activity focuses their attention on key explanatory issues (Reiser, 2004).

This example also illustrates the importance of engaging students in practices to help them develop understanding of disciplinary core ideas while also giving teachers information to guide instruction. In this case, the teacher's active probing of students' ideas demonstrates the way that formative assessment strategies can be effectively used as a part of instruction. The discussion of the models not only reveals the students' understanding about the phenomenon, but also allows the teacher to evaluate progress, uncover problematic issues, and help students construct and refine their models.

Example 5: Movement of Water

The committee chose this example to show how a teacher can monitor developing understanding in the course of a lesson. "Clicker technology"[7] is used to obtain individual student responses that inform teachers of what the students have learned from an activity and which are then the basis for structuring small-group discussions that address misunderstandings. This task assesses both understanding of a concept as it develops in the course of a lesson and students' discussion skills. The assessments are used formatively and are closely tied to classroom instruction.

In the previous example (Example 4), the teacher orchestrates a discussion in which students present alternative points of view and then come to consensus about a disciplinary core idea through the practice of argumentation. However, many teachers may find it challenging to track students' thinking while also promoting the development of understanding for the whole class. The example on the movement of air was developed as part of a program for helping teachers learn to lead students in "assessment conversations" (Duschl and Gitomer, 1997).[8] In the

[7]Clicker technology, also known as classroom response systems, allows students to use handheld clickers to respond to questions from a teacher. The responses are gathered by a central receiver and immediately tallied for the teacher—or the whole class—to see.

[8]This example is taken from the Contingent Pedagogies Project, which provides formative assessment tools for middle schools and supports teachers in integrating assessment activities into discussions for both small groups and entire classes. Of the students who responded to the task, 46 percent were Latino. For more information, see http://contingentpedagogies.org [October 2013].

task, middle school students engage in argumentation about disciplinary core ideas in earth science. As with the previous example, the formative assessment activity is more than just the initial question posed to students; it also includes the discussion that follows from student responses to it and teachers' decisions about what to do next, after she brings the discussion to a close.

In this activity, which also takes place in a single class session, the teacher structures a conversation about how the movement of water affects the deposition of surface and subsurface materials. The activity involves disciplinary core ideas (similar to Earth's systems in the NGSS) and engages students in practices, including modeling and constructing examples. It also requires students to reason about models of geosphere-hydrosphere interactions, which is an example of the crosscutting concept pertaining to systems and system models.[9]

Teachers use classroom clicker technology to pose multiple-choice questions that are carefully designed to elicit students' ideas related to the movement of water. These questions have been tested in classrooms, and the response choices reflect common student ideas, including those that are especially problematic. In the course of both small-group and whole-class discussions, students construct and challenge possible explanations of the process of deposition. If students have difficulty in developing explanations, teachers can guide students to activities designed to improve their understanding, such as interpreting models of the deposition of surface and subsurface materials.

When students begin this activity, they will just have completed a set of investigations of weathering, erosion, and deposition that are part of a curriculum on investigating Earth systems.[10] Students will have had the opportunity to build physical models of these phenomena and frame hypotheses about how water will move sediment using stream tables.[11] The teacher begins the formative assessment activity by projecting on a screen a question about the process of deposition designed to check students' understanding of the activities they have completed: see Figure 4-3 for a sample question. Students select their answers using clickers.

[9]The specific NGSS core idea addressed is similar to MS-ESS2.C: "How do the properties and movement of water shape Earth's surface and affect its systems?" The closest NGSS performance expectation is MS-ESS2-c: "Construct an explanation based on evidence for how geoscience processes have changed Earth's surface at varying time and spatial scales."

[10]This curriculum, for middle school students, was developed by the American Geosciences Institute. For more information, see http://www.agiweb.org/education/ies [July 2013].

[11]Stream tables are models of stream flows set up in large boxes filled with sedimentary material and tilted so that water can flow through.

FIGURE 4-3 Sample question for Example 5, "Movement of Water."
SOURCE: NASA/GSFC/JPL/LaRC, MISR Science Team (2013) and Los Angeles County Museum of Art (2013).

Pairs or small groups of students then discuss their reasoning and offer explanations for their choices to the whole class. Teachers help students begin the small-group discussions by asking why someone might select A, B, or C, implying that any of them could be a reasonable response. Teachers press students for their reasoning and invite them to compare their own reasoning to that of others, using specific discussion strategies (see Michaels and O'Connor, 2011; National Research Council, 2007). After discussing their reasoning, students again vote, using their clickers. In this example, the student responses recorded using the clicker technology are scorable. A separate set of assessments (not discussed here) produces scores to evaluate the efficacy of the project as a whole.

The program materials include a set of "contingent activities" for teachers to use if students have difficulty meeting a performance expectation related to an investigation. Teachers use students' responses to decide which contingent activities are needed, and thus they use the activity as an informal formative assessment. In these activities, students might be asked to interpret models, construct explanations, and make predictions using those models as a way to deepen their understanding of Earth systems. In this example about the movement of air, students who are having difficulty understanding can view an animation of deposition and then make a prediction about a pattern they might expect to find at the mouth of a river where sediment is being deposited.

The aim of this kind of assessment activity is to guide teachers in using assessment techniques to improve student learning outcomes.[12] The techniques used in this example demonstrate a means of rapidly assessing how well students have mastered a complex combination of practices and concepts in the midst of a lesson, which allows teachers to immediately address areas students do not understand well. The contingent activities that provide alternative ways for students to master the core ideas (by engaging in particular practices) are an integral component of the formative assessment process.

Example 6: Biodiversity in the Schoolyard

The committee chose this example to show the use of multiple interrelated tasks to assess a disciplinary core idea, biodiversity, with multiple science practices. As part of an extended unit, students complete four assessment tasks. The first three serve formative purposes and are designed to function close to instruction, informing the teacher about how well students have learned key concepts and mastered practices. The last assessment task serves a summative purpose, as an end-of-unit test, and is an example of a proximal assessment. The tasks address concepts related to biodiversity and science practices in an integrated fashion.

This set of four assessment tasks was designed to provide evidence of 5th-grade students' developing proficiency with a body of knowledge that blends a disciplinary core idea (biodiversity; LS4 in the NGSS; see Box 2-1 in Chapter 2) and a crosscutting concept (patterns) with three different practices: planning and carrying out investigations, analyzing and interpreting data, and constructing explanations (see Songer et al., 2009; Gotwals and Songer, 2013). These tasks, developed by researchers as part of an examination of the development of complex reasoning, are intended for use in an extended unit of study.[13]

[12]A quasi-experimental study compared the learning gains for students in classes that used the approach of the Contingent Pedagogies Project with gains for students in other classes in the same school district that used the same curriculum but not that approach. The students whose teachers used the Contingent Pedagogies Project demonstrated greater proficiency in earth science objectives than did students in classrooms in which teachers only had access to the regular curriculum materials (Penuel et al., 2012).

[13]The tasks were given to a sample of 6th-grade students in the Detroit Public School System, the majority of whom were racial/ethnic minority students (for details, see Songer et al., 2009).